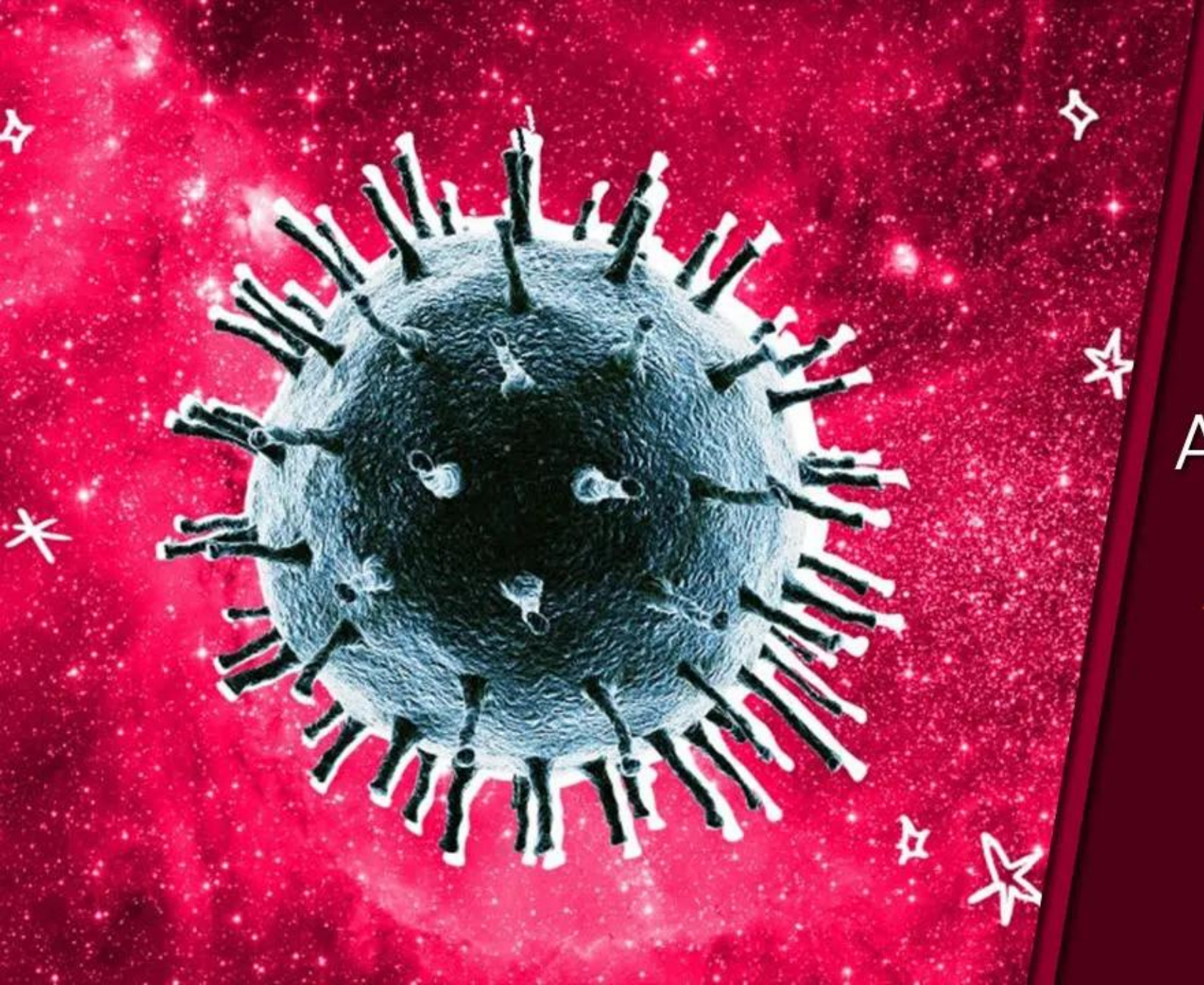

Astrovirology

NASA Workshop Without Walls

1
00:00:06,389 --> 00:00:02,550
um greetings everyone thanks for joining

2
00:00:09,910 --> 00:00:06,399
us for the second of our two day

3
00:00:11,669 --> 00:00:09,920
uh how two half day uh adventures with

4
00:00:12,870 --> 00:00:11,679
astrovirology

5
00:00:15,430 --> 00:00:12,880
we were very pleased with the

6
00:00:17,670 --> 00:00:15,440
participation yesterday we had

7
00:00:20,950 --> 00:00:17,680
in total uh more than a hundred people

8
00:00:23,910 --> 00:00:20,960
join us um throughout the

9
00:00:25,830 --> 00:00:23,920
four hour event and so

10
00:00:28,390 --> 00:00:25,840
no doubt today we'll have people coming

11
00:00:29,750 --> 00:00:28,400
on as they get out of lunch and meetings

12
00:00:31,589 --> 00:00:29,760
and teaching and all those things that

13
00:00:33,590 --> 00:00:31,599

people do

14
00:00:35,830 --> 00:00:33,600
we gave a somewhat lengthier

15
00:00:37,750 --> 00:00:35,840
introduction yesterday but i'm going to

16
00:00:39,030 --> 00:00:37,760
truncate it make it very short today

17
00:00:40,549 --> 00:00:39,040
because we want to get on to our

18
00:00:42,630 --> 00:00:40,559
speakers

19
00:00:44,389 --> 00:00:42,640
this workshop is sponsored by the nasa

20
00:00:47,510 --> 00:00:44,399
astrobiology institute

21
00:00:50,150 --> 00:00:47,520
in our last few months of existence and

22
00:00:52,630 --> 00:00:50,160
astrovirology is something that

23
00:00:55,510 --> 00:00:52,640
requires a great deal more development

24
00:00:57,510 --> 00:00:55,520
within the astrobiology community and we

25
00:00:58,869 --> 00:00:57,520
think it is a very valuable direction to

26
00:01:00,709 --> 00:00:58,879
go down

27
00:01:02,389 --> 00:01:00,719
and so there are two things planned for

28
00:01:04,549 --> 00:01:02,399
this uh

29
00:01:08,310 --> 00:01:04,559
this group as it goes forward into the

30
00:01:12,149 --> 00:01:08,320
future and uh one is a journal article

31
00:01:14,950 --> 00:01:12,159
that is essentially uh building on uh

32
00:01:16,710 --> 00:01:14,960
the article that um ken steadman and

33
00:01:20,149 --> 00:01:16,720
colleagues published i think it was last

34
00:01:22,070 --> 00:01:20,159
year and um uh see if we can you know

35
00:01:25,830 --> 00:01:22,080
pull this uh community which is

36
00:01:27,749 --> 00:01:25,840
certainly larger than i feared and

37
00:01:30,870 --> 00:01:27,759
making me very hopeful for the future of

38
00:01:33,030 --> 00:01:30,880

this into a journal article

39
00:01:35,749 --> 00:01:33,040
for submission as soon as reasonably

40
00:01:37,510 --> 00:01:35,759
possible to the journal astrobiology

41
00:01:39,109 --> 00:01:37,520
we've been in contact with sherry cady

42
00:01:41,190 --> 00:01:39,119
the executive

43
00:01:43,270 --> 00:01:41,200
editor of that in order to facilitate

44
00:01:45,830 --> 00:01:43,280
that and so she knows that this will be

45
00:01:48,550 --> 00:01:45,840
coming and then the ultimate idea of

46
00:01:51,910 --> 00:01:48,560
course is to distill the important uh

47
00:01:53,429 --> 00:01:51,920
portions uh the high-level messages from

48
00:01:54,830 --> 00:01:53,439
the community

49
00:01:57,270 --> 00:01:54,840
on the topic of

50
00:01:58,950 --> 00:01:57,280
astrovirology and to be able to put that

51
00:02:01,109 --> 00:01:58,960
in as a white paper

52
00:02:02,789 --> 00:02:01,119
for the upcoming planetary decadal

53
00:02:04,550 --> 00:02:02,799
process conducted by the national

54
00:02:07,990 --> 00:02:04,560
academy sciences

55
00:02:10,949 --> 00:02:08,000
our latest intel on that is that um

56
00:02:12,949 --> 00:02:10,959
solicitation for those papers will begin

57
00:02:15,670 --> 00:02:12,959
uh in approximately the february time

58
00:02:18,070 --> 00:02:15,680
frame we don't know when um and so that

59
00:02:19,910 --> 00:02:18,080
is still a bit nebulous but that is what

60
00:02:21,589 --> 00:02:19,920
we're aiming for and then of course

61
00:02:23,750 --> 00:02:21,599
those papers will be read and

62
00:02:25,830 --> 00:02:23,760
adjudicated by the panels that are

63
00:02:29,030 --> 00:02:25,840

constituted as part of that national

64
00:02:31,270 --> 00:02:29,040
academies activity and sometime in the

65
00:02:33,830 --> 00:02:31,280
following year it will

66
00:02:35,830 --> 00:02:33,840
come out as the decadal survey of

67
00:02:38,229 --> 00:02:35,840
planetary science and that has

68
00:02:41,110 --> 00:02:38,239
significant consequences for

69
00:02:43,589 --> 00:02:41,120
what nasa is willing to fund in terms of

70
00:02:45,030 --> 00:02:43,599
basic science in terms of missions in

71
00:02:47,350 --> 00:02:45,040
terms of

72
00:02:50,470 --> 00:02:47,360
how all of these topics and technologies

73
00:02:52,869 --> 00:02:50,480
fit in with the nasa goals and

74
00:02:55,190 --> 00:02:52,879
activities so uh with that i will go

75
00:02:58,070 --> 00:02:55,200
ahead and turn that over to

76
00:03:00,790 --> 00:02:58,080
ken stedman i guess he's gonna run it

77
00:03:02,869 --> 00:03:00,800
and uh go for it can

78
00:03:06,790 --> 00:03:02,879
okay thanks again penny and thanks to

79
00:03:08,710 --> 00:03:06,800
everybody both at nii central and all

80
00:03:10,869 --> 00:03:08,720
the rest of you from throughout the

81
00:03:12,390 --> 00:03:10,879
world as far as we can tell

82
00:03:14,630 --> 00:03:12,400
i'm not sure how many of you know the

83
00:03:16,790 --> 00:03:14,640
reasoning behind splitting this into two

84
00:03:18,390 --> 00:03:16,800
half-day sessions part of the idea was

85
00:03:19,830 --> 00:03:18,400
to be able to accommodate as many people

86
00:03:22,149 --> 00:03:19,840
in as many different time zones as

87
00:03:24,149 --> 00:03:22,159
possible and so that was one of the

88
00:03:26,789 --> 00:03:24,159

big impetus for splitting it up in terms

89
00:03:30,070 --> 00:03:26,799
of two half days but given yesterday i'm

90
00:03:33,350 --> 00:03:30,080
really happy that i had about

91
00:03:35,270 --> 00:03:33,360
uh a day and a half to digest some of

92
00:03:37,670 --> 00:03:35,280
the stuff that we went over yesterday it

93
00:03:41,190 --> 00:03:37,680
was awesome it was great we had huge

94
00:03:43,589 --> 00:03:41,200
amounts of information um and i will try

95
00:03:45,430 --> 00:03:43,599
to mention a little bit of that at the

96
00:03:46,390 --> 00:03:45,440
very end and try and collect a few

97
00:03:48,470 --> 00:03:46,400
pieces

98
00:03:50,630 --> 00:03:48,480
as we move along to the rest of this i

99
00:03:53,750 --> 00:03:50,640
wanted to emphasize what any mentioned

100
00:03:56,390 --> 00:03:53,760
about a paper and coming together for a

101
00:03:58,070 --> 00:03:56,400
white paper later on i know gary sent

102
00:04:00,710 --> 00:03:58,080
out a

103
00:04:03,350 --> 00:04:00,720
it was a chat or a tweet or something

104
00:04:05,750 --> 00:04:03,360
one of these things in terms of people

105
00:04:06,630 --> 00:04:05,760
who are interested please get in touch

106
00:04:09,830 --> 00:04:06,640
with

107
00:04:11,429 --> 00:04:09,840
gary or i or both and we will do a

108
00:04:12,630 --> 00:04:11,439
follow-up to

109
00:04:15,830 --> 00:04:12,640
this meeting

110
00:04:19,590 --> 00:04:15,840
probably early next week in terms of

111
00:04:22,710 --> 00:04:19,600
trying to organize exactly who what

112
00:04:24,070 --> 00:04:22,720
where when etc as far as papers are

113
00:04:27,189 --> 00:04:24,080

concerned

114
00:04:29,830 --> 00:04:27,199
for the next step i think everybody has

115
00:04:33,110 --> 00:04:29,840
noticed that there's also a chat window

116
00:04:36,310 --> 00:04:33,120
that seems to be very active we will

117
00:04:37,510 --> 00:04:36,320
collect all of those messages and post

118
00:04:40,710 --> 00:04:37,520
them to

119
00:04:43,590 --> 00:04:40,720
the nai website after

120
00:04:45,590 --> 00:04:43,600
this whole meeting is done

121
00:04:48,710 --> 00:04:45,600
maybe tomorrow but probably again early

122
00:04:51,830 --> 00:04:48,720
next week the same thing is true for the

123
00:04:52,629 --> 00:04:51,840
recordings of the presentations and

124
00:04:53,909 --> 00:04:52,639
the

125
00:04:57,350 --> 00:04:53,919
actual

126
00:04:59,110 --> 00:04:57,360
visuals that all of the presenters have

127
00:05:00,710 --> 00:04:59,120
shared with us i know there are a couple

128
00:05:03,029 --> 00:05:00,720
of things that people are not ready to

129
00:05:04,629 --> 00:05:03,039
share which is totally fine um we've

130
00:05:06,790 --> 00:05:04,639
just asked that those people send those

131
00:05:09,590 --> 00:05:06,800
to us and then we will do our best to go

132
00:05:11,830 --> 00:05:09,600
ahead and get those posted

133
00:05:15,189 --> 00:05:11,840
one other

134
00:05:17,749 --> 00:05:15,199
request and that is um for people to

135
00:05:20,310 --> 00:05:17,759
ask lots of questions afterwards um each

136
00:05:22,469 --> 00:05:20,320
of the talks um the first talk we're

137
00:05:23,909 --> 00:05:22,479
going to have from evelyn andreessen's

138
00:05:26,310 --> 00:05:23,919

is actually set up to be a 30-minute

139
00:05:27,590 --> 00:05:26,320
talk other than the 20-minute talks that

140
00:05:29,110 --> 00:05:27,600
we have later

141
00:05:31,830 --> 00:05:29,120
i'm also with

142
00:05:33,110 --> 00:05:31,840
10 minutes for questions and again as i

143
00:05:34,870 --> 00:05:33,120
mentioned yesterday particularly for the

144
00:05:35,830 --> 00:05:34,880
people who are new today

145
00:05:36,550 --> 00:05:35,840
if

146
00:05:39,590 --> 00:05:36,560
we

147
00:05:42,390 --> 00:05:39,600
have too many questions you will get cut

148
00:05:44,230 --> 00:05:42,400
off um so that we can stay on time

149
00:05:46,870 --> 00:05:44,240
those can always continue offline or in

150
00:05:49,110 --> 00:05:46,880
the chat box um or if we've got some

151
00:05:52,150 --> 00:05:49,120
extra time that gives us a little break

152
00:05:54,550 --> 00:05:52,160
so we can go and refill our coffee or

153
00:05:56,629 --> 00:05:54,560
have a physiology break or whatever

154
00:05:59,749 --> 00:05:56,639
happens to be necessary

155
00:06:01,029 --> 00:05:59,759
at that particular time i am just again

156
00:06:03,029 --> 00:06:01,039
quick comment as far as yesterday is

157
00:06:06,309 --> 00:06:03,039
concerned i thought that it went

158
00:06:09,029 --> 00:06:06,319
extremely well um thanks again to the

159
00:06:09,830 --> 00:06:09,039
people at nai central particularly marco

160
00:06:13,430 --> 00:06:09,840
for

161
00:06:15,590 --> 00:06:13,440
making sure that everything worked and

162
00:06:18,309 --> 00:06:15,600
again i was very pleasantly surprised

163
00:06:20,150 --> 00:06:18,319

that it worked as well as it did um if

164
00:06:23,430 --> 00:06:20,160
people have questions particularly the

165
00:06:26,150 --> 00:06:23,440
presenters in terms of moving things

166
00:06:29,110 --> 00:06:26,160
through advancing your slides etc the

167
00:06:31,189 --> 00:06:29,120
key is the little share button down at

168
00:06:32,309 --> 00:06:31,199
the bottom of your zoom screen

169
00:06:34,469 --> 00:06:32,319
and then you should be able to share

170
00:06:36,150 --> 00:06:34,479
your screen with everyone if you don't

171
00:06:38,070 --> 00:06:36,160
want to see yourself while you're

172
00:06:41,590 --> 00:06:38,080
talking because we all just love to look

173
00:06:43,430 --> 00:06:41,600
at ourselves um there's a also a button

174
00:06:45,270 --> 00:06:43,440
which you can press to say to mute your

175
00:06:48,309 --> 00:06:45,280
own video so

176
00:06:50,790 --> 00:06:48,319
i know some people appreciate that

177
00:06:53,510 --> 00:06:50,800
that's all that i had for

178
00:06:56,350 --> 00:06:53,520
this section um we're mostly going to be

179
00:06:59,189 --> 00:06:56,360
talking about virus ecology

180
00:07:01,510 --> 00:06:59,199
exobiology the three organizers are

181
00:07:02,629 --> 00:07:01,520
going to get to regale you with what

182
00:07:05,029 --> 00:07:02,639
they're doing

183
00:07:07,110 --> 00:07:05,039
um and i'll try and wrap all of it up at

184
00:07:09,029 --> 00:07:07,120
the very end do we have any more

185
00:07:10,309 --> 00:07:09,039
questions for anyone at this point

186
00:07:15,830 --> 00:07:10,319
particularly for any of the people who

187
00:07:19,589 --> 00:07:17,589
no we're all good

188
00:07:21,510 --> 00:07:19,599

gary and i just got our presentations in

189
00:07:23,270 --> 00:07:21,520
today so we told everyone to get them in

190
00:07:25,189 --> 00:07:23,280
earlier and we managed to be the the

191
00:07:29,270 --> 00:07:25,199
latest ones i think catherine was way

192
00:07:29,280 --> 00:07:32,469
no

193
00:07:36,070 --> 00:07:34,390
catherine who

194
00:07:37,430 --> 00:07:36,080
just got sent in

195
00:07:38,870 --> 00:07:37,440
i guess i guess that's the prerogative

196
00:07:41,589 --> 00:07:38,880
of being an organizer right you can

197
00:07:43,589 --> 00:07:41,599
break your own rules

198
00:07:44,710 --> 00:07:43,599
we're not specifically that right

199
00:07:45,830 --> 00:07:44,720
and i wasn't supposed to admit it you

200
00:07:47,830 --> 00:07:45,840
see that's what happens when if gary

201
00:07:49,990 --> 00:07:47,840
when you let me do the introductions

202
00:07:51,749 --> 00:07:50,000
okay so um is

203
00:07:54,390 --> 00:07:51,759
um well just so real quickly again so

204
00:07:55,510 --> 00:07:54,400
our our speakers this afternoon this

205
00:07:57,110 --> 00:07:55,520
afternoon

206
00:07:59,589 --> 00:07:57,120
california time

207
00:08:02,150 --> 00:07:59,599
at oregon time evelyn andreessen's

208
00:08:03,430 --> 00:08:02,160
followed by nigel goldenfeld followed by

209
00:08:04,629 --> 00:08:03,440
simo

210
00:08:07,510 --> 00:08:04,639
followed by

211
00:08:09,270 --> 00:08:07,520
gary the guy in the headless fade shirt

212
00:08:11,270 --> 00:08:09,280
there

213
00:08:14,550 --> 00:08:11,280

we switched our times oh you switch your

214
00:08:16,469 --> 00:08:14,560
times okay so yeah so kathy's gonna go

215
00:08:19,430 --> 00:08:16,479
kathy first then

216
00:08:21,510 --> 00:08:19,440
uh gary and then i will try and

217
00:08:23,990 --> 00:08:21,520
wrap things up and talk a little bit

218
00:08:27,510 --> 00:08:24,000
about what we're doing but mostly about

219
00:08:28,790 --> 00:08:27,520
sort of plans for next stages i'm also

220
00:08:30,150 --> 00:08:28,800
going to try and give a little bit of

221
00:08:31,670 --> 00:08:30,160
background some of the things i noticed

222
00:08:34,230 --> 00:08:31,680
from the

223
00:08:35,190 --> 00:08:34,240
chats um not everybody is on the same

224
00:08:37,029 --> 00:08:35,200
page

225
00:08:40,149 --> 00:08:37,039
some of us have been doing viruses for

226
00:08:43,750 --> 00:08:40,159
multiple decades some of us for

227
00:08:46,230 --> 00:08:43,760
yesterday morning um so i'll give a

228
00:08:47,430 --> 00:08:46,240
little bit of an introduction in terms

229
00:08:50,790 --> 00:08:47,440
of some of the things we're thinking

230
00:08:52,550 --> 00:08:50,800
about from a virus point of view and

231
00:08:54,470 --> 00:08:52,560
also as a quick reminder again before we

232
00:08:55,269 --> 00:08:54,480
get evelyn started

233
00:08:57,350 --> 00:08:55,279
is

234
00:08:59,269 --> 00:08:57,360
that there are some great references

235
00:09:00,470 --> 00:08:59,279
thanks gary i think it was it kathy for

236
00:09:02,550 --> 00:09:00,480
putting it together

237
00:09:05,750 --> 00:09:02,560
on the website for

238
00:09:07,269 --> 00:09:05,760

a number of the articles which are

239
00:09:09,269 --> 00:09:07,279
really nice background with the

240
00:09:10,790 --> 00:09:09,279
exception of course of mine

241
00:09:12,630 --> 00:09:10,800
which you can then go and get some more

242
00:09:15,269 --> 00:09:12,640
details on what's going on there and if

243
00:09:17,590 --> 00:09:15,279
people have trouble accessing any of

244
00:09:20,230 --> 00:09:17,600
those i'm happy to

245
00:09:21,590 --> 00:09:20,240
share um that article in particular and

246
00:09:23,990 --> 00:09:21,600
i'm sure other people have access to

247
00:09:28,790 --> 00:09:26,870
so without further ado um

248
00:09:30,310 --> 00:09:28,800
evelyn if you're ready to go i think

249
00:09:33,590 --> 00:09:30,320
that marco will be able to switch you

250
00:09:34,630 --> 00:09:33,600
over and we should be good to go

251
00:09:35,829 --> 00:09:34,640
okay

252
00:09:40,790 --> 00:09:35,839
um

253
00:09:47,350 --> 00:09:41,580
and

254
00:09:50,389 --> 00:09:49,670
all right

255
00:09:52,870 --> 00:09:50,399
so

256
00:09:56,230 --> 00:09:52,880
thank you very much for inviting me um

257
00:09:58,230 --> 00:09:56,240
to speak here about virus taxonomy

258
00:09:59,590 --> 00:09:58,240
um and i hope i will do it justice

259
00:10:04,630 --> 00:09:59,600
because i've seen that there's some

260
00:10:10,230 --> 00:10:06,949
basically i'm i'll try to tell you uh

261
00:10:13,670 --> 00:10:10,240
what is virus taxonomy and and why

262
00:10:14,710 --> 00:10:13,680
should you care um feel free to uh tweet

263
00:10:16,870 --> 00:10:14,720

me

264
00:10:18,790 --> 00:10:16,880
uh and i wanted to quickly share our

265
00:10:21,350 --> 00:10:18,800
beautiful new building i work in the

266
00:10:22,870 --> 00:10:21,360
quadrant institute in norwich in the uk

267
00:10:25,190 --> 00:10:22,880
and we just moved into a new building

268
00:10:27,829 --> 00:10:25,200
and everybody's loving it so just i

269
00:10:30,389 --> 00:10:27,839
wanted to share that to you guys

270
00:10:33,190 --> 00:10:30,399
so because i'm the first speaker of the

271
00:10:35,750 --> 00:10:33,200
day i thought maybe i'll just start very

272
00:10:38,790 --> 00:10:35,760
general and very philosophical with what

273
00:10:40,710 --> 00:10:38,800
is a virus then i'll try to explain to

274
00:10:43,190 --> 00:10:40,720
you what taxonomy is

275
00:10:45,030 --> 00:10:43,200
and then i'll finish off with the now

276
00:10:46,550 --> 00:10:45,040
what

277
00:10:47,829 --> 00:10:46,560
so when you look at

278
00:10:49,509 --> 00:10:47,839
viruses

279
00:10:51,110 --> 00:10:49,519
a lot of virologists are very much

280
00:10:53,670 --> 00:10:51,120
specialists

281
00:10:56,069 --> 00:10:53,680
and if you look at the first picture of

282
00:10:58,150 --> 00:10:56,079
poliovirus a lot of people think of a

283
00:11:01,030 --> 00:10:58,160
virus system something small something

284
00:11:03,350 --> 00:11:01,040
simple something roundish

285
00:11:05,590 --> 00:11:03,360
but then there's another human virus

286
00:11:08,710 --> 00:11:05,600
called ebola virus which is and looks

287
00:11:10,949 --> 00:11:08,720
nothing like it is way bigger is

288
00:11:12,949 --> 00:11:10,959

stringy ribbon-like

289
00:11:14,790 --> 00:11:12,959
so the first definition of a virus

290
00:11:16,630 --> 00:11:14,800
something small and round

291
00:11:18,949 --> 00:11:16,640
already falls away

292
00:11:19,670 --> 00:11:18,959
if you then look at bacteriophages which

293
00:11:21,430 --> 00:11:19,680
is

294
00:11:23,430 --> 00:11:21,440
which are my personal favorites and if

295
00:11:25,910 --> 00:11:23,440
you look at gary's t-shirt obviously

296
00:11:28,470 --> 00:11:25,920
they're also his personal favorite um in

297
00:11:30,630 --> 00:11:28,480
the middle is a feature that was that

298
00:11:33,990 --> 00:11:30,640
that my student and i isolated during my

299
00:11:35,670 --> 00:11:34,000
uh phd which we call limestone

300
00:11:37,110 --> 00:11:35,680
um but

301
00:11:39,350 --> 00:11:37,120
and of course i wanted to put that in

302
00:11:41,190 --> 00:11:39,360
because um it looks like a lunar lander

303
00:11:45,350 --> 00:11:41,200
and we're talking about space and

304
00:11:49,829 --> 00:11:47,910
there's loads more

305
00:11:51,590 --> 00:11:49,839
i've put in an example of an archaeal

306
00:11:53,990 --> 00:11:51,600
virus and it's the lipid tricks virus

307
00:11:56,550 --> 00:11:54,000
and it looks pretty awesome and it has a

308
00:11:59,030 --> 00:11:56,560
little hook and it looks nothing like

309
00:12:01,030 --> 00:11:59,040
anything other than what you had in mind

310
00:12:02,629 --> 00:12:01,040
of what a virus should look like

311
00:12:04,470 --> 00:12:02,639
and it doesn't end there because the

312
00:12:06,470 --> 00:12:04,480
last one i have in my image is a mimi

313
00:12:08,870 --> 00:12:06,480

virus and that mimi actually stands for

314
00:12:11,430 --> 00:12:08,880
microbe mimicking and it's because this

315
00:12:13,509 --> 00:12:11,440
is a virus that can be looked at through

316
00:12:16,150 --> 00:12:13,519
a light microscope instead of an

317
00:12:18,069 --> 00:12:16,160
electron microscope so

318
00:12:20,389 --> 00:12:18,079
a lot of people have this an idea of

319
00:12:22,230 --> 00:12:20,399
what a virus is and

320
00:12:24,949 --> 00:12:22,240
each time a new virus gets discovered

321
00:12:27,110 --> 00:12:24,959
that idea kind of needs to shift so what

322
00:12:28,949 --> 00:12:27,120
where we're at now is that a virus is a

323
00:12:30,870 --> 00:12:28,959
biological entity

324
00:12:35,030 --> 00:12:30,880
um it

325
00:12:36,949 --> 00:12:35,040
goes into a cell infects it um sells of

326
00:12:38,470 --> 00:12:36,959
all the means of life and it needs it to

327
00:12:39,750 --> 00:12:38,480
replicate

328
00:12:43,190 --> 00:12:39,760
and

329
00:12:45,110 --> 00:12:43,200
it's a bit of a vague description but

330
00:12:47,110 --> 00:12:45,120
it's it's the best we can do at the time

331
00:12:48,710 --> 00:12:47,120
because each time a new virus comes

332
00:12:50,389 --> 00:12:48,720
along it kind of shifts our paradigm a

333
00:12:53,030 --> 00:12:50,399
little

334
00:12:55,030 --> 00:12:53,040
but in its easiest form

335
00:12:56,629 --> 00:12:55,040
of virion the

336
00:12:59,910 --> 00:12:56,639
replicating free

337
00:13:00,949 --> 00:12:59,920
standing part of a virus is just nucleic

338
00:13:03,750 --> 00:13:00,959

acid

339
00:13:05,829 --> 00:13:03,760
with a shell around it and if you look

340
00:13:08,629 --> 00:13:05,839
at the ones on the left which is uh

341
00:13:10,389 --> 00:13:08,639
poliovirus it can be very simple and if

342
00:13:13,910 --> 00:13:10,399
you look at the mimi virus again it can

343
00:13:16,949 --> 00:13:13,920
get really complicated um with

344
00:13:18,949 --> 00:13:16,959
multiple layers of capsid um and these

345
00:13:21,269 --> 00:13:18,959
weird spikes and

346
00:13:23,509 --> 00:13:21,279
and a port that opens and again it can

347
00:13:24,870 --> 00:13:23,519
get really complicated so if we're

348
00:13:27,430 --> 00:13:24,880
looking at

349
00:13:29,670 --> 00:13:27,440
viruses outside of earth

350
00:13:32,870 --> 00:13:29,680
we kind of need to keep an open mind

351
00:13:35,350 --> 00:13:32,880
about what they can look like

352
00:13:37,590 --> 00:13:35,360
i'm i already said

353
00:13:39,750 --> 00:13:37,600
small is not a

354
00:13:41,509 --> 00:13:39,760
good definition for a virus anymore

355
00:13:43,750 --> 00:13:41,519
because if you look at human viruses

356
00:13:45,350 --> 00:13:43,760
alone they go from and this is to scale

357
00:13:47,990 --> 00:13:45,360
amongst each other you go from really

358
00:13:50,470 --> 00:13:48,000
tiny and a part of a virus to actually

359
00:13:53,189 --> 00:13:50,480
quite large for ebola virus

360
00:13:55,430 --> 00:13:53,199
if you were to compare them through to

361
00:13:59,829 --> 00:13:55,440
a bacteria that would be

362
00:14:00,949 --> 00:13:59,839
like this um the smallest bacteria would

363
00:14:02,949 --> 00:14:00,959

eclipse

364
00:14:04,870 --> 00:14:02,959
all of the viruses here

365
00:14:06,389 --> 00:14:04,880
and if you were to compare that to a

366
00:14:07,910 --> 00:14:06,399
human hair

367
00:14:10,470 --> 00:14:07,920
it would just look like this because the

368
00:14:11,750 --> 00:14:10,480
human hair would cover everything

369
00:14:15,509 --> 00:14:11,760
so not only

370
00:14:16,470 --> 00:14:15,519
what the viruses look like is vastly

371
00:14:18,550 --> 00:14:16,480
diverse

372
00:14:19,910 --> 00:14:18,560
also their genetic material is

373
00:14:22,310 --> 00:14:19,920
very diverse

374
00:14:24,629 --> 00:14:22,320
so for example you have

375
00:14:27,350 --> 00:14:24,639
viruses that are made up of dna and

376
00:14:28,550 --> 00:14:27,360
viruses are made up of rna but not just

377
00:14:29,750 --> 00:14:28,560
double-stranded dna you have

378
00:14:31,430 --> 00:14:29,760
single-stranded dna you have

379
00:14:33,910 --> 00:14:31,440
double-stranded rna you have

380
00:14:35,829 --> 00:14:33,920
single-stranded rna in a positive sense

381
00:14:38,710 --> 00:14:35,839
in a negative sense and then you even

382
00:14:41,269 --> 00:14:38,720
have viruses that have dna that reverse

383
00:14:42,629 --> 00:14:41,279
transcribe to dna to rna and into dna

384
00:14:46,230 --> 00:14:42,639
again

385
00:14:48,470 --> 00:14:46,240
um so if we look at

386
00:14:49,430 --> 00:14:48,480
what a virus might look like someplace

387
00:14:51,670 --> 00:14:49,440
else

388
00:14:54,550 --> 00:14:51,680

what we want is actually a molecule that

389
00:14:55,990 --> 00:14:54,560
can get that can get carry genetic

390
00:14:58,069 --> 00:14:56,000
information

391
00:15:02,710 --> 00:14:58,079
that can replicate and that's packaged

392
00:15:07,110 --> 00:15:04,550
one of the big questions or one of the

393
00:15:09,110 --> 00:15:07,120
big questions for some people is are

394
00:15:10,470 --> 00:15:09,120
viruses alive

395
00:15:14,230 --> 00:15:10,480
so

396
00:15:16,230 --> 00:15:14,240
people publish a paper say saying that

397
00:15:18,069 --> 00:15:16,240
viruses should not be in the tree of

398
00:15:20,069 --> 00:15:18,079
life and of course then other people

399
00:15:21,750 --> 00:15:20,079
publish another paper saying no viruses

400
00:15:24,710 --> 00:15:21,760
should be in the tree of life and

401
00:15:26,870 --> 00:15:24,720
there's people who are more nuanced and

402
00:15:28,710 --> 00:15:26,880
there's people who say why are you

403
00:15:31,030 --> 00:15:28,720
having this conversation

404
00:15:32,629 --> 00:15:31,040
and i'm i'm kind of on the train of the

405
00:15:34,790 --> 00:15:32,639
people who are asking why are you having

406
00:15:36,550 --> 00:15:34,800
this conversation because it's all a

407
00:15:38,870 --> 00:15:36,560
matter of semantics and and does it

408
00:15:43,269 --> 00:15:38,880
really matter the viruses don't care

409
00:15:46,629 --> 00:15:44,790
for those of you who don't know about

410
00:15:48,389 --> 00:15:46,639
the tree of life i'm sure

411
00:15:50,389 --> 00:15:48,399
all of you do but then it's in its

412
00:15:52,629 --> 00:15:50,399
simplest form it describes

413
00:15:55,990 --> 00:15:52,639

bacteria archaea and eukarya as the

414
00:15:59,350 --> 00:15:57,430
and you can see that

415
00:16:00,629 --> 00:15:59,360
people and animals and fungi and plants

416
00:16:03,670 --> 00:16:00,639
are actually quite close together on

417
00:16:06,550 --> 00:16:03,680
that tree if you expanded

418
00:16:08,949 --> 00:16:06,560
to look at all the potential

419
00:16:10,949 --> 00:16:08,959
types of bacteria and types of archaea

420
00:16:14,389 --> 00:16:10,959
you can see that the eukaryotes kind of

421
00:16:16,629 --> 00:16:14,399
get dwarfed by the diversity of

422
00:16:19,030 --> 00:16:16,639
unicellular life

423
00:16:21,350 --> 00:16:19,040
if you look at even further and you try

424
00:16:23,110 --> 00:16:21,360
to add all the viruses on this tree of

425
00:16:26,470 --> 00:16:23,120
life you can see that the virus

426
00:16:29,189 --> 00:16:26,480
diversity actually even eclipses that of

427
00:16:31,670 --> 00:16:29,199
all cellular diversity

428
00:16:33,670 --> 00:16:31,680
meaning that

429
00:16:35,670 --> 00:16:33,680
when you do taxonomy

430
00:16:38,310 --> 00:16:35,680
stuff gets really complicated really

431
00:16:41,829 --> 00:16:39,670
so

432
00:16:44,629 --> 00:16:41,839
when i say what about taxonomy i might

433
00:16:46,790 --> 00:16:44,639
want to start with what is taxonomy and

434
00:16:48,829 --> 00:16:46,800
as a good researcher the first thing you

435
00:16:51,910 --> 00:16:48,839
do is you go to

436
00:16:53,269 --> 00:16:51,920
wikipedia and wikipedia

437
00:16:55,350 --> 00:16:53,279
tells you not to confuse it with

438
00:16:57,990 --> 00:16:55,360

taxidermy so any anybody of you are

439
00:16:59,670 --> 00:16:58,000
confused and are actually at the wrong

440
00:17:01,910 --> 00:16:59,680
place please please leave the

441
00:17:03,910 --> 00:17:01,920
conversation now and we'll go on with

442
00:17:05,590 --> 00:17:03,920
taxonomy which is the practice and

443
00:17:07,669 --> 00:17:05,600
science of classification of things or

444
00:17:09,909 --> 00:17:07,679
concepts including the principles that

445
00:17:11,990 --> 00:17:09,919
underlie such concepts

446
00:17:14,710 --> 00:17:12,000
now i always like to share some some

447
00:17:17,350 --> 00:17:14,720
funny quotes about that because

448
00:17:19,510 --> 00:17:17,360
a person called cowan in the 70s told us

449
00:17:21,429 --> 00:17:19,520
that taxonomy is written by taxonomists

450
00:17:23,829 --> 00:17:21,439
for taxonomists and in this form the

451
00:17:26,309 --> 00:17:23,839
subject is so dull that few if any

452
00:17:27,829 --> 00:17:26,319
non-taxonomists are tempted to read it

453
00:17:29,510 --> 00:17:27,839
and it's the most subjective branch of

454
00:17:32,549 --> 00:17:29,520
any biological discipline and in many

455
00:17:33,990 --> 00:17:32,559
ways more of an art than a science and

456
00:17:35,590 --> 00:17:34,000
us taxonomists like to think of

457
00:17:37,669 --> 00:17:35,600
ourselves as artists able to perceive

458
00:17:39,350 --> 00:17:37,679
form shape color and relationships that

459
00:17:41,430 --> 00:17:39,360
are hidden from the gates of the more

460
00:17:43,350 --> 00:17:41,440
mundane scientists

461
00:17:45,909 --> 00:17:43,360
and this is not this does not

462
00:17:48,310 --> 00:17:45,919
necessarily reflect my own opinion but

463
00:17:52,470 --> 00:17:48,320

when it comes to taxonomy um a lot in a

464
00:17:55,510 --> 00:17:52,480
lot of cases people agree to disagree

465
00:17:57,669 --> 00:17:55,520
and uh recently i found this paper

466
00:17:59,510 --> 00:17:57,679
to unite taxonomy in the first two lines

467
00:18:01,990 --> 00:17:59,520
the first line struck my attention

468
00:18:04,150 --> 00:18:02,000
because it asks people what do you think

469
00:18:06,070 --> 00:18:04,160
when you think of taxonomy is this an

470
00:18:08,070 --> 00:18:06,080
18th century gentleman in breaches well

471
00:18:10,230 --> 00:18:08,080
i can tell you i'm not an 18th century

472
00:18:12,630 --> 00:18:10,240
gentleman in breaches and how did i get

473
00:18:15,590 --> 00:18:12,640
involved in taxonomy then

474
00:18:18,549 --> 00:18:15,600
well it all started in 2011 when i was

475
00:18:21,430 --> 00:18:18,559
uh um attending the evergreen page

476
00:18:23,750 --> 00:18:21,440
biology meeting in olympia washington um

477
00:18:26,070 --> 00:18:23,760
and took this nice picture or had this

478
00:18:27,430 --> 00:18:26,080
nice picture taken of me on uh climbing

479
00:18:28,630 --> 00:18:27,440
mount veneer

480
00:18:30,710 --> 00:18:28,640
and

481
00:18:33,590 --> 00:18:30,720
what happened there is there was a

482
00:18:35,909 --> 00:18:33,600
couple of scientists who just isolated a

483
00:18:37,830 --> 00:18:35,919
new phage and we all realized that it

484
00:18:39,590 --> 00:18:37,840
was very much the same so we got

485
00:18:42,630 --> 00:18:39,600
together

486
00:18:44,070 --> 00:18:42,640
and wrote a paper suggesting a new genus

487
00:18:46,950 --> 00:18:44,080
of phages

488
00:18:49,510 --> 00:18:46,960

called fiona-like virus

489
00:18:56,150 --> 00:18:52,310
that was not the end of it because

490
00:18:59,590 --> 00:18:56,160
writing a paper is not enough to make

491
00:19:02,710 --> 00:18:59,600
taxonomy official as i soon find out

492
00:19:05,350 --> 00:19:02,720
found out so what did i do then

493
00:19:08,870 --> 00:19:05,360
i went to look at what

494
00:19:10,549 --> 00:19:08,880
body is in charge of taxonomy

495
00:19:13,830 --> 00:19:10,559
so and that is the international

496
00:19:15,830 --> 00:19:13,840
committee on taxonomy of viruses of ictv

497
00:19:18,789 --> 00:19:15,840
of which i now am a part of and it all

498
00:19:21,029 --> 00:19:18,799
started at that one conference

499
00:19:22,470 --> 00:19:21,039
so the objectives of the ictv are to

500
00:19:24,950 --> 00:19:22,480
develop an internationally agreed

501
00:19:27,270 --> 00:19:24,960
taxonomy for viruses

502
00:19:30,070 --> 00:19:27,280
develop internationally green names for

503
00:19:32,470 --> 00:19:30,080
virus taxa to communicate the decisions

504
00:19:33,590 --> 00:19:32,480
and to maintain an index of the agreed

505
00:19:35,029 --> 00:19:33,600
names

506
00:19:36,950 --> 00:19:35,039
so if you want to learn more about it

507
00:19:37,990 --> 00:19:36,960
you can go to the website

508
00:19:40,230 --> 00:19:38,000
um so

509
00:19:41,510 --> 00:19:40,240
it's it's a committee so a committee has

510
00:19:43,830 --> 00:19:41,520
a structure

511
00:19:45,590 --> 00:19:43,840
there's an executive committee um filled

512
00:19:47,909 --> 00:19:45,600
with all of these big shots and

513
00:19:49,990 --> 00:19:47,919

scientists some of which are in the

514
00:19:51,990 --> 00:19:50,000
audience um the main is made of

515
00:19:53,510 --> 00:19:52,000
subcommittees who each have an area of

516
00:19:55,029 --> 00:19:53,520
expertise and these subcommittees have

517
00:19:57,350 --> 00:19:55,039
study groups and each of these study

518
00:19:59,350 --> 00:19:57,360
groups have even more specialized areas

519
00:20:01,110 --> 00:19:59,360
of expertise

520
00:20:03,110 --> 00:20:01,120
so what happens when you have a new

521
00:20:05,270 --> 00:20:03,120
virus and you want to have it officially

522
00:20:07,830 --> 00:20:05,280
classified or when you discover a new

523
00:20:08,750 --> 00:20:07,840
group of viruses that fit in a taxon

524
00:20:10,950 --> 00:20:08,760
it's a

525
00:20:12,470 --> 00:20:10,960
semi-convoluted process that that can

526
00:20:13,750 --> 00:20:12,480
take quite a while

527
00:20:16,549 --> 00:20:13,760
starting with

528
00:20:18,310 --> 00:20:16,559
identifying what type of virus it is so

529
00:20:19,830 --> 00:20:18,320
what study group or subcommittee you

530
00:20:21,990 --> 00:20:19,840
need to go to and write the taxonomy

531
00:20:23,510 --> 00:20:22,000
proposal and then you submit it and you

532
00:20:25,350 --> 00:20:23,520
might need to revise and then it gets

533
00:20:26,470 --> 00:20:25,360
submitted to the sub-commit

534
00:20:28,390 --> 00:20:26,480
subcommittee chair and then the

535
00:20:30,149 --> 00:20:28,400
subcommittee chair submits it to the

536
00:20:32,870 --> 00:20:30,159
executive committee and the executive

537
00:20:35,590 --> 00:20:32,880
committee meets once a year and

538
00:20:37,669 --> 00:20:35,600

discusses all the proposals and

539
00:20:39,990 --> 00:20:37,679
the proposals can then be accepted or

540
00:20:41,510 --> 00:20:40,000
asked for revisions or rejected and

541
00:20:43,510 --> 00:20:41,520
after that all of the accepted and

542
00:20:46,470 --> 00:20:43,520
revised proposals are voted on by all

543
00:20:48,870 --> 00:20:46,480
the members of the ictv and once these

544
00:20:50,549 --> 00:20:48,880
are ratified then there's a taxonomy

545
00:20:53,350 --> 00:20:50,559
gets updated there's a new master

546
00:20:56,310 --> 00:20:53,360
species list and this links to external

547
00:20:58,390 --> 00:20:56,320
databases including the ncbi taxonomy

548
00:21:01,350 --> 00:20:58,400
database

549
00:21:03,510 --> 00:21:01,360
so that just goes to show you that it's

550
00:21:05,590 --> 00:21:03,520
it's not as straightforward as just

551
00:21:07,510 --> 00:21:05,600
publishing a paper saying hey look i

552
00:21:09,830 --> 00:21:07,520
have a new genus or hey look i have a

553
00:21:12,149 --> 00:21:09,840
new species

554
00:21:13,830 --> 00:21:12,159
a quick note on what a taxon is so

555
00:21:15,510 --> 00:21:13,840
basically it's a box

556
00:21:18,470 --> 00:21:15,520
and

557
00:21:22,789 --> 00:21:18,480
humans want to put stuff in boxes

558
00:21:25,110 --> 00:21:24,230
the truth

559
00:21:26,950 --> 00:21:25,120
but

560
00:21:28,070 --> 00:21:26,960
a lot of these boxes are meaningful and

561
00:21:29,830 --> 00:21:28,080
helpful

562
00:21:31,669 --> 00:21:29,840
to go forward and i hope that in the

563
00:21:33,510 --> 00:21:31,679

next part of my presentation i can

564
00:21:34,710 --> 00:21:33,520
actually convince you convince you of

565
00:21:37,029 --> 00:21:34,720
that

566
00:21:38,149 --> 00:21:37,039
so in this case um our

567
00:21:41,190 --> 00:21:38,159
box

568
00:21:44,149 --> 00:21:41,200
strategy is hierarchical so um it's a

569
00:21:46,870 --> 00:21:44,159
bit like the russian dolls um every

570
00:21:49,029 --> 00:21:46,880
smaller taxon fits in the large the next

571
00:21:52,470 --> 00:21:49,039
larger one and we go all the way from

572
00:21:55,190 --> 00:21:52,480
realm uh which ends in the suffix viria

573
00:21:58,230 --> 00:21:55,200
down to the species

574
00:22:00,630 --> 00:21:58,240
mostly ends in the suffix suffix virus

575
00:22:04,070 --> 00:22:00,640
and anything in between but not all of

576
00:22:07,590 --> 00:22:05,909
so with all of this this information

577
00:22:09,590 --> 00:22:07,600
you're you're probably thinking why

578
00:22:11,990 --> 00:22:09,600
should i care so i'm going to try to

579
00:22:14,310 --> 00:22:12,000
link this taxonomy to virus discovery

580
00:22:16,549 --> 00:22:14,320
and astrovirology

581
00:22:18,230 --> 00:22:16,559
and for that i want to take you down to

582
00:22:19,190 --> 00:22:18,240
antarctica

583
00:22:23,909 --> 00:22:19,200
and

584
00:22:25,750 --> 00:22:23,919
the dry valleys in antarctica are

585
00:22:28,549 --> 00:22:25,760
actually a very good

586
00:22:30,310 --> 00:22:28,559
proxy to study what life on another

587
00:22:31,350 --> 00:22:30,320
planet might look like

588
00:22:33,270 --> 00:22:31,360

because

589
00:22:35,430 --> 00:22:33,280
it's the coldest and driest place on

590
00:22:37,110 --> 00:22:35,440
earth and when the explorers first

591
00:22:39,350 --> 00:22:37,120
arrived there they thought that there

592
00:22:41,510 --> 00:22:39,360
was nothing there

593
00:22:44,310 --> 00:22:41,520
but actually there's quite a lot of life

594
00:22:46,390 --> 00:22:44,320
and a lot of it is microbial life

595
00:22:49,270 --> 00:22:46,400
and where there's microbial life there's

596
00:22:50,630 --> 00:22:49,280
viruses of microbes so actually this

597
00:22:52,630 --> 00:22:50,640
picture that you see is full of

598
00:22:54,789 --> 00:22:52,640
bacteriophages

599
00:22:57,190 --> 00:22:54,799
virus infecting bacteria but it's also

600
00:22:58,950 --> 00:22:57,200
full of other viruses of microbes and

601
00:23:01,590 --> 00:22:58,960
one of the interesting features there

602
00:23:04,710 --> 00:23:01,600
are are hyperlids so when you walk into

603
00:23:06,470 --> 00:23:04,720
this desert-like environment

604
00:23:09,909 --> 00:23:06,480
and unfortunately i've never been there

605
00:23:12,630 --> 00:23:09,919
only my my boss has been there um you

606
00:23:14,549 --> 00:23:12,640
pick up a rock a quartz rock it's

607
00:23:16,310 --> 00:23:14,559
translucent and what you see underneath

608
00:23:17,909 --> 00:23:16,320
is this community and it's a complex

609
00:23:20,710 --> 00:23:17,919
community of cyanobacteria and

610
00:23:23,270 --> 00:23:20,720
heterotrophic bacteria who managed to

611
00:23:25,830 --> 00:23:23,280
find find a niche for themselves

612
00:23:27,669 --> 00:23:25,840
protected from the sun from uv radiation

613
00:23:29,190 --> 00:23:27,679

from wind

614
00:23:30,390 --> 00:23:29,200
and from

615
00:23:31,830 --> 00:23:30,400
any of the

616
00:23:34,789 --> 00:23:31,840
harsh

617
00:23:37,669 --> 00:23:34,799
conditions that are present there

618
00:23:40,070 --> 00:23:37,679
so when we investigated these antarctic

619
00:23:41,750 --> 00:23:40,080
soil viruses and how did we do that we

620
00:23:43,029 --> 00:23:41,760
did that by

621
00:23:50,149 --> 00:23:43,039
sequencing

622
00:23:51,909 --> 00:23:50,159
their dna genomes we saw that most of

623
00:23:53,269 --> 00:23:51,919
the sequence data that we got was

624
00:23:57,590 --> 00:23:53,279
unknown

625
00:24:00,230 --> 00:23:57,600
and 5 to 15 of the data that we did find

626
00:24:03,110 --> 00:24:00,240
had an equivalent in the database

627
00:24:05,350 --> 00:24:03,120
so when we looked at what was that known

628
00:24:09,190 --> 00:24:05,360
fraction in antarctic soul it was mostly

629
00:24:10,950 --> 00:24:09,200
bacteriophages and then um to my

630
00:24:14,630 --> 00:24:10,960
the first time i i saw one of these

631
00:24:16,549 --> 00:24:14,640
plots um horrifyingly i saw all of these

632
00:24:18,470 --> 00:24:16,559
unclassified cypher very day and

633
00:24:20,549 --> 00:24:18,480
classified this unclassified that

634
00:24:22,630 --> 00:24:20,559
unclassified that and i thought oh no

635
00:24:26,789 --> 00:24:22,640
i'm i'm now on the taxonomy committee

636
00:24:27,590 --> 00:24:26,799
this is now my job to fix this

637
00:24:29,350 --> 00:24:27,600
so

638
00:24:32,549 --> 00:24:29,360

it basically comes down to two main

639
00:24:34,230 --> 00:24:32,559
challenges when you look at

640
00:24:37,510 --> 00:24:34,240
an environment like that we have the

641
00:24:41,029 --> 00:24:37,520
unknowns and we have the unclassifieds

642
00:24:43,190 --> 00:24:41,039
and the unknowns are actually have

643
00:24:45,350 --> 00:24:43,200
a really apt name

644
00:24:47,430 --> 00:24:45,360
because

645
00:24:49,590 --> 00:24:47,440
they're called microbial dark matter at

646
00:24:51,750 --> 00:24:49,600
the moment and for viruses that that

647
00:24:53,430 --> 00:24:51,760
would then be microbial dark matter

648
00:24:56,390 --> 00:24:53,440
because you can break the unknown

649
00:24:58,470 --> 00:24:56,400
problem into four quadrants you have a

650
00:25:01,110 --> 00:24:58,480
protein that's known

651
00:25:03,029 --> 00:25:01,120
in a virus type that's known so those

652
00:25:04,950 --> 00:25:03,039
are the well-known proteins but you can

653
00:25:07,990 --> 00:25:04,960
have a protein of which the function is

654
00:25:09,830 --> 00:25:08,000
known but that's in an unknown virus

655
00:25:12,310 --> 00:25:09,840
genome so those are potentially new

656
00:25:14,710 --> 00:25:12,320
lineages of viruses

657
00:25:16,390 --> 00:25:14,720
and then you have um

658
00:25:18,549 --> 00:25:16,400
a known lineage

659
00:25:20,470 --> 00:25:18,559
of virus or a known genome but it has

660
00:25:22,390 --> 00:25:20,480
this one protein in it that's unknown or

661
00:25:24,470 --> 00:25:22,400
multiple proteins that are known so then

662
00:25:26,470 --> 00:25:24,480
you have potentially new functions

663
00:25:28,789 --> 00:25:26,480

but where those two meet and what we

664
00:25:30,950 --> 00:25:28,799
call the unknown unknowns that's what

665
00:25:33,269 --> 00:25:30,960
microbial dark matter is

666
00:25:36,149 --> 00:25:33,279
and that unfortunately is not something

667
00:25:38,470 --> 00:25:36,159
that a taxonomy committee can fix

668
00:25:40,789 --> 00:25:38,480
so then that brings me to what the

669
00:25:43,430 --> 00:25:40,799
taxonomy committee can fix and that's

670
00:25:45,750 --> 00:25:43,440
the unclassifieds so that's what we try

671
00:25:48,870 --> 00:25:45,760
to do is try to bring order to the known

672
00:25:50,230 --> 00:25:48,880
virus sphere how do we do that by

673
00:25:52,470 --> 00:25:50,240
clustering

674
00:25:54,549 --> 00:25:52,480
genomes by type

675
00:25:57,110 --> 00:25:54,559
looking at their shared protein content

676
00:25:59,830 --> 00:25:57,120
looking at evolutionary relationships of

677
00:26:01,990 --> 00:25:59,840
certain signature genes and even looking

678
00:26:04,310 --> 00:26:02,000
for conserved structural traits that

679
00:26:09,029 --> 00:26:04,320
cannot be resolved by

680
00:26:13,350 --> 00:26:11,990
so in the past couple of years i've

681
00:26:16,789 --> 00:26:13,360
tried to

682
00:26:18,390 --> 00:26:16,799
help people who are developing new tools

683
00:26:20,149 --> 00:26:18,400
with my

684
00:26:22,070 --> 00:26:20,159
wisdom that i've gathered in the last

685
00:26:23,830 --> 00:26:22,080
years of of taxonomy and one of the

686
00:26:25,590 --> 00:26:23,840
tools that was developed recently is

687
00:26:26,470 --> 00:26:25,600
called gravity

688
00:26:28,390 --> 00:26:26,480

and

689
00:26:31,750 --> 00:26:28,400
this is what what some of the output

690
00:26:33,669 --> 00:26:31,760
looks like and what we try to do is we

691
00:26:35,269 --> 00:26:33,679
we can use this for any type of virus

692
00:26:36,789 --> 00:26:35,279
but the example that i'm giving you now

693
00:26:39,269 --> 00:26:36,799
is for bacteriophages because

694
00:26:41,830 --> 00:26:39,279
bacteriophages are my main love

695
00:26:44,870 --> 00:26:41,840
um so we just tried to distill the

696
00:26:45,909 --> 00:26:44,880
difference between two genomes into one

697
00:26:50,830 --> 00:26:45,919
number

698
00:26:53,350 --> 00:26:50,840
so we looked at

699
00:26:56,549 --> 00:26:53,360
how um

700
00:26:58,549 --> 00:26:56,559
how different or how similar

701
00:27:00,390 --> 00:26:58,559
proteins are that are encoded by the

702
00:27:02,230 --> 00:27:00,400
genomes and look at all of the proteins

703
00:27:04,870 --> 00:27:02,240
and then look at their

704
00:27:07,269 --> 00:27:04,880
genome gene order

705
00:27:09,669 --> 00:27:07,279
and build that into a number combine

706
00:27:11,990 --> 00:27:09,679
those numbers into a composite

707
00:27:14,870 --> 00:27:12,000
generalized occurred distance and then

708
00:27:17,269 --> 00:27:14,880
have this matrix of all the genomes

709
00:27:20,070 --> 00:27:17,279
versus each other and how much they

710
00:27:22,870 --> 00:27:20,080
share so a zero is that the all of the

711
00:27:24,230 --> 00:27:22,880
proteins are 100 similar and are in the

712
00:27:26,070 --> 00:27:24,240
same order

713
00:27:28,310 --> 00:27:26,080

in the genome and then the one is

714
00:27:29,990 --> 00:27:28,320
nothing similar and

715
00:27:32,389 --> 00:27:30,000
obviously if nothing similar nothing is

716
00:27:34,630 --> 00:27:32,399
in the same order and when we do that

717
00:27:36,149 --> 00:27:34,640
for all the known viruses we can see

718
00:27:41,190 --> 00:27:36,159
patterns

719
00:27:43,990 --> 00:27:41,200
looking for to create boxes and in this

720
00:27:45,269 --> 00:27:44,000
case we created a new box called herrell

721
00:27:48,710 --> 00:27:45,279
veride

722
00:27:52,470 --> 00:27:48,720
which was a new family

723
00:27:54,310 --> 00:27:52,480
a different but similar strategy um

724
00:27:56,149 --> 00:27:54,320
was developed by

725
00:27:58,630 --> 00:27:56,159
the lab of matt sullivan and led by

726
00:28:00,070 --> 00:27:58,640
hobin yang and ben baldac

727
00:28:02,870 --> 00:28:00,080
and this

728
00:28:04,789 --> 00:28:02,880
this is a network-based approach so

729
00:28:07,430 --> 00:28:04,799
a part of it is similar you look at the

730
00:28:10,630 --> 00:28:07,440
genes that are encoded by

731
00:28:13,350 --> 00:28:10,640
genomes of viruses in this case it's

732
00:28:16,549 --> 00:28:13,360
bacterial anarchial viruses

733
00:28:18,389 --> 00:28:16,559
and if two genomes share a gene

734
00:28:21,110 --> 00:28:18,399
then they can get connected in the

735
00:28:24,710 --> 00:28:21,120
network by

736
00:28:27,110 --> 00:28:24,720
an edge so one of the lines and the more

737
00:28:30,870 --> 00:28:27,120
genes two genomes have in common the

738
00:28:33,269 --> 00:28:30,880

closer together they are in the network

739
00:28:35,750 --> 00:28:33,279
so we use that again to define this

740
00:28:38,070 --> 00:28:35,760
family herald variday and its internal

741
00:28:40,789 --> 00:28:38,080
structure

742
00:28:43,909 --> 00:28:40,799
but what is cool about this method is

743
00:28:45,269 --> 00:28:43,919
that you can expand it into the dark

744
00:28:46,230 --> 00:28:45,279
virus sphere

745
00:28:48,470 --> 00:28:46,240
so

746
00:28:50,710 --> 00:28:48,480
what you see here is the same picture as

747
00:28:53,110 --> 00:28:50,720
the slide before and everything that was

748
00:28:54,950 --> 00:28:53,120
in the slide before is colored in red

749
00:28:58,789 --> 00:28:54,960
but now um

750
00:29:01,669 --> 00:28:58,799
it got expanded by unknown dna viruses

751
00:29:04,630 --> 00:29:01,679
from the global ocean virum data set and

752
00:29:07,029 --> 00:29:04,640
then and this is in total almost 17 000

753
00:29:10,230 --> 00:29:07,039
virus genomes that are added

754
00:29:12,950 --> 00:29:10,240
and again we we're looking for patterns

755
00:29:14,149 --> 00:29:12,960
and what we see is that the majority of

756
00:29:15,909 --> 00:29:14,159
viruses

757
00:29:18,870 --> 00:29:15,919
that are in

758
00:29:20,389 --> 00:29:18,880
environments that are not isolated that

759
00:29:23,190 --> 00:29:20,399
they're new

760
00:29:25,830 --> 00:29:23,200
and don't cluster or

761
00:29:28,950 --> 00:29:25,840
cluster very far away from the known

762
00:29:32,149 --> 00:29:28,960
viruses so why do we want to look at it

763
00:29:36,389 --> 00:29:32,159

like this well we want to create

764
00:29:37,830 --> 00:29:36,399
meaningful groups so that we actually

765
00:29:40,710 --> 00:29:37,840
can use this

766
00:29:43,190 --> 00:29:40,720
so i've just highlighted two um because

767
00:29:47,029 --> 00:29:43,200
we because i know about them

768
00:29:48,630 --> 00:29:47,039
so um on the left hand side in green you

769
00:29:50,630 --> 00:29:48,640
have the group that i was talking about

770
00:29:52,710 --> 00:29:50,640
before which have which has bacterial

771
00:29:54,870 --> 00:29:52,720
features that are potentially useful in

772
00:29:55,990 --> 00:29:54,880
combating bacterial contamination in

773
00:29:58,470 --> 00:29:56,000
food

774
00:30:00,389 --> 00:29:58,480
so if you have unknown viruses that

775
00:30:01,830 --> 00:30:00,399
cluster their unknown bacteriophages

776
00:30:03,269 --> 00:30:01,840
maybe you can use them for the same

777
00:30:05,029 --> 00:30:03,279
application

778
00:30:07,269 --> 00:30:05,039
on the other hand on the right hand side

779
00:30:09,190 --> 00:30:07,279
you have a group of bacteriophages that

780
00:30:11,669 --> 00:30:09,200
are known to integrate into bacterial

781
00:30:13,430 --> 00:30:11,679
genomes and have the potential to carry

782
00:30:15,510 --> 00:30:13,440
a toxin

783
00:30:18,470 --> 00:30:15,520
and in this case a lot of these carry a

784
00:30:21,750 --> 00:30:18,480
shiga toxin that can go into e coli and

785
00:30:24,549 --> 00:30:21,760
cause severe diarrheal illness so if

786
00:30:26,870 --> 00:30:24,559
again if you know that your your viruses

787
00:30:30,149 --> 00:30:26,880
or your bacteriophages cluster here then

788
00:30:31,750 --> 00:30:30,159

you have the first indication that

789
00:30:32,870 --> 00:30:31,760
this might be going on with these as

790
00:30:36,710 --> 00:30:32,880
well

791
00:30:38,389 --> 00:30:36,720
so the ultimate goal is to

792
00:30:40,870 --> 00:30:38,399
make a picture of this of all the

793
00:30:43,110 --> 00:30:40,880
viruses out there and then color all the

794
00:30:44,230 --> 00:30:43,120
dots with meaningful groups and

795
00:30:48,149 --> 00:30:44,240
functions

796
00:30:51,269 --> 00:30:48,159
which is of course not ambitious at all

797
00:30:54,230 --> 00:30:51,279
and this reminded me of of a figure that

798
00:30:55,430 --> 00:30:54,240
probably a lot of you have seen before

799
00:30:58,549 --> 00:30:55,440
which is

800
00:31:01,669 --> 00:30:58,559
a epifluorescent microscopy picture of a

801
00:31:03,669 --> 00:31:01,679
drop of sea water so

802
00:31:06,630 --> 00:31:03,679
all the nucleic acid here in a drop of

803
00:31:09,909 --> 00:31:06,640
seawater is dyed with cyber green and

804
00:31:12,230 --> 00:31:09,919
all of the tiny dots are virus particles

805
00:31:14,549 --> 00:31:12,240
and all of the bigger dots are bacteria

806
00:31:16,230 --> 00:31:14,559
and you can see that viruses and

807
00:31:17,509 --> 00:31:16,240
bacteria are

808
00:31:19,430 --> 00:31:17,519
basically

809
00:31:23,750 --> 00:31:19,440
everywhere

810
00:31:26,149 --> 00:31:23,760
and this obviously reminded me of

811
00:31:28,230 --> 00:31:26,159
what it looks like um at the final

812
00:31:29,350 --> 00:31:28,240
frontier

813
00:31:33,990 --> 00:31:29,360

and

814
00:31:35,350 --> 00:31:34,000
you ask is this coincidence absolutely

815
00:31:38,710 --> 00:31:35,360
and

816
00:31:41,430 --> 00:31:38,720
with that i come to the end of my talk

817
00:31:43,269 --> 00:31:41,440
and i'd like to acknowledge

818
00:31:45,269 --> 00:31:43,279
my two main

819
00:31:46,870 --> 00:31:45,279
taxonomy mentors andrew krabinski and

820
00:31:49,509 --> 00:31:46,880
rob levine

821
00:31:52,070 --> 00:31:49,519
all of the members of the ictv my

822
00:31:54,149 --> 00:31:52,080
committee and my study group

823
00:31:57,509 --> 00:31:54,159
and all of my colleagues at the quadrant

824
00:32:08,070 --> 00:31:57,519
institute and my funders the ddsrc

825
00:32:16,070 --> 00:32:09,750
does anyone have any questions before i

826
00:32:19,830 --> 00:32:17,909
all right evelyn i i think it might be

827
00:32:22,070 --> 00:32:19,840
nice to talk about this some

828
00:32:23,269 --> 00:32:22,080
if we're able to get taxonomy for

829
00:32:25,830 --> 00:32:23,279
viruses

830
00:32:27,509 --> 00:32:25,840
you know with microbial taxonomy once we

831
00:32:29,990 --> 00:32:27,519
have a certain taxonomic group we can

832
00:32:33,350 --> 00:32:30,000
start inferring things that they can do

833
00:32:34,950 --> 00:32:33,360
so from taxonomy of viruses could we

834
00:32:37,269 --> 00:32:34,960
basically look at a certain level and

835
00:32:39,350 --> 00:32:37,279
determine how big that virus is or who

836
00:32:42,470 --> 00:32:39,360
that virus infects or

837
00:32:43,909 --> 00:32:42,480
maybe a biome that it lives in

838
00:32:45,830 --> 00:32:43,919

um i think

839
00:32:47,110 --> 00:32:45,840
that's the ultimate goal

840
00:32:48,149 --> 00:32:47,120
and

841
00:32:50,470 --> 00:32:48,159
um

842
00:32:52,950 --> 00:32:50,480
what we're trying to do now with these

843
00:32:55,430 --> 00:32:52,960
big data is looking for patterns and

844
00:32:57,190 --> 00:32:55,440
then see if the groups that we see

845
00:33:00,310 --> 00:32:57,200
um have something

846
00:33:03,430 --> 00:33:00,320
shared so for a lot of these you can

847
00:33:06,470 --> 00:33:03,440
it's quite it's quite easy to do to say

848
00:33:08,149 --> 00:33:06,480
in broad terms this is probably a

849
00:33:10,230 --> 00:33:08,159
bacterial virus this is probably an

850
00:33:12,549 --> 00:33:10,240
archaeal virus this is probably a human

851
00:33:15,269 --> 00:33:12,559
virus and then

852
00:33:17,350 --> 00:33:15,279
specifically for certain groups that are

853
00:33:19,190 --> 00:33:17,360
well investigated like some of the human

854
00:33:22,389 --> 00:33:19,200
pathogens it will be quite

855
00:33:24,389 --> 00:33:22,399
straightforward to say that

856
00:33:26,470 --> 00:33:24,399
but for example for the bacteriophages a

857
00:33:28,070 --> 00:33:26,480
lot of it is is dark matter at the

858
00:33:30,230 --> 00:33:28,080
moment because

859
00:33:32,149 --> 00:33:30,240
um you could see on the tree of life

860
00:33:34,389 --> 00:33:32,159
that there's so many different bacteria

861
00:33:36,470 --> 00:33:34,399
and most of them um a lot of them have

862
00:33:38,070 --> 00:33:36,480
not been cultured yet so we don't have

863
00:33:39,269 --> 00:33:38,080

any information about the viruses that

864
00:33:40,149 --> 00:33:39,279
infect them

865
00:33:41,430 --> 00:33:40,159
so

866
00:33:43,909 --> 00:33:41,440
it will be

867
00:33:45,669 --> 00:33:43,919
very very difficult to extract

868
00:33:54,630 --> 00:33:45,679
information

869
00:33:57,990 --> 00:33:56,549
i guess i'll ask another question then

870
00:34:00,470 --> 00:33:58,000
um

871
00:34:03,029 --> 00:34:00,480
so we know that viruses don't have this

872
00:34:05,509 --> 00:34:03,039
universal marker gene analogous to the

873
00:34:06,870 --> 00:34:05,519
16s or 18s

874
00:34:09,430 --> 00:34:06,880
and you already talked about trying to

875
00:34:12,389 --> 00:34:09,440
use shared protein content

876
00:34:14,230 --> 00:34:12,399
to connect viruses but we also know that

877
00:34:16,470 --> 00:34:14,240
viruses have

878
00:34:18,790 --> 00:34:16,480
all forms of nucleic acid

879
00:34:20,629 --> 00:34:18,800
so how do we

880
00:34:21,990 --> 00:34:20,639
try to bring together all the different

881
00:34:23,909 --> 00:34:22,000
types of viruses you know we have

882
00:34:25,589 --> 00:34:23,919
double-stranded dna single-stranded

883
00:34:27,109 --> 00:34:25,599
double-stranded rna single-stranded do

884
00:34:29,589 --> 00:34:27,119
you imagine

885
00:34:32,629 --> 00:34:29,599
one big viral tree one day or or keeping

886
00:34:35,750 --> 00:34:32,639
them separate to start or just any um

887
00:34:38,550 --> 00:34:35,760
knowledge on that so what what um a

888
00:34:41,750 --> 00:34:38,560

group of people at the ictv are are

889
00:34:43,510 --> 00:34:41,760
doing and this is research that um some

890
00:34:45,829 --> 00:34:43,520
groups across the world are doing is

891
00:34:47,190 --> 00:34:45,839
look at the very deep evolutionary

892
00:34:50,069 --> 00:34:47,200
relationships

893
00:34:52,310 --> 00:34:50,079
and then you kind of go and look at

894
00:34:54,430 --> 00:34:52,320
protein folds in the capsid

895
00:34:57,910 --> 00:34:54,440
and then you can for example link

896
00:35:00,710 --> 00:34:57,920
adenoviruses with prd-1 bacteriophages

897
00:35:03,030 --> 00:35:00,720
or you can link

898
00:35:04,310 --> 00:35:03,040
tailed bacteriophages with herpes

899
00:35:08,710 --> 00:35:04,320
viruses

900
00:35:09,349 --> 00:35:08,720
um

901
00:35:25,589 --> 00:35:09,359
a

902
00:35:27,589 --> 00:35:25,599
root

903
00:35:30,390 --> 00:35:27,599
so there's there's some controversy

904
00:35:32,310 --> 00:35:30,400
there um so

905
00:35:34,790 --> 00:35:32,320
there are methods for example the

906
00:35:37,670 --> 00:35:34,800
gravity that was developed by um

907
00:35:39,109 --> 00:35:37,680
ghan ayusakun and peter simmons um they

908
00:35:40,710 --> 00:35:39,119
broke it up in a different baltimore

909
00:35:42,390 --> 00:35:40,720
group so they have a different plot like

910
00:35:45,829 --> 00:35:42,400
that for each of the

911
00:35:48,790 --> 00:35:45,839
each of the baltimore groups

912
00:35:50,230 --> 00:35:48,800
i see some uh questions in the chat

913
00:35:51,829 --> 00:35:50,240

as well but

914
00:35:53,990 --> 00:35:51,839
yes go start with nigel then we'll go to

915
00:35:58,470 --> 00:35:54,000
the chat yeah

916
00:35:59,750 --> 00:35:58,480
um i've got a naive question if i may um

917
00:36:01,589 --> 00:35:59,760
so

918
00:36:04,069 --> 00:36:01,599
you're talking all the time as if it's

919
00:36:06,310 --> 00:36:04,079
obvious that if you do viral taxonomy

920
00:36:08,470 --> 00:36:06,320
you're going to end up with the tree

921
00:36:10,950 --> 00:36:08,480
but um

922
00:36:14,150 --> 00:36:10,960
in fact the only reason that happens uh

923
00:36:16,710 --> 00:36:14,160
for uh you say bacteria or the three

924
00:36:19,510 --> 00:36:16,720
domains of life is because

925
00:36:21,990 --> 00:36:19,520
we're in a in an era where uh massive

926
00:36:23,750 --> 00:36:22,000
horizontal gene transfer has been uh

927
00:36:26,790 --> 00:36:23,760
turned off at least according to the

928
00:36:28,310 --> 00:36:26,800
theory that carl rose and i promulgated

929
00:36:30,230 --> 00:36:28,320
and so on so

930
00:36:32,630 --> 00:36:30,240
so i would have if you'd asked me to

931
00:36:34,150 --> 00:36:32,640
guess i would have guessed that that

932
00:36:36,069 --> 00:36:34,160
for the reasons that you said about

933
00:36:37,829 --> 00:36:36,079
there being no marker genes and so on

934
00:36:39,670 --> 00:36:37,839
that viruses

935
00:36:42,470 --> 00:36:39,680
would form some kind of network where

936
00:36:46,150 --> 00:36:42,480
the links would be whatever quality you

937
00:36:47,829 --> 00:36:46,160
want to associate it to them so um to

938
00:36:49,910 --> 00:36:47,839

what extent is the thinking that that

939
00:36:51,190 --> 00:36:49,920
we're looking for trees really built

940
00:36:53,270 --> 00:36:51,200
into this

941
00:36:55,270 --> 00:36:53,280
into viral taxonomy so so this is

942
00:36:58,870 --> 00:36:55,280
exactly why i wanted to show the network

943
00:37:01,910 --> 00:36:58,880
um because i think so the tree is kind

944
00:37:04,630 --> 00:37:01,920
of the the framework that we are trying

945
00:37:07,990 --> 00:37:04,640
to force ourselves into because that's

946
00:37:09,510 --> 00:37:08,000
what's um what makes taxonomy uh most

947
00:37:13,030 --> 00:37:09,520
easily digestible if you have a

948
00:37:14,790 --> 00:37:13,040
hierarchical system right um but but we

949
00:37:16,150 --> 00:37:14,800
do see and and with bacteriophages

950
00:37:19,109 --> 00:37:16,160
there's a huge group of bacteria

951
00:37:21,910 --> 00:37:19,119
features that that have um

952
00:37:23,349 --> 00:37:21,920
a rampant mosaicism as it's usually

953
00:37:25,190 --> 00:37:23,359
called it's it's a minority of

954
00:37:28,069 --> 00:37:25,200
bacteriophages but there's it's still a

955
00:37:29,670 --> 00:37:28,079
significant group so one of the one of

956
00:37:32,150 --> 00:37:29,680
the things and this is my personal

957
00:37:33,670 --> 00:37:32,160
opinion um is that

958
00:37:36,150 --> 00:37:33,680
we will

959
00:37:38,150 --> 00:37:36,160
at certain levels of taxonomy will have

960
00:37:40,470 --> 00:37:38,160
blurry lines and we'll say we cannot

961
00:37:42,950 --> 00:37:40,480
resolve this at this at

962
00:37:45,109 --> 00:37:42,960
this level but we can go to the groups

963
00:37:46,550 --> 00:37:45,119

higher and we can go to the groups

964
00:37:49,109 --> 00:37:46,560
uh lower but

965
00:37:51,589 --> 00:37:49,119
at these certain levels everything's

966
00:37:53,910 --> 00:37:51,599
there's too much gene exchange and and

967
00:37:55,990 --> 00:37:53,920
that's what what what is clear from

968
00:37:58,230 --> 00:37:56,000
those networks right i mean i think this

969
00:38:01,109 --> 00:37:58,240
goes back to the debate between carl

970
00:38:03,589 --> 00:38:01,119
rose and ernst mayer in the 1990s where

971
00:38:05,190 --> 00:38:03,599
they argued uh over the decade at cross

972
00:38:07,190 --> 00:38:05,200
purposes about what was the purpose of

973
00:38:10,150 --> 00:38:07,200
taxonomy and for mayor it was

974
00:38:12,550 --> 00:38:10,160
classification and for woes it was

975
00:38:13,990 --> 00:38:12,560
um just the result of what happens if

976
00:38:15,589 --> 00:38:14,000
you try to determine the evolutionary

977
00:38:18,630 --> 00:38:15,599
history of life

978
00:38:20,069 --> 00:38:18,640
and and so those goals are different and

979
00:38:21,430 --> 00:38:20,079
they and they may give you completely

980
00:38:24,230 --> 00:38:21,440
different structures

981
00:38:26,150 --> 00:38:24,240
yeah and i think maybe maybe some people

982
00:38:29,990 --> 00:38:26,160
get mad at me if i say this but for me

983
00:38:31,109 --> 00:38:30,000
taxonomy is about creating useful boxers

984
00:38:32,630 --> 00:38:31,119
um

985
00:38:35,430 --> 00:38:32,640
to help people for the future not

986
00:38:36,390 --> 00:38:35,440
necessarily to to

987
00:38:38,069 --> 00:38:36,400
answer

988
00:38:40,390 --> 00:38:38,079

all the questions about evolutionary

989
00:38:42,870 --> 00:38:40,400
history yeah i mean that wasn't mayor's

990
00:38:45,750 --> 00:38:42,880
view about the purpose of taxonomy was

991
00:38:49,030 --> 00:38:45,760
classification here yeah thank you a

992
00:38:52,630 --> 00:38:49,990
okay

993
00:38:55,589 --> 00:38:52,640
i see a question about um fragmented

994
00:38:58,069 --> 00:38:55,599
genomes um recovering from virums and

995
00:39:01,109 --> 00:38:58,079
how this affects classification and yes

996
00:39:03,109 --> 00:39:01,119
this is this is a something that i've

997
00:39:05,510 --> 00:39:03,119
struggled with myself

998
00:39:07,750 --> 00:39:05,520
um when you when you have a firearm data

999
00:39:10,069 --> 00:39:07,760
set and you know for example you have

1000
00:39:12,790 --> 00:39:10,079
rotavirus signatures in there and

1001
00:39:14,710 --> 00:39:12,800
rotavirus has 11 segments um it's a

1002
00:39:16,310 --> 00:39:14,720
double stranded rna virus

1003
00:39:17,589 --> 00:39:16,320
and you

1004
00:39:19,030 --> 00:39:17,599
you basically

1005
00:39:20,950 --> 00:39:19,040
um

1006
00:39:23,349 --> 00:39:20,960
what do you do if you have two of the

1007
00:39:27,109 --> 00:39:23,359
same segments in the same sample

1008
00:39:28,870 --> 00:39:27,119
i there's there's no easy way to say

1009
00:39:30,550 --> 00:39:28,880
these two segments belong together

1010
00:39:33,030 --> 00:39:30,560
because there's always a possibility

1011
00:39:35,030 --> 00:39:33,040
that in in one set of the very own

1012
00:39:36,550 --> 00:39:35,040
particles they get packaged separately

1013
00:39:38,150 --> 00:39:36,560

as in another set

1014
00:39:41,510 --> 00:39:38,160
so um

1015
00:39:45,750 --> 00:39:44,710
there might be um a machine learning way

1016
00:39:51,510 --> 00:39:45,760
to

1017
00:39:53,430 --> 00:39:51,520
answers ourselves and we cannot see it

1018
00:39:55,030 --> 00:39:53,440
with our eyes but that there's a machine

1019
00:40:05,109 --> 00:39:55,040
out there that's smarter and will fix

1020
00:40:08,390 --> 00:40:07,109
ken do you have other questions

1021
00:40:09,510 --> 00:40:08,400
yeah so there were another couple

1022
00:40:12,150 --> 00:40:09,520
questions from the chat that may have

1023
00:40:14,309 --> 00:40:12,160
gotten off your screen now so um do we

1024
00:40:16,390 --> 00:40:14,319
do taxonomy on contigs and can we get

1025
00:40:18,710 --> 00:40:16,400
different taxonomies for two contigs

1026
00:40:20,309 --> 00:40:18,720
from the same phage

1027
00:40:23,030 --> 00:40:20,319
um

1028
00:40:25,750 --> 00:40:23,040
yes okay so we do

1029
00:40:28,950 --> 00:40:25,760
we um we are technically allowed to do

1030
00:40:31,990 --> 00:40:28,960
taxonomy on context um but we only want

1031
00:40:34,309 --> 00:40:32,000
to do taxonomy and context that we have

1032
00:40:36,870 --> 00:40:34,319
very good evidence that they represent a

1033
00:40:39,109 --> 00:40:36,880
nearly complete

1034
00:40:40,390 --> 00:40:39,119
genome so

1035
00:40:45,510 --> 00:40:40,400
um

1036
00:40:48,069 --> 00:40:45,520
that was led by simon

1037
00:40:51,349 --> 00:40:48,079
rue and that kind of describes the

1038
00:40:53,829 --> 00:40:51,359

framework that we're working in um

1039
00:40:57,349 --> 00:40:53,839
with regards to what we're calling uvics

1040
00:40:59,829 --> 00:40:57,359
or uncultivated virus genomes

1041
00:41:01,510 --> 00:40:59,839
and yes it is entirely possible that

1042
00:41:03,109 --> 00:41:01,520
automated programs

1043
00:41:07,750 --> 00:41:03,119
will put two contexts that actually

1044
00:41:11,349 --> 00:41:09,589
taxon

1045
00:41:13,750 --> 00:41:11,359
and it all depends on what program you

1046
00:41:16,069 --> 00:41:13,760
use if you for for example if you use

1047
00:41:18,069 --> 00:41:16,079
the gene gene sharing networks of v

1048
00:41:20,309 --> 00:41:18,079
contact then

1049
00:41:22,550 --> 00:41:20,319
your two contexts of the same genome

1050
00:41:23,829 --> 00:41:22,560
will probably come into the same cluster

1051
00:41:25,430 --> 00:41:23,839
because if there's a near enough

1052
00:41:27,510 --> 00:41:25,440
reference they will the two contexts

1053
00:41:29,910 --> 00:41:27,520
will cluster together but if you for

1054
00:41:32,870 --> 00:41:29,920
example use a signature gene that is

1055
00:41:35,430 --> 00:41:32,880
signal copy as most of them are

1056
00:41:38,309 --> 00:41:35,440
then obviously it will only be present

1057
00:41:39,910 --> 00:41:38,319
in one of the context of the genome so

1058
00:41:42,470 --> 00:41:39,920
therefore

1059
00:41:47,670 --> 00:41:42,480
they will never come into the same taxon

1060
00:41:52,790 --> 00:41:50,390
i hope that answered the question

1061
00:41:54,630 --> 00:41:52,800
was wasn't my question there's some uh

1062
00:41:57,349 --> 00:41:54,640
interesting discussions going on in

1063
00:42:01,750 --> 00:41:57,359

there in the chat as well here

1064
00:42:04,870 --> 00:42:01,760
i i can't follow all of it yet um

1065
00:42:06,390 --> 00:42:04,880
did somebody asked about rna viruses

1066
00:42:08,390 --> 00:42:06,400
um

1067
00:42:10,550 --> 00:42:08,400
yeah so we have

1068
00:42:13,270 --> 00:42:10,560
we have a classification system for all

1069
00:42:16,230 --> 00:42:13,280
types of viruses um and

1070
00:42:18,150 --> 00:42:16,240
not all of the methods

1071
00:42:20,230 --> 00:42:18,160
can be carried over

1072
00:42:22,710 --> 00:42:20,240
because a lot of the the gene sharing

1073
00:42:24,790 --> 00:42:22,720
network the rna viruses are really small

1074
00:42:27,589 --> 00:42:24,800
and they include a very limited number

1075
00:42:29,589 --> 00:42:27,599
of proteins therefore your

1076
00:42:31,430 --> 00:42:29,599
resolution is not as big as for the

1077
00:42:33,190 --> 00:42:31,440
large dna viruses where there's a lot of

1078
00:42:34,710 --> 00:42:33,200
proteins where you can create beautiful

1079
00:42:36,230 --> 00:42:34,720
networks so

1080
00:42:39,270 --> 00:42:36,240
um

1081
00:42:41,990 --> 00:42:39,280
usually for rna viruses you go down to

1082
00:42:43,349 --> 00:42:42,000
single gene phylogenetics

1083
00:42:44,790 --> 00:42:43,359
yeah and i think this is something

1084
00:42:47,190 --> 00:42:44,800
you've addressed already another

1085
00:42:49,270 --> 00:42:47,200
question um was you know can such a

1086
00:42:51,270 --> 00:42:49,280
taxonomic tree be used to look at the

1087
00:42:54,309 --> 00:42:51,280
core components of a hypothetical virus

1088
00:42:56,829 --> 00:42:54,319

common ancestor um similar to luca what

1089
00:43:01,349 --> 00:42:56,839
are your thoughts on that

1090
00:43:04,790 --> 00:43:01,359
um yeah so there's definitely not a core

1091
00:43:06,069 --> 00:43:04,800
viral ancestor um because viruses

1092
00:43:08,230 --> 00:43:06,079
there's

1093
00:43:10,710 --> 00:43:08,240
multiple events uh in the past where

1094
00:43:13,910 --> 00:43:10,720
viruses arose and and i think some

1095
00:43:16,950 --> 00:43:13,920
people in in the chat are probably more

1096
00:43:19,910 --> 00:43:16,960
experts on this than i am um but there's

1097
00:43:22,470 --> 00:43:19,920
definitely um a multitude of different

1098
00:43:25,589 --> 00:43:22,480
lineages and there's no such thing as a

1099
00:43:27,589 --> 00:43:25,599
core viral ancestor is this proven

1100
00:43:30,470 --> 00:43:27,599
already i thought that was still a

1101
00:43:34,950 --> 00:43:31,910
yeah well

1102
00:43:36,069 --> 00:43:34,960
it's a good question um

1103
00:43:40,550 --> 00:43:36,079
you can't

1104
00:43:40,560 --> 00:43:44,069
yes

1105
00:43:44,079 --> 00:43:47,910
how

1106
00:43:52,630 --> 00:43:49,910
so it's it's it's all in the realm of

1107
00:43:55,510 --> 00:43:52,640
the theoretical right um and in in this

1108
00:43:56,790 --> 00:43:55,520
case you have viruses that are so vastly

1109
00:43:58,710 --> 00:43:56,800
different

1110
00:44:00,550 --> 00:43:58,720
um that there's

1111
00:44:01,910 --> 00:44:00,560
literally nothing

1112
00:44:04,470 --> 00:44:01,920
um

1113
00:44:08,710 --> 00:44:04,480

connecting them except for

1114
00:44:11,510 --> 00:44:08,720
that we call them a virus

1115
00:44:13,510 --> 00:44:11,520
what you just said would argue that it

1116
00:44:15,670 --> 00:44:13,520
is not yet proven that there is a common

1117
00:44:17,589 --> 00:44:15,680
origin of viruses but that's a different

1118
00:44:20,150 --> 00:44:17,599
thing than to say it's proven that

1119
00:44:22,710 --> 00:44:20,160
viruses do not have a common origin

1120
00:44:24,630 --> 00:44:22,720
for luca of cells it seems it is proven

1121
00:44:27,349 --> 00:44:24,640
there is a luca if you look at things

1122
00:44:30,870 --> 00:44:27,359
like uh i i see professor gogarten is in

1123
00:44:32,870 --> 00:44:30,880
the chat um his paper is on the on the

1124
00:44:34,390 --> 00:44:32,880
early prokaryotic

1125
00:44:37,750 --> 00:44:34,400
i mean sorry

1126
00:44:40,069 --> 00:44:37,760
hallmark genes that existed preluca

1127
00:44:42,230 --> 00:44:40,079
you know it's it's fairly obvious that

1128
00:44:44,630 --> 00:44:42,240
there must be luca but i'm not sure that

1129
00:44:49,430 --> 00:44:44,640
it's been actually shown that there

1130
00:44:49,440 --> 00:44:53,349
well equally um

1131
00:44:58,470 --> 00:44:54,950
yeah i'm not quite sure how you want me

1132
00:45:00,950 --> 00:44:58,480
to respond to this um because there's

1133
00:45:03,109 --> 00:45:00,960
there's multiple different um

1134
00:45:05,829 --> 00:45:03,119
origins that have been

1135
00:45:08,309 --> 00:45:05,839
deduced from the data

1136
00:45:12,309 --> 00:45:08,319
and linking these different origins to

1137
00:45:13,829 --> 00:45:12,319
one common origin i wouldn't know how to

1138
00:45:15,589 --> 00:45:13,839

go about that

1139
00:45:18,150 --> 00:45:15,599
at this point with the information we

1140
00:45:20,470 --> 00:45:18,160
have

1141
00:45:22,790 --> 00:45:20,480
stop being a devil's advocate here

1142
00:45:24,790 --> 00:45:22,800
there's you know

1143
00:45:27,109 --> 00:45:24,800
there's no data pointing to the idea

1144
00:45:29,030 --> 00:45:27,119
that rna viruses came from dna viruses

1145
00:45:30,630 --> 00:45:29,040
or vice versa so let's just immediately

1146
00:45:33,109 --> 00:45:30,640
assume there's two separate common

1147
00:45:36,790 --> 00:45:33,119
ancestors i i'm

1148
00:45:38,790 --> 00:45:36,800
bringing it up because if we say that

1149
00:45:40,710 --> 00:45:38,800
the question is settled that means

1150
00:45:41,990 --> 00:45:40,720
there's no point in actually doing a

1151
00:45:44,950 --> 00:45:42,000
study on it

1152
00:45:47,270 --> 00:45:44,960
right so if the question isn't settled

1153
00:45:48,950 --> 00:45:47,280
then i think it's could be worth someone

1154
00:45:50,950 --> 00:45:48,960
investigating so we shouldn't shut down

1155
00:45:53,109 --> 00:45:50,960
the line of inquiry unless

1156
00:45:54,790 --> 00:45:53,119
there's evidence that that line should

1157
00:45:56,790 --> 00:45:54,800
just be dead

1158
00:45:58,710 --> 00:45:56,800
i will i'm sorry to jump in here i'll

1159
00:46:00,950 --> 00:45:58,720
talk a little bit about this at the end

1160
00:46:01,750 --> 00:46:00,960
and this is actually might be a good

1161
00:46:03,670 --> 00:46:01,760
uh

1162
00:46:05,910 --> 00:46:03,680
way to sort of you know push off some of

1163
00:46:07,589 --> 00:46:05,920

these discussions um i also have this

1164
00:46:09,990 --> 00:46:07,599
paper which i'll put in the chat here

1165
00:46:11,430 --> 00:46:10,000
the the most recent

1166
00:46:14,150 --> 00:46:11,440
uh

1167
00:46:16,069 --> 00:46:14,160
review by cooper vick dulcin doljan

1168
00:46:17,829 --> 00:46:16,079
kunin which you may or may not agree

1169
00:46:19,829 --> 00:46:17,839
with um but we can talk a little bit

1170
00:46:21,270 --> 00:46:19,839
more about that a little bit later on it

1171
00:46:23,589 --> 00:46:21,280
was uh earlier this year and again i'll

1172
00:46:25,750 --> 00:46:23,599
put a link i don't think we have that in

1173
00:46:28,230 --> 00:46:25,760
our our list here about origins of

1174
00:46:29,190 --> 00:46:28,240
viruses but um i agree with you jason i

1175
00:46:31,190 --> 00:46:29,200
think that

1176
00:46:32,950 --> 00:46:31,200
you def we definitely have to think

1177
00:46:35,589 --> 00:46:32,960
about a lot of these things in a broader

1178
00:46:37,829 --> 00:46:35,599
sense which is exactly why we're doing

1179
00:46:39,910 --> 00:46:37,839
this whole workshop in the first place

1180
00:46:42,470 --> 00:46:39,920
is to get people talking about these

1181
00:46:45,190 --> 00:46:42,480
kinds of things and then hopefully

1182
00:46:47,190 --> 00:46:45,200
convincing various funding agencies that

1183
00:46:49,990 --> 00:46:47,200
it's something that they want to support

1184
00:46:52,309 --> 00:46:50,000
so that's um that i think is sort of the

1185
00:46:54,950 --> 00:46:52,319
the impetus at least that was part of my

1186
00:46:56,790 --> 00:46:54,960
impetus in terms of trying to put this

1187
00:46:59,349 --> 00:46:56,800
whole thing together so um great

1188
00:47:01,670 --> 00:46:59,359

thoughts and um love to hear

1189
00:47:03,349 --> 00:47:01,680
lots of people's input as we as we move

1190
00:47:05,270 --> 00:47:03,359
forward on this

1191
00:47:06,550 --> 00:47:05,280
yeah i think we needed mark krupovic for

1192
00:47:08,630 --> 00:47:06,560
this

1193
00:47:11,190 --> 00:47:08,640
unfortunately mark couldn't make it um

1194
00:47:13,030 --> 00:47:11,200
we there was a miscommunication mia

1195
00:47:14,309 --> 00:47:13,040
called bummer called me at culpa in

1196
00:47:17,270 --> 00:47:14,319
terms of doing that apparently he's in

1197
00:47:20,230 --> 00:47:17,280
the uk right now so yeah

1198
00:47:22,470 --> 00:47:20,240
you can find him there so

1199
00:47:23,270 --> 00:47:22,480
well it is getting quite late here so i

1200
00:47:25,910 --> 00:47:23,280
don't

1201
00:47:27,270 --> 00:47:25,920
think i'll be able to finish the whole

1202
00:47:30,309 --> 00:47:27,280
session

1203
00:47:32,470 --> 00:47:30,319
yes thank you so much by the way for

1204
00:47:34,710 --> 00:47:32,480
for doing this and as you say such a

1205
00:47:38,230 --> 00:47:34,720
late hour it's greatly appreciated oh no

1206
00:47:40,870 --> 00:47:39,670
um

1207
00:47:42,950 --> 00:47:40,880
if there's

1208
00:47:45,589 --> 00:47:42,960
any more questions any more any more

1209
00:47:47,589 --> 00:47:45,599
questions now do we do a a three-minute

1210
00:47:50,230 --> 00:47:47,599
break and let nigel get himself set up

1211
00:47:52,069 --> 00:47:50,240
that sounds like a good deal

1212
00:48:03,990 --> 00:47:52,079
thank you so much evelyn and again

1213
00:48:09,589 --> 00:48:07,670

so nigel are you more or less set up

1214
00:48:10,470 --> 00:48:09,599
uh yes do i just need to push the share

1215
00:48:11,510 --> 00:48:10,480
button

1216
00:48:13,589 --> 00:48:11,520
you should be able to and they'll have

1217
00:48:15,349 --> 00:48:13,599
to pick which screen you want to use and

1218
00:48:17,270 --> 00:48:15,359
then maybe we'll give it a couple of

1219
00:48:18,790 --> 00:48:17,280
minutes just in case people are um

1220
00:48:20,390 --> 00:48:18,800
coming from one place or another to try

1221
00:48:23,349 --> 00:48:20,400
and catch your time i get this message

1222
00:48:26,390 --> 00:48:23,359
host disabled participant screen sharing

1223
00:48:27,670 --> 00:48:26,400
ah then marco needs to help us

1224
00:48:29,829 --> 00:48:27,680
this is where we kept

1225
00:48:42,790 --> 00:48:29,839
you should have permission now nigel

1226
00:48:46,630 --> 00:48:44,950
okay does it how how does that look ken

1227
00:48:49,190 --> 00:48:46,640
it looks great thank you

1228
00:48:51,030 --> 00:48:49,200
good okay

1229
00:48:53,430 --> 00:48:51,040
and i have all these sort of buttons and

1230
00:48:55,190 --> 00:48:53,440
things on the top and and everything

1231
00:48:59,910 --> 00:48:55,200
you don't see those right or do you not

1232
00:48:59,920 --> 00:49:05,190
is there a way i can make them go away

1233
00:49:05,200 --> 00:49:09,510
um you could try dragging it away

1234
00:49:14,470 --> 00:49:12,790
if i if i hide the video panel

1235
00:49:21,430 --> 00:49:14,480
how do i get it back if i click that

1236
00:49:36,230 --> 00:49:23,190
i'll just leave it there it's not it's

1237
00:49:39,670 --> 00:49:38,309
and and i get to talk for 20 minutes

1238
00:49:42,150 --> 00:49:39,680

that's right and then there's a few

1239
00:49:43,030 --> 00:49:42,160
minutes of discussion correct 20 plus

1240
00:49:45,589 --> 00:49:43,040
10.

1241
00:49:46,950 --> 00:49:45,599
so um and if you like i can wave at you

1242
00:49:49,270 --> 00:49:46,960
when you've got about you know five or

1243
00:49:51,670 --> 00:49:49,280
ten minutes to go no it's okay i i i

1244
00:49:52,549 --> 00:49:51,680
sure sure do though i i have a timer

1245
00:49:55,430 --> 00:49:52,559
okay

1246
00:49:56,950 --> 00:49:55,440
perfect so if you're ready to go

1247
00:50:00,630 --> 00:49:56,960
take it away

1248
00:50:01,670 --> 00:50:00,640
is everybody else back from from breaks

1249
00:50:02,470 --> 00:50:01,680
if they aren't that's their problem

1250
00:50:09,190 --> 00:50:02,480
right

1251
00:50:10,710 --> 00:50:09,200
welcome also from uh from england um

1252
00:50:13,750 --> 00:50:10,720
so um

1253
00:50:16,390 --> 00:50:13,760
i am going to uh change gears we're

1254
00:50:20,549 --> 00:50:16,400
going to talk um about the the effects

1255
00:50:22,470 --> 00:50:20,559
of viruses on ecosystems and um

1256
00:50:24,549 --> 00:50:22,480
and essentially what i'm going to try to

1257
00:50:26,549 --> 00:50:24,559
point out is that um

1258
00:50:29,510 --> 00:50:26,559
in at least in the bacterial world

1259
00:50:32,230 --> 00:50:29,520
viruses tend to get a a bad rap uh

1260
00:50:34,790 --> 00:50:32,240
they're regarded as as predators and and

1261
00:50:37,030 --> 00:50:34,800
bad things and uh and generally the word

1262
00:50:39,510 --> 00:50:37,040
virus is synonymous with uh something

1263
00:50:42,950 --> 00:50:39,520

bad and um and what i'm going to try to

1264
00:50:46,069 --> 00:50:42,960
uh to argue is that actually uh viruses

1265
00:50:48,069 --> 00:50:46,079
are a fantastic example um of what um

1266
00:50:51,349 --> 00:50:48,079
ends up being a sort of a multi-level

1267
00:50:53,589 --> 00:50:51,359
selection and that uh and viruses uh

1268
00:50:56,790 --> 00:50:53,599
through their activity can actually

1269
00:50:57,829 --> 00:50:56,800
generate a good uh in an ecosystem help

1270
00:51:00,790 --> 00:50:57,839
it to

1271
00:51:04,230 --> 00:51:00,800
accelerate its evolution and generate uh

1272
00:51:07,270 --> 00:51:04,240
organismal diversity and and indeed even

1273
00:51:09,510 --> 00:51:07,280
a rich uh set of uh ecotypes

1274
00:51:11,109 --> 00:51:09,520
so um let me just uh get through the

1275
00:51:13,430 --> 00:51:11,119
acknowledgement so the beginning ones

1276
00:51:16,710 --> 00:51:13,440
this is work with the hong yang shi who

1277
00:51:19,829 --> 00:51:16,720
is now the academic cylinder in taiwan

1278
00:51:21,430 --> 00:51:19,839
uh this was uh uh we were benefited from

1279
00:51:23,430 --> 00:51:21,440
a lot of discussions with tim rogers

1280
00:51:27,910 --> 00:51:23,440
forest war penny chisholm and paul

1281
00:51:30,150 --> 00:51:27,920
farkowski and supporters through the nai

1282
00:51:32,390 --> 00:51:30,160
and um

1283
00:51:34,230 --> 00:51:32,400
so we're very grateful uh to their

1284
00:51:35,990 --> 00:51:34,240
wonderful support

1285
00:51:39,990 --> 00:51:36,000
so let me start then with a central

1286
00:51:42,470 --> 00:51:40,000
question uh for uh astro for for biology

1287
00:51:44,710 --> 00:51:42,480
that really came from nasa and this is

1288
00:51:47,990 --> 00:51:44,720

quite an inspiring uh document in my

1289
00:51:50,710 --> 00:51:48,000
view the nasa astrology roadmap from uh

1290
00:51:52,470 --> 00:51:50,720
from over 10 years or so ago

1291
00:51:54,870 --> 00:51:52,480
one of the objectives of that was to try

1292
00:51:57,270 --> 00:51:54,880
to understand the co-evolution of

1293
00:51:59,910 --> 00:51:57,280
microbial communities and particularly

1294
00:52:01,990 --> 00:51:59,920
to understand the metabolic and genetic

1295
00:52:03,829 --> 00:52:02,000
interactions in microbial communities

1296
00:52:06,870 --> 00:52:03,839
including viruses

1297
00:52:09,510 --> 00:52:06,880
and and how these have a shaped

1298
00:52:11,750 --> 00:52:09,520
evolution and maybe determined uh major

1299
00:52:14,549 --> 00:52:11,760
geochemical processes and changes on

1300
00:52:17,589 --> 00:52:14,559
earth so typically we tend to think of

1301
00:52:18,470 --> 00:52:17,599
bacteria and viruses as representing uh

1302
00:52:19,910 --> 00:52:18,480
predator

1303
00:52:21,030 --> 00:52:19,920
and prey

1304
00:52:23,270 --> 00:52:21,040
uh uh

1305
00:52:26,470 --> 00:52:23,280
virus has been predator the bacteria

1306
00:52:28,309 --> 00:52:26,480
being the uh being the prey but um but

1307
00:52:30,549 --> 00:52:28,319
i'm i'm going to try to argue that this

1308
00:52:32,390 --> 00:52:30,559
view is a little bit um

1309
00:52:35,109 --> 00:52:32,400
is a little bit too naive and too

1310
00:52:37,750 --> 00:52:35,119
simplistic and that the relationships

1311
00:52:39,990 --> 00:52:37,760
between the organisms

1312
00:52:43,270 --> 00:52:40,000
that you see

1313
00:52:45,030 --> 00:52:43,280

first hide a much more nuanced uh

1314
00:52:46,950 --> 00:52:45,040
dynamic which you can which you can

1315
00:52:49,030 --> 00:52:46,960
understand the other thing i want to

1316
00:52:50,790 --> 00:52:49,040
talk about is the environment as well

1317
00:52:53,109 --> 00:52:50,800
typically when we talk about predator

1318
00:52:54,790 --> 00:52:53,119
prey dynamics we don't really think

1319
00:52:56,710 --> 00:52:54,800
about the environment at least

1320
00:52:59,349 --> 00:52:56,720
ecologists tend not to and we're going

1321
00:53:00,790 --> 00:52:59,359
to say a bit more about that

1322
00:53:04,069 --> 00:53:00,800
on the viral

1323
00:53:06,470 --> 00:53:04,079
side and the the the question is how are

1324
00:53:09,510 --> 00:53:06,480
uh viral and bacterial communities

1325
00:53:12,230 --> 00:53:09,520
stabilized because uh if you have um

1326
00:53:15,030 --> 00:53:12,240
highly diverse microphage communities

1327
00:53:17,510 --> 00:53:15,040
naively and theoretically uh you would

1328
00:53:19,510 --> 00:53:17,520
expect that stability would be a huge

1329
00:53:21,589 --> 00:53:19,520
problem because of the exponential

1330
00:53:22,470 --> 00:53:21,599
growth of phage due to their large birth

1331
00:53:26,230 --> 00:53:22,480
size

1332
00:53:28,390 --> 00:53:26,240
so it's a very it's still a problematic

1333
00:53:30,390 --> 00:53:28,400
uh issue how are such communities

1334
00:53:32,549 --> 00:53:30,400
stabilized uh what controls their

1335
00:53:34,630 --> 00:53:32,559
diversity and of course one simple

1336
00:53:36,150 --> 00:53:34,640
answer is that you have spatial uh

1337
00:53:38,790 --> 00:53:36,160
heterogeneity you have spatial

1338
00:53:40,630 --> 00:53:38,800

variations you have reservoir effects

1339
00:53:43,510 --> 00:53:40,640
and so on and that certainly is going to

1340
00:53:45,670 --> 00:53:43,520
tend to promote coexistence but still

1341
00:53:49,030 --> 00:53:45,680
when you do simple calculations it's

1342
00:53:50,950 --> 00:53:49,040
hard to uh to see how it is that uh

1343
00:53:53,670 --> 00:53:50,960
microbial communities

1344
00:53:55,910 --> 00:53:53,680
can be stable in the presence of viruses

1345
00:53:57,190 --> 00:53:55,920
and how does one modulate the virulence

1346
00:53:57,990 --> 00:53:57,200
of the phage

1347
00:54:09,670 --> 00:53:58,000
to

1348
00:54:12,870 --> 00:54:09,680
what actually is life itself

1349
00:54:15,589 --> 00:54:12,880
now i would uh argue along with everett

1350
00:54:19,990 --> 00:54:15,599
shock and others um that really the

1351
00:54:23,510 --> 00:54:20,000
purpose of life is to uh help planets

1352
00:54:26,069 --> 00:54:23,520
approach um equilibrium and what is the

1353
00:54:29,349 --> 00:54:26,079
basic idea here well the idea here is

1354
00:54:32,950 --> 00:54:29,359
that as our planet is formed uh it it

1355
00:54:35,670 --> 00:54:32,960
cools and and uh and um

1356
00:54:39,990 --> 00:54:35,680
huge numbers of meta-stable states are

1357
00:54:42,710 --> 00:54:40,000
created due to the existence of

1358
00:54:45,589 --> 00:54:42,720
redox gradients uh biochemical radiance

1359
00:54:47,589 --> 00:54:45,599
and so on and so forth and uh and the

1360
00:54:49,750 --> 00:54:47,599
system will only be able to

1361
00:54:52,710 --> 00:54:49,760
approach equilibrium

1362
00:54:54,549 --> 00:54:52,720
when that chemical energy can be uh

1363
00:54:56,710 --> 00:54:54,559

released

1364
00:54:59,109 --> 00:54:56,720
so the idea then is that the way to

1365
00:55:01,270 --> 00:54:59,119
think about the earth is that it is is

1366
00:55:02,470 --> 00:55:01,280
it is essentially um

1367
00:55:03,990 --> 00:55:02,480
a uh

1368
00:55:06,309 --> 00:55:04,000
a a a

1369
00:55:08,630 --> 00:55:06,319
a system like shown on the right here

1370
00:55:10,630 --> 00:55:08,640
where you have a battery which is the

1371
00:55:12,789 --> 00:55:10,640
creation of the various uh chemical

1372
00:55:15,270 --> 00:55:12,799
potential gradients and then what you

1373
00:55:18,630 --> 00:55:15,280
have across that are two uh

1374
00:55:20,870 --> 00:55:18,640
pathways an abiotic pathway and a

1375
00:55:24,069 --> 00:55:20,880
living pathway is shown here as as

1376
00:55:26,870 --> 00:55:24,079
parallel resistors and the idea is that

1377
00:55:29,109 --> 00:55:26,880
the environment provides energy input

1378
00:55:31,750 --> 00:55:29,119
that drives both of these processes the

1379
00:55:34,069 --> 00:55:31,760
abiotic physical chemical processes and

1380
00:55:36,069 --> 00:55:34,079
the living systems want but living

1381
00:55:38,069 --> 00:55:36,079
systems have an advantage because they

1382
00:55:40,230 --> 00:55:38,079
can use novel catalytic pathways to

1383
00:55:42,470 --> 00:55:40,240
accelerate reaction leaks and out

1384
00:55:44,870 --> 00:55:42,480
compete abiotic processes

1385
00:55:46,390 --> 00:55:44,880
in order to release these of these free

1386
00:55:48,309 --> 00:55:46,400
energy gradients

1387
00:55:51,190 --> 00:55:48,319
now how is it the living systems can do

1388
00:55:53,349 --> 00:55:51,200

that whereas abiotic systems cannot well

1389
00:55:55,430 --> 00:55:53,359
the idea is that living systems discover

1390
00:55:57,430 --> 00:55:55,440
these pathways through the flow of

1391
00:56:00,470 --> 00:55:57,440
information so in other words

1392
00:56:02,710 --> 00:56:00,480
evolutionary mechanisms impact ecology

1393
00:56:04,789 --> 00:56:02,720
and the global planetary environment and

1394
00:56:06,230 --> 00:56:04,799
what i want to talk about mostly for

1395
00:56:09,270 --> 00:56:06,240
most of this talk

1396
00:56:10,950 --> 00:56:09,280
is the is the special role of horizontal

1397
00:56:13,750 --> 00:56:10,960
gene transfer as an evolutionary

1398
00:56:16,630 --> 00:56:13,760
mechanism that is unusually effective in

1399
00:56:18,390 --> 00:56:16,640
discovering novelty and allowing uh

1400
00:56:19,270 --> 00:56:18,400
systems to

1401
00:56:21,109 --> 00:56:19,280
um

1402
00:56:22,870 --> 00:56:21,119
to to to to

1403
00:56:24,870 --> 00:56:22,880
equilibrate their environment at least

1404
00:56:27,589 --> 00:56:24,880
to tend to equilibrate it now the

1405
00:56:29,910 --> 00:56:27,599
organism that's the style of this talk

1406
00:56:32,870 --> 00:56:29,920
is the uh is proclaimed prochlorococcus

1407
00:56:34,950 --> 00:56:32,880
it's a tiny marine cyanobacteria and

1408
00:56:37,270 --> 00:56:34,960
they're phased uh and i'm going to use

1409
00:56:39,750 --> 00:56:37,280
that as an example of how life uses

1410
00:56:42,549 --> 00:56:39,760
information flow to thrive and occupy

1411
00:56:46,390 --> 00:56:42,559
new planetary feedback

1412
00:56:48,150 --> 00:56:46,400
uh planetary nations now ecosystems uh

1413
00:56:50,950 --> 00:56:48,160

can be can be thought of in in the

1414
00:56:53,589 --> 00:56:50,960
following way and on one level one has

1415
00:56:55,270 --> 00:56:53,599
thermodynamics which is the energy input

1416
00:56:57,430 --> 00:56:55,280
that drives everything that goes on

1417
00:56:59,349 --> 00:56:57,440
inside the ecosystem but then when you

1418
00:57:01,190 --> 00:56:59,359
drill down a bit further into ecology

1419
00:57:03,270 --> 00:57:01,200
you'll find that it has two aspects

1420
00:57:05,430 --> 00:57:03,280
currently not well unified one is

1421
00:57:07,990 --> 00:57:05,440
population dynamics and the other is

1422
00:57:09,750 --> 00:57:08,000
metabolism and right now we don't really

1423
00:57:11,990 --> 00:57:09,760
have a good way of unifying both of

1424
00:57:14,870 --> 00:57:12,000
these descriptions and then the third

1425
00:57:16,309 --> 00:57:14,880
aspect is information flow

1426
00:57:20,150 --> 00:57:16,319
so what i'm going to talk about in the

1427
00:57:22,230 --> 00:57:20,160
marine micro virus ecosystem is the the

1428
00:57:24,710 --> 00:57:22,240
the environment is the photon gradient

1429
00:57:27,030 --> 00:57:24,720
the population dynamics is cyanobacteria

1430
00:57:29,030 --> 00:57:27,040
and cyanophage and the predation and the

1431
00:57:31,510 --> 00:57:29,040
information flow is the mechanisms of

1432
00:57:33,109 --> 00:57:31,520
horizontal gene transfer

1433
00:57:35,670 --> 00:57:33,119
so what i'm going to show you now is

1434
00:57:37,190 --> 00:57:35,680
that through horizontal gene transfer

1435
00:57:40,069 --> 00:57:37,200
through the collective dynamics of

1436
00:57:41,910 --> 00:57:40,079
cyanobacteria and cyanophage one can

1437
00:57:44,950 --> 00:57:41,920
understand the range expansion and knee

1438
00:57:45,990 --> 00:57:44,960

stratification of a prochlorococcus and

1439
00:57:47,990 --> 00:57:46,000
what we're going to see is that

1440
00:57:50,789 --> 00:57:48,000
predation really stabilizes the

1441
00:57:52,150 --> 00:57:50,799
ecosystem and allows it to find it its

1442
00:57:54,549 --> 00:57:52,160
nations

1443
00:57:56,549 --> 00:57:54,559
so again the usual assumption then is

1444
00:57:59,109 --> 00:57:56,559
that bacteria are prey phage of

1445
00:58:00,630 --> 00:57:59,119
predators but uh what we're going to see

1446
00:58:02,870 --> 00:58:00,640
is that the relationship is much more

1447
00:58:05,349 --> 00:58:02,880
complex there's a co-evolutionary arms

1448
00:58:07,829 --> 00:58:05,359
race between the bacteria and the phage

1449
00:58:09,990 --> 00:58:07,839
through the cell surface receptors which

1450
00:58:12,230 --> 00:58:10,000
the tail sequences of the phage

1451
00:58:15,030 --> 00:58:12,240
lock into and so what we're going to see

1452
00:58:17,430 --> 00:58:15,040
is that there's a you get a diversity of

1453
00:58:19,670 --> 00:58:17,440
bacterial genome sequences as a result

1454
00:58:22,470 --> 00:58:19,680
of that co-evolutionary arms race at the

1455
00:58:24,309 --> 00:58:22,480
same time horizontal gene transfer

1456
00:58:26,470 --> 00:58:24,319
between the bacteria which is mediated

1457
00:58:28,390 --> 00:58:26,480
by the phage gives rise to genetic

1458
00:58:30,630 --> 00:58:28,400
diversity

1459
00:58:32,870 --> 00:58:30,640
the prochlorococcus is the world's most

1460
00:58:35,270 --> 00:58:32,880
abundant photosynthetic organism it has

1461
00:58:39,510 --> 00:58:35,280
a highly screening scrublines genome

1462
00:58:41,750 --> 00:58:39,520
1700 genes in about 1.65 megabases it's

1463
00:58:44,150 --> 00:58:41,760

very dilute in the ocean it fixes the

1464
00:58:46,950 --> 00:58:44,160
same amount of carbon as global cropland

1465
00:58:49,589 --> 00:58:46,960
and very importantly it has no real

1466
00:58:50,950 --> 00:58:49,599
immune systems no crispr there's no

1467
00:58:53,910 --> 00:58:50,960
prophages

1468
00:58:55,750 --> 00:58:53,920
very limited restriction modification so

1469
00:58:58,150 --> 00:58:55,760
the genetic diversity

1470
00:59:00,950 --> 00:58:58,160
arises from genes acquired by horizontal

1471
00:59:02,309 --> 00:59:00,960
gene transfer code for cell surface

1472
00:59:04,710 --> 00:59:02,319
receptors

1473
00:59:09,030 --> 00:59:04,720
the the the diversity is huge it has a

1474
00:59:11,190 --> 00:59:09,040
the pangenome is 84 000 genes so uh four

1475
00:59:13,510 --> 00:59:11,200
or five times bigger than the human

1476
00:59:16,230 --> 00:59:13,520
genome and especially there is niche

1477
00:59:18,309 --> 00:59:16,240
stratification so over here you can see

1478
00:59:20,230 --> 00:59:18,319
as a function of depth in the ocean

1479
00:59:23,510 --> 00:59:20,240
light intensity this is the euphotic

1480
00:59:25,829 --> 00:59:23,520
zone and these uh these two curves in

1481
00:59:28,789 --> 00:59:25,839
blue and yellow show you the population

1482
00:59:31,510 --> 00:59:28,799
and logarithms of two particular

1483
00:59:34,549 --> 00:59:31,520
ecotypes a low light adapted strain and

1484
00:59:37,910 --> 00:59:36,309
strain there are other eco types here

1485
00:59:39,829 --> 00:59:37,920
but this coarse grained picture is

1486
00:59:40,789 --> 00:59:39,839
enough for what i want to talk about

1487
00:59:42,390 --> 00:59:40,799
and then down here you have the

1488
00:59:45,349 --> 00:59:42,400

cyanophages which are lytic

1489
00:59:47,750 --> 00:59:45,359
double-stranded dna phages usually in t4

1490
00:59:49,510 --> 00:59:47,760
or t7 groups and the most important

1491
00:59:51,510 --> 00:59:49,520
thing about them is they carry

1492
00:59:56,870 --> 00:59:51,520
photosynthesis genes

1493
01:00:01,190 --> 00:59:56,880
have uh uh our photosystem two genes and

1494
01:00:03,349 --> 01:00:01,200
what um uh matt sullivan at al uh at al

1495
01:00:05,990 --> 01:00:03,359
being the penny chisholm lab uh

1496
01:00:08,470 --> 01:00:06,000
discovered uh over 15 years or so around

1497
01:00:11,030 --> 01:00:08,480
about 15 years or so ago was that these

1498
01:00:12,870 --> 01:00:11,040
genes have been swapped back and forth

1499
01:00:15,190 --> 01:00:12,880
uh over evolutionary time between the

1500
01:00:17,270 --> 01:00:15,200
cyanobacteria and the sand ephedras and

1501
01:00:19,109 --> 01:00:17,280
we know that from from looking at the at

1502
01:00:21,190 --> 01:00:19,119
the trees of these genes

1503
01:00:22,390 --> 01:00:21,200
the other thing about this is that the

1504
01:00:25,030 --> 01:00:22,400
phage

1505
01:00:26,950 --> 01:00:25,040
benefits from the photosynthesis genes

1506
01:00:29,270 --> 01:00:26,960
in other words they can actually use

1507
01:00:32,309 --> 01:00:29,280
them and they use them because they have

1508
01:00:33,670 --> 01:00:32,319
a higher burst size during license and

1509
01:00:36,309 --> 01:00:33,680
this has been shown by a number of

1510
01:00:37,990 --> 01:00:36,319
studies especially from uh debbie

1511
01:00:40,710 --> 01:00:38,000
lindell

1512
01:00:43,589 --> 01:00:40,720
so what did solomon at how conclude

1513
01:00:45,910 --> 01:00:43,599

after they they uncovered this they they

1514
01:00:48,309 --> 01:00:45,920
hypothesized that host-like genes

1515
01:00:50,549 --> 01:00:48,319
acquired by phages undergo a period of

1516
01:00:52,470 --> 01:00:50,559
diversification in the phase genomes and

1517
01:00:55,109 --> 01:00:52,480
then serve as a genetic reservoir for

1518
01:00:57,589 --> 01:00:55,119
their hosts and then what happens is

1519
01:00:59,510 --> 01:00:57,599
that genetic strains across these pools

1520
01:01:02,150 --> 01:00:59,520
leads to evolutionary change for both

1521
01:01:04,390 --> 01:01:02,160
the host and the phage and so the idea

1522
01:01:06,789 --> 01:01:04,400
is you've got microphage interactions

1523
01:01:08,870 --> 01:01:06,799
create a global reservoir of genes that

1524
01:01:11,510 --> 01:01:08,880
benefit both the microbes and the phages

1525
01:01:13,349 --> 01:01:11,520
but exactly how that comes about

1526
01:01:16,309 --> 01:01:13,359
has never been made clear and that's

1527
01:01:18,789 --> 01:01:16,319
what i want to explain to you next

1528
01:01:20,870 --> 01:01:18,799
so what we want to do is build a minimal

1529
01:01:22,789 --> 01:01:20,880
model not a not a model cursed with

1530
01:01:24,950 --> 01:01:22,799
excessive realism but a minimal

1531
01:01:27,750 --> 01:01:24,960
stochastic model of the interaction

1532
01:01:31,990 --> 01:01:27,760
between bacteria and virus and our goal

1533
01:01:34,309 --> 01:01:32,000
is to ask is it possible that one can

1534
01:01:36,630 --> 01:01:34,319
derive the existence of a collective

1535
01:01:39,109 --> 01:01:36,640
mutually beneficial state between

1536
01:01:41,349 --> 01:01:39,119
microbes and viruses that emerges

1537
01:01:43,510 --> 01:01:41,359
through these antagonistic interactions

1538
01:01:45,589 --> 01:01:43,520

of predation and the answer is yes and

1539
01:01:47,510 --> 01:01:45,599
i'm going to show you how

1540
01:01:50,549 --> 01:01:47,520
now i'm going to start by giving you the

1541
01:01:52,390 --> 01:01:50,559
basic idea and then i'm going to rapidly

1542
01:01:55,029 --> 01:01:52,400
go through a sequence of models very

1543
01:01:56,789 --> 01:01:55,039
quickly just to show you uh the results

1544
01:01:58,069 --> 01:01:56,799
of calculations that we did but i'm not

1545
01:01:59,829 --> 01:01:58,079
going to have time to go into great

1546
01:02:02,150 --> 01:01:59,839
detail so i want to first of all give

1547
01:02:04,150 --> 01:02:02,160
you the basic idea first of all so the

1548
01:02:06,630 --> 01:02:04,160
idea is this we're drawing here the

1549
01:02:09,109 --> 01:02:06,640
bacteriosphere and the vivosphere plus

1550
01:02:12,950 --> 01:02:09,119
and minus refer to genes that are good

1551
01:02:14,309 --> 01:02:12,960
or bad beneficial or deleterious and uh

1552
01:02:16,390 --> 01:02:14,319
we're just we're not saying what the

1553
01:02:17,270 --> 01:02:16,400
genes do for the moment but but never

1554
01:02:19,349 --> 01:02:17,280
mind

1555
01:02:21,829 --> 01:02:19,359
so the idea is first of all that the

1556
01:02:23,750 --> 01:02:21,839
viruses gain genes from the bacteria

1557
01:02:25,670 --> 01:02:23,760
through a horizontal gene transfer so

1558
01:02:28,069 --> 01:02:25,680
that's what's shown in this first

1559
01:02:30,630 --> 01:02:28,079
the first diagram here now the basic

1560
01:02:32,230 --> 01:02:30,640
point is that in the virus sphere the

1561
01:02:35,990 --> 01:02:32,240
viruses

1562
01:02:38,950 --> 01:02:36,000
have more mutations and so what happens

1563
01:02:41,109 --> 01:02:38,960

is now you have a greater diversity of

1564
01:02:42,230 --> 01:02:41,119
these genes in the virus sphere shown

1565
01:02:44,309 --> 01:02:42,240
here

1566
01:02:45,910 --> 01:02:44,319
no reason to be positive or negative

1567
01:02:48,630 --> 01:02:45,920
particularly just

1568
01:02:50,789 --> 01:02:48,640
just randomly mutating

1569
01:02:53,270 --> 01:02:50,799
and those genes can be transferred back

1570
01:02:55,510 --> 01:02:53,280
uh to the bacteria as i've shown here

1571
01:02:57,670 --> 01:02:55,520
now the bacteria on the other hand

1572
01:02:59,990 --> 01:02:57,680
have a streamlined genome and so they

1573
01:03:02,710 --> 01:03:00,000
have a huge selection pressure only to

1574
01:03:05,270 --> 01:03:02,720
keep the beneficial genes and so those

1575
01:03:06,549 --> 01:03:05,280
get enriched in the bacterial population

1576
01:03:08,950 --> 01:03:06,559
as shown

1577
01:03:10,630 --> 01:03:08,960
at the bottom here and then when the

1578
01:03:12,470 --> 01:03:10,640
process starts again those get

1579
01:03:14,710 --> 01:03:12,480
horizontally transferred into the virus

1580
01:03:16,950 --> 01:03:14,720
sphere those genes

1581
01:03:19,190 --> 01:03:16,960
those enriched genes those better genes

1582
01:03:22,710 --> 01:03:19,200
are now back in the biosphere and so the

1583
01:03:25,430 --> 01:03:22,720
viruses are benefiting from the bacteria

1584
01:03:27,270 --> 01:03:25,440
that have filtered the the good genes

1585
01:03:30,150 --> 01:03:27,280
now this is actually important because

1586
01:03:32,710 --> 01:03:30,160
the viruses uh by having rapid mutation

1587
01:03:35,430 --> 01:03:32,720
are prone to muller's ratchet and so

1588
01:03:38,789 --> 01:03:35,440

might be expected to rapidly degrade but

1589
01:03:41,029 --> 01:03:38,799
the bacteria by being a filter provide a

1590
01:03:43,349 --> 01:03:41,039
way of filtering out the good genes and

1591
01:03:46,309 --> 01:03:43,359
preserving uh the healthy state of the

1592
01:03:48,150 --> 01:03:46,319
biosphere from the synthesis genes

1593
01:03:49,990 --> 01:03:48,160
now i'm going to very quickly very

1594
01:03:51,829 --> 01:03:50,000
quickly walk through a sequence of molds

1595
01:03:54,069 --> 01:03:51,839
i don't expect you to follow every step

1596
01:03:55,829 --> 01:03:54,079
here i think the last steps are the ones

1597
01:03:58,470 --> 01:03:55,839
that are the most important i just want

1598
01:03:59,990 --> 01:03:58,480
to show you how horizontal gene transfer

1599
01:04:02,230 --> 01:04:00,000
works and how it can stabilize

1600
01:04:03,990 --> 01:04:02,240
ecosystems and the first model i'm just

1601
01:04:05,829 --> 01:04:04,000
going to not talk about photosynthesis

1602
01:04:07,910 --> 01:04:05,839
genes at all i'm going to talk about a

1603
01:04:10,069 --> 01:04:07,920
mythical gene that we'll call a barrier

1604
01:04:12,549 --> 01:04:10,079
to predation gene and in these minimal

1605
01:04:15,270 --> 01:04:12,559
models what we say is a gene has a

1606
01:04:18,069 --> 01:04:15,280
particular number associated to it and

1607
01:04:19,990 --> 01:04:18,079
if a bacteria and a phage meet the

1608
01:04:22,630 --> 01:04:20,000
entity with the bigger

1609
01:04:24,309 --> 01:04:22,640
number is the one that survives so for

1610
01:04:25,750 --> 01:04:24,319
example if a virus with a bigger

1611
01:04:27,750 --> 01:04:25,760
predation gene

1612
01:04:29,750 --> 01:04:27,760
value meets a bacteria with a smaller

1613
01:04:31,670 --> 01:04:29,760

one then the bacteria is killed and

1614
01:04:34,150 --> 01:04:31,680
lysed and so on if on the other hand the

1615
01:04:36,630 --> 01:04:34,160
bacterium has a bigger value of the gene

1616
01:04:38,309 --> 01:04:36,640
then the virus attack is ineffective so

1617
01:04:39,430 --> 01:04:38,319
that's the way one kind of tends to

1618
01:04:42,230 --> 01:04:39,440
model

1619
01:04:43,990 --> 01:04:42,240
bacteria virus interactions

1620
01:04:45,430 --> 01:04:44,000
so one writes these down and i'm not

1621
01:04:47,829 --> 01:04:45,440
going to go through these equations in

1622
01:04:49,589 --> 01:04:47,839
detail as essentially like chemical

1623
01:04:51,349 --> 01:04:49,599
reactions in between

1624
01:04:53,589 --> 01:04:51,359
the different bacteria and the different

1625
01:04:55,750 --> 01:04:53,599
phage the most important thing to notice

1626
01:04:57,990 --> 01:04:55,760
here is that these are models these are

1627
01:05:00,710 --> 01:04:58,000
individual level models and what you

1628
01:05:04,230 --> 01:05:00,720
find is that if you have just a regular

1629
01:05:07,510 --> 01:05:04,240
lot caval tail model what happens is the

1630
01:05:09,990 --> 01:05:07,520
bacteria will make phage

1631
01:05:12,549 --> 01:05:10,000
and here the the population dynamics and

1632
01:05:13,349 --> 01:05:12,559
what eventually happens is the the phage

1633
01:05:15,910 --> 01:05:13,359
um

1634
01:05:17,990 --> 01:05:15,920
uh will growth because of their burst

1635
01:05:20,230 --> 01:05:18,000
size and eventually they will collapse

1636
01:05:22,789 --> 01:05:20,240
the population so that's the first

1637
01:05:25,190 --> 01:05:22,799
paradox so then what happens is we

1638
01:05:27,270 --> 01:05:25,200

consider a second model where now we

1639
01:05:29,910 --> 01:05:27,280
allow there to be a mutations both in

1640
01:05:31,589 --> 01:05:29,920
the bacteria and in the phage so what

1641
01:05:34,230 --> 01:05:31,599
i'm showing you in this bottom panel on

1642
01:05:36,470 --> 01:05:34,240
the left here is the value of the of

1643
01:05:38,630 --> 01:05:36,480
this predation gene the barrier to

1644
01:05:41,510 --> 01:05:38,640
predation gene high

1645
01:05:43,750 --> 01:05:41,520
means a good uh value and this is time

1646
01:05:45,270 --> 01:05:43,760
along the horizontal axis and what you

1647
01:05:48,390 --> 01:05:45,280
see here is that once you switch on

1648
01:05:50,789 --> 01:05:48,400
mutation in both bacteria and phage they

1649
01:05:53,589 --> 01:05:50,799
co-evolve together and the population

1650
01:05:56,150 --> 01:05:53,599
becomes stabilized both the bacteria so

1651
01:05:58,150 --> 01:05:56,160
the bacteria and the phage

1652
01:06:00,150 --> 01:05:58,160
have populations that persist and they

1653
01:06:01,910 --> 01:06:00,160
don't crash

1654
01:06:03,670 --> 01:06:01,920
now what happens of course though is

1655
01:06:05,910 --> 01:06:03,680
that the viral mutation rate is much

1656
01:06:07,270 --> 01:06:05,920
greater than the bacterial mutation rate

1657
01:06:10,069 --> 01:06:07,280
and so even if you have this

1658
01:06:12,789 --> 01:06:10,079
co-evolutionary arms race eventually

1659
01:06:13,750 --> 01:06:12,799
the population becomes unstable as shown

1660
01:06:15,510 --> 01:06:13,760
here

1661
01:06:17,750 --> 01:06:15,520
so now what you do is switch on

1662
01:06:20,230 --> 01:06:17,760
horizontal gene transfer and the precise

1663
01:06:22,950 --> 01:06:20,240

mechanism for that is is not really

1664
01:06:25,670 --> 01:06:22,960
important for these uh calculations but

1665
01:06:27,430 --> 01:06:25,680
here we used a generalized transduction

1666
01:06:29,829 --> 01:06:27,440
and what happens then is that by

1667
01:06:34,549 --> 01:06:29,839
swapping the genes back and forth that

1668
01:06:36,549 --> 01:06:34,559
allows the bacteria to uh to receive

1669
01:06:40,390 --> 01:06:36,559
a revolved genes from the phage and so

1670
01:06:41,910 --> 01:06:40,400
the population then becomes stabilized

1671
01:06:48,390 --> 01:06:41,920
so um

1672
01:06:50,230 --> 01:06:48,400
gene transfer ensures a persistent

1673
01:06:51,990 --> 01:06:50,240
coexisting state

1674
01:06:54,390 --> 01:06:52,000
between bacterium phase now i skipped

1675
01:06:56,870 --> 01:06:54,400
over the slide by accident and and this

1676
01:06:58,789 --> 01:06:56,880
here is meant to be what happens when

1677
01:07:00,950 --> 01:06:58,799
the genes that you're swapping are now

1678
01:07:03,430 --> 01:07:00,960
photosynthesis genes which can be used

1679
01:07:05,430 --> 01:07:03,440
both by the bacteria and the phage so

1680
01:07:07,589 --> 01:07:05,440
once again without horizontal gene

1681
01:07:09,829 --> 01:07:07,599
transfer these the population is

1682
01:07:13,829 --> 01:07:09,839
unstable but with horizontal gene

1683
01:07:16,069 --> 01:07:13,839
transfer the populations get stabilized

1684
01:07:18,150 --> 01:07:16,079
okay now the problem is this as i said

1685
01:07:20,549 --> 01:07:18,160
before uh the

1686
01:07:23,349 --> 01:07:20,559
the genes are not equally good or bad

1687
01:07:25,670 --> 01:07:23,359
typically genes are

1688
01:07:27,750 --> 01:07:25,680

mutations are neutral or deleterious and

1689
01:07:30,150 --> 01:07:27,760
so the probability distribution for the

1690
01:07:31,510 --> 01:07:30,160
quality of the gene the score

1691
01:07:32,230 --> 01:07:31,520
is going to be

1692
01:07:59,029 --> 01:07:32,240
a

1693
01:08:00,549 --> 01:07:59,039
this effect

1694
01:08:02,630 --> 01:08:00,559
you see very quickly that the

1695
01:08:05,510 --> 01:08:02,640
populations would collapse but once

1696
01:08:07,190 --> 01:08:05,520
again horizontal gene transfer prevents

1697
01:08:10,150 --> 01:08:07,200
muller's ratchet by allowing the

1698
01:08:12,950 --> 01:08:10,160
bacteria to filter back to the viruses

1699
01:08:14,630 --> 01:08:12,960
the the genes with a with beneficial

1700
01:08:16,149 --> 01:08:14,640
values

1701
01:08:18,229 --> 01:08:16,159
so what happens is that when you look at

1702
01:08:21,269 --> 01:08:18,239
the phase diagram of the system i'm

1703
01:08:23,510 --> 01:08:21,279
sorry you look at the fraction of

1704
01:08:25,829 --> 01:08:23,520
of beneficial photosynthesis genes in

1705
01:08:28,470 --> 01:08:25,839
the population as a function of the

1706
01:08:30,789 --> 01:08:28,480
horizontal gene transfer rate what you

1707
01:08:33,030 --> 01:08:30,799
find is this at small horizontal

1708
01:08:34,630 --> 01:08:33,040
transfer rate the transfer is not strong

1709
01:08:36,070 --> 01:08:34,640
enough to supply good genes to beat

1710
01:08:37,669 --> 01:08:36,080
mueller's ratchet and the system is

1711
01:08:40,070 --> 01:08:37,679
unstable

1712
01:08:42,709 --> 01:08:40,080
for intermediate values horizontal gene

1713
01:08:44,709 --> 01:08:42,719

transfer balances moles ratchet and then

1714
01:08:47,269 --> 01:08:44,719
for larger values of horizontal gene

1715
01:08:49,430 --> 01:08:47,279
transfer all the phages have transferred

1716
01:08:51,669 --> 01:08:49,440
have got all the good genes so further

1717
01:08:53,430 --> 01:08:51,679
horizontal gene transfer now starts to

1718
01:08:55,110 --> 01:08:53,440
transfer bad genes and the system

1719
01:08:57,510 --> 01:08:55,120
collapses again

1720
01:08:59,669 --> 01:08:57,520
so what you what you have here is a

1721
01:09:01,749 --> 01:08:59,679
collective state if you turn off the

1722
01:09:04,309 --> 01:09:01,759
horizontal gene transfer either from

1723
01:09:06,550 --> 01:09:04,319
bacteria to virus or virus to bacteria

1724
01:09:09,590 --> 01:09:06,560
and what you see in this simulation is

1725
01:09:12,309 --> 01:09:09,600
that the uh the system very quickly

1726
01:09:15,189 --> 01:09:12,319
degrades and is not stable

1727
01:09:16,550 --> 01:09:15,199
now the last the last uh part of this is

1728
01:09:18,870 --> 01:09:16,560
what happens when you do this in a

1729
01:09:21,189 --> 01:09:18,880
photic gradient so now the birth rate of

1730
01:09:24,070 --> 01:09:21,199
the bacteria is a function of the photon

1731
01:09:25,990 --> 01:09:24,080
density and what we're going to look at

1732
01:09:28,470 --> 01:09:26,000
is what happens when we start with a

1733
01:09:30,550 --> 01:09:28,480
population of low light adapted bacteria

1734
01:09:32,070 --> 01:09:30,560
and we ask what happens as the system

1735
01:09:33,990 --> 01:09:32,080
evolves in time

1736
01:09:37,189 --> 01:09:34,000
so let's see so we're going to run the

1737
01:09:40,070 --> 01:09:37,199
movie so down here uh at the bottom you

1738
01:09:41,829 --> 01:09:40,080

see the uh the bacteria the red of the

1739
01:09:42,950 --> 01:09:41,839
viruses the yellow of the low light

1740
01:09:45,110 --> 01:09:42,960
adapted

1741
01:09:47,189 --> 01:09:45,120
bacteria they're trying to evolve

1742
01:09:48,789 --> 01:09:47,199
towards the light but in the highlight

1743
01:09:51,829 --> 01:09:48,799
adapted uh

1744
01:09:53,749 --> 01:09:51,839
region the light intensity is deadly and

1745
01:09:55,669 --> 01:09:53,759
the frequency of light is different from

1746
01:09:59,189 --> 01:09:55,679
the low light adapted ones so the

1747
01:10:01,669 --> 01:09:59,199
bacteria are forced to uh mutate very

1748
01:10:04,550 --> 01:10:01,679
rapidly in order to be able to occupy

1749
01:10:07,270 --> 01:10:04,560
the highlight adapted zone and they do

1750
01:10:10,310 --> 01:10:07,280
that through the viral transfer and now

1751
01:10:12,709 --> 01:10:10,320
what you can see is the emergence of a

1752
01:10:15,270 --> 01:10:12,719
uh a

1753
01:10:17,830 --> 01:10:15,280
highlight adapted ecotype moved down up

1754
01:10:21,270 --> 01:10:17,840
here near the top of the ocean and a low

1755
01:10:22,790 --> 01:10:21,280
light adapted one near the bottom

1756
01:10:24,870 --> 01:10:22,800
so what i've shown you then is that the

1757
01:10:27,510 --> 01:10:24,880
bacteria and the phage actually help

1758
01:10:29,510 --> 01:10:27,520
each other you have collective behavior

1759
01:10:32,870 --> 01:10:29,520
despite there being predation the

1760
01:10:36,950 --> 01:10:35,030
the bacteria have a slow mutation rate

1761
01:10:38,790 --> 01:10:36,960
compared to that of the phage through

1762
01:10:40,870 --> 01:10:38,800
horizontal gene transfer they benefit

1763
01:10:43,189 --> 01:10:40,880

from the phage high mutation rate but at

1764
01:10:44,790 --> 01:10:43,199
the cost of predation the phage have a

1765
01:10:46,950 --> 01:10:44,800
fast mutation rate compared to the

1766
01:10:48,950 --> 01:10:46,960
bacteria they would deteriorate by

1767
01:10:51,350 --> 01:10:48,960
mullis ratchet but they can benefit from

1768
01:10:53,430 --> 01:10:51,360
the bacteria's low mutation rate and

1769
01:10:55,669 --> 01:10:53,440
they can acquire evolved photosynthesis

1770
01:10:58,229 --> 01:10:55,679
genes that prevent the degradation of

1771
01:11:00,550 --> 01:10:58,239
the phage so this collective phenomenon

1772
01:11:02,870 --> 01:11:00,560
creates a huge pan genome and leads to

1773
01:11:05,910 --> 01:11:02,880
range expansion niche uh

1774
01:11:08,149 --> 01:11:05,920
and each stratification so this minimal

1775
01:11:10,550 --> 01:11:08,159
model explains all of these facts about

1776
01:11:12,310 --> 01:11:10,560
folklore caucus in one fell swoop it

1777
01:11:15,110 --> 01:11:12,320
tells you the role of the highly

1778
01:11:17,110 --> 01:11:15,120
streamlined streamlines genome it tells

1779
01:11:19,189 --> 01:11:17,120
you that the fact that there are no

1780
01:11:20,870 --> 01:11:19,199
crispr or prophages means that the

1781
01:11:22,870 --> 01:11:20,880
bacteria are balancing the risk of

1782
01:11:25,270 --> 01:11:22,880
predation with the benefit they get from

1783
01:11:28,149 --> 01:11:25,280
horizontal gene transfer you get a huge

1784
01:11:30,870 --> 01:11:28,159
pan genome due to the co-evolutionary

1785
01:11:32,709 --> 01:11:30,880
arms race and niche stratification

1786
01:11:34,950 --> 01:11:32,719
emerges from the viral mediated

1787
01:11:36,709 --> 01:11:34,960
utilization of the photon gradient and

1788
01:11:38,709 --> 01:11:36,719

the end result of that is that when you

1789
01:11:40,470 --> 01:11:38,719
look at the cyanophage they carry

1790
01:11:42,390 --> 01:11:40,480
photosynthesis genes because they

1791
01:11:45,350 --> 01:11:42,400
benefit from the improved photosynthesis

1792
01:11:47,270 --> 01:11:45,360
genes for a larger birth size

1793
01:11:49,270 --> 01:11:47,280
so this is uh this is what i've told you

1794
01:11:51,030 --> 01:11:49,280
i don't need to go through it again

1795
01:11:53,669 --> 01:11:51,040
what i've shown you then to take her

1796
01:11:55,830 --> 01:11:53,679
message is that an ecosystem can evolve

1797
01:11:57,669 --> 01:11:55,840
to occupying niches where it can utilize

1798
01:12:00,149 --> 01:11:57,679
the energy source provided by a photon

1799
01:12:02,229 --> 01:12:00,159
gradient and this happens due to

1800
01:12:04,149 --> 01:12:02,239
information flow and mediated by

1801
01:12:05,910 --> 01:12:04,159
horizontal gene transfer

1802
01:12:07,830 --> 01:12:05,920
so let me finish and thank my

1803
01:12:08,630 --> 01:12:07,840
collaborator hong yan

1804
01:12:11,030 --> 01:12:08,640
thank

1805
01:12:27,510 --> 01:12:11,040
my funding source and

1806
01:12:27,520 --> 01:12:32,709
we're going to open it up for questions

1807
01:12:36,390 --> 01:12:34,070
i know we have a couple in the group

1808
01:12:39,430 --> 01:12:36,400
chat so if no one has any burning ones

1809
01:12:41,189 --> 01:12:39,440
i'll repeat some from the group chat

1810
01:12:44,310 --> 01:12:41,199
report

1811
01:12:45,510 --> 01:12:44,320
um one comment was made um in regards to

1812
01:12:48,870 --> 01:12:45,520
what you were talking about before the

1813
01:12:51,189 --> 01:12:48,880

benefits of the bacterium and the virus

1814
01:12:53,669 --> 01:12:51,199
saying that um there's not a clear

1815
01:12:55,189 --> 01:12:53,679
benefit of the phage that we see it

1816
01:13:00,070 --> 01:12:55,199
means that they have to acquire the same

1817
01:13:04,390 --> 01:13:01,990
the the the uh

1818
01:13:07,110 --> 01:13:04,400
the the phage are acquiring

1819
01:13:08,870 --> 01:13:07,120
uh the for the photosynthesis genes back

1820
01:13:10,149 --> 01:13:08,880
and forth from the from the bacteria

1821
01:13:11,990 --> 01:13:10,159
that that's right they are being

1822
01:13:13,430 --> 01:13:12,000
transferred multiple times there's this

1823
01:13:15,750 --> 01:13:13,440
phylogenetic evidence that this has

1824
01:13:18,390 --> 01:13:15,760
happened in in folklore caucus

1825
01:13:20,630 --> 01:13:18,400
and and their phages and and and that's

1826
01:13:21,830 --> 01:13:20,640
what this model uh would predict would

1827
01:13:23,110 --> 01:13:21,840
happen

1828
01:13:25,030 --> 01:13:23,120
and it by the way

1829
01:13:26,229 --> 01:13:25,040
the the ocean system is not the only

1830
01:13:28,790 --> 01:13:26,239
system where this could happen this

1831
01:13:32,790 --> 01:13:28,800
could happen in the lithic communities

1832
01:13:34,790 --> 01:13:32,800
as well it could happen in salt um

1833
01:13:37,430 --> 01:13:34,800
in in salt um

1834
01:13:40,390 --> 01:13:37,440
stratified bacterial communities uh salt

1835
01:13:42,430 --> 01:13:40,400
salt maps and so on so so any any

1836
01:13:45,910 --> 01:13:42,440
community where you have um

1837
01:13:47,990 --> 01:13:45,920
stratification uh and uh and and strong

1838
01:13:51,189 --> 01:13:48,000

selection and this mechanism credit

1839
01:13:54,149 --> 01:13:52,870
did that answer i didn't see who the

1840
01:13:57,510 --> 01:13:54,159
question it was but did i answer the

1841
01:14:02,790 --> 01:13:59,270
we have another question that kind of

1842
01:14:04,149 --> 01:14:02,800
ties in with it um

1843
01:14:07,750 --> 01:14:04,159
uh i'm trying to figure out how to

1844
01:14:11,350 --> 01:14:09,430
you see i'm talking about

1845
01:14:15,270 --> 01:14:11,360
um ken

1846
01:14:17,030 --> 01:14:15,280
i'm not sure but is it the um

1847
01:14:18,630 --> 01:14:17,040
the cyanobig so i've had it well why

1848
01:14:20,550 --> 01:14:18,640
don't i'll do my question i think it

1849
01:14:22,709 --> 01:14:20,560
relates to one of these and that has to

1850
01:14:24,709 --> 01:14:22,719
do with um

1851
01:14:26,229 --> 01:14:24,719
you mentioned that the model suggests

1852
01:14:29,830 --> 01:14:26,239
that you have a very large pan genome

1853
01:14:32,310 --> 01:14:29,840
and proclar caucus um partly because

1854
01:14:35,590 --> 01:14:32,320
they have this you know lack of crispr

1855
01:14:38,470 --> 01:14:35,600
casts or other defense mechanisms do

1856
01:14:41,270 --> 01:14:38,480
people see the same sizes of pan genomes

1857
01:14:42,630 --> 01:14:41,280
in systems that do have more protection

1858
01:14:45,430 --> 01:14:42,640
how much is known about that i don't

1859
01:14:47,350 --> 01:14:45,440
know i'm just curious i i no i i don't

1860
01:14:50,630 --> 01:14:47,360
think they do this this this is a very

1861
01:14:53,350 --> 01:14:50,640
large pan genome by bacterial standards

1862
01:14:56,550 --> 01:14:53,360
and uh and and um

1863
01:14:58,550 --> 01:14:56,560

and we our model uh predicts that this

1864
01:15:02,070 --> 01:14:58,560
is really an outcome of the fact that

1865
01:15:03,910 --> 01:15:02,080
these these tiny bacteria just don't

1866
01:15:06,630 --> 01:15:03,920
have these of these defenses so they're

1867
01:15:18,950 --> 01:15:06,640
forced to make a living in this way

1868
01:15:24,950 --> 01:15:21,750
i guess i'll follow up with a question

1869
01:15:26,149 --> 01:15:24,960
can you use the model to guess which

1870
01:15:27,510 --> 01:15:26,159
genes would be

1871
01:15:30,550 --> 01:15:27,520
acquired

1872
01:15:33,189 --> 01:15:30,560
for a pathway so for photosynthesis we

1873
01:15:34,790 --> 01:15:33,199
typically see psba and pspd for

1874
01:15:37,990 --> 01:15:34,800
photosystem 2.

1875
01:15:40,229 --> 01:15:38,000
can you make predictions from that

1876
01:15:42,630 --> 01:15:40,239
so i think you i think you could

1877
01:15:44,709 --> 01:15:42,640
although we we haven't tried to do that

1878
01:15:47,110 --> 01:15:44,719
um i mean i think the most important

1879
01:15:49,270 --> 01:15:47,120
thing is whatever is the

1880
01:15:51,750 --> 01:15:49,280
the rule the rule is that whatever is

1881
01:15:53,990 --> 01:15:51,760
the the strongest form of selection that

1882
01:15:56,790 --> 01:15:54,000
is going on those will be the genes that

1883
01:15:58,470 --> 01:15:56,800
will be swapped back and forth between

1884
01:15:59,590 --> 01:15:58,480
between the phage i mean the bacterial

1885
01:16:00,550 --> 01:15:59,600
community

1886
01:16:04,830 --> 01:16:00,560
will

1887
01:16:08,630 --> 01:16:04,840
this avenue to

1888
01:16:09,830 --> 01:16:08,640

um you know to to survive and so if you

1889
01:16:12,310 --> 01:16:09,840
want to know what is the strongest

1890
01:16:14,229 --> 01:16:12,320
selection pressure that the host is is

1891
01:16:16,149 --> 01:16:14,239
is um undergoing

1892
01:16:17,910 --> 01:16:16,159
look at look at what genes are carried

1893
01:16:21,189 --> 01:16:17,920
by the phage and ask how did they get

1894
01:16:23,590 --> 01:16:21,199
there and that would be the reason

1895
01:16:25,669 --> 01:16:23,600
so so so um

1896
01:16:26,870 --> 01:16:25,679
so could you have predicted this

1897
01:16:28,310 --> 01:16:26,880
well

1898
01:16:29,830 --> 01:16:28,320
probably you could have done for the

1899
01:16:31,430 --> 01:16:29,840
reason that you that you said we unders

1900
01:16:33,990 --> 01:16:31,440
we understand a bit about photosystem

1901
01:16:35,110 --> 01:16:34,000
too so you you would have uh

1902
01:16:36,709 --> 01:16:35,120
you might have predicted that these

1903
01:16:37,990 --> 01:16:36,719
would be the genes because that's

1904
01:16:40,070 --> 01:16:38,000
obviously the form of selection that's

1905
01:16:41,910 --> 01:16:40,080
going to be important for this system

1906
01:16:43,750 --> 01:16:41,920
but uh but you could do the same thing

1907
01:16:46,870 --> 01:16:43,760
and we've suggested this to

1908
01:16:48,470 --> 01:16:46,880
for for other ecosystems uh one one

1909
01:16:50,470 --> 01:16:48,480
would actually be able to predict what

1910
01:16:51,830 --> 01:16:50,480
what what are certainly what what those

1911
01:16:55,270 --> 01:16:51,840
genes would be i think and certainly

1912
01:16:59,669 --> 01:16:57,510
yeah good job nigel i have a comment

1913
01:17:02,229 --> 01:16:59,679

this is penny um

1914
01:17:05,669 --> 01:17:02,239
it's not really a question it's just an

1915
01:17:08,310 --> 01:17:05,679
observation and that is that

1916
01:17:11,990 --> 01:17:08,320
the value of having this very large

1917
01:17:13,990 --> 01:17:12,000
reservoir of information available

1918
01:17:15,510 --> 01:17:14,000
for the photosynthesis process seems

1919
01:17:18,470 --> 01:17:15,520
analogous to me

1920
01:17:21,030 --> 01:17:18,480
uh in terms of the nitrogen fixation uh

1921
01:17:23,590 --> 01:17:21,040
process which is widely distributed

1922
01:17:25,830 --> 01:17:23,600
through so many many different organisms

1923
01:17:29,030 --> 01:17:25,840
and is so fundamental and there are many

1924
01:17:33,030 --> 01:17:29,040
variants of course as we now know of the

1925
01:17:34,390 --> 01:17:33,040
nitrogenous uh genes and so forth and it

1926
01:17:37,750 --> 01:17:34,400
seems like it's a

1927
01:17:40,390 --> 01:17:37,760
similar circumstance with um similar

1928
01:17:42,630 --> 01:17:40,400
import to the biosphere as a whole

1929
01:17:45,430 --> 01:17:42,640
yes i agree i suppose from the planetary

1930
01:17:46,310 --> 01:17:45,440
perspective this is this is this is a

1931
01:17:47,750 --> 01:17:46,320
weight

1932
01:17:51,030 --> 01:17:47,760
from the planet's point of view this is

1933
01:17:53,430 --> 01:17:51,040
a way to ensure uh the the the

1934
01:17:55,189 --> 01:17:53,440
the greatest amount of uh dissipation

1935
01:17:59,590 --> 01:17:55,199
free energy gradients

1936
01:18:02,310 --> 01:17:59,600
um so yeah so so yes so so life life is

1937
01:18:04,550 --> 01:18:02,320
is doing that for the planet as it were

1938
01:18:06,550 --> 01:18:04,560

and so that seems actually sort of gaian

1939
01:18:08,470 --> 01:18:06,560
to me

1940
01:18:10,790 --> 01:18:08,480
well i don't know whether it's gaian or

1941
01:18:13,430 --> 01:18:10,800
not but but but it's but it certainly is

1942
01:18:15,590 --> 01:18:13,440
a dynamical mechanism that um

1943
01:18:20,149 --> 01:18:15,600
that that is implied by the existence of

1944
01:18:22,310 --> 01:18:20,159
horizontal gene transfer um and the the

1945
01:18:24,229 --> 01:18:22,320
the way that that

1946
01:18:27,750 --> 01:18:24,239
involves the population dynamics along

1947
01:18:29,030 --> 01:18:27,760
with the metabolism of the organisms

1948
01:18:31,750 --> 01:18:29,040
it doesn't it doesn't necessarily mean

1949
01:18:34,149 --> 01:18:31,760
that the system can't go unstable

1950
01:18:35,270 --> 01:18:34,159
but but it certainly is a it certainly

1951
01:18:36,149 --> 01:18:35,280
is a way

1952
01:18:37,110 --> 01:18:36,159
for

1953
01:18:37,830 --> 01:18:37,120
um

1954
01:18:39,750 --> 01:18:37,840
for

1955
01:18:41,510 --> 01:18:39,760
well i i think you know i think one of

1956
01:18:42,470 --> 01:18:41,520
the burning questions is you know is

1957
01:18:44,310 --> 01:18:42,480
life

1958
01:18:46,709 --> 01:18:44,320
an inevitable consequence of the laws of

1959
01:18:48,390 --> 01:18:46,719
physics and you know is life something

1960
01:18:51,830 --> 01:18:48,400
that is genetic in a planetary

1961
01:18:53,189 --> 01:18:51,840
environment and this the sort of

1962
01:18:55,350 --> 01:18:53,199
argument i was presenting at the

1963
01:18:57,830 --> 01:18:55,360

beginning

1964
01:18:59,910 --> 01:18:57,840
which many years ago i i think i heard

1965
01:19:02,630 --> 01:18:59,920
first from everett shock

1966
01:19:04,870 --> 01:19:02,640
that sort of argument you know regards

1967
01:19:08,229 --> 01:19:04,880
life as a planetary process and and as

1968
01:19:10,790 --> 01:19:08,239
something that is a is a way for mata to

1969
01:19:14,070 --> 01:19:10,800
to organize as an information bearing

1970
01:19:16,229 --> 01:19:14,080
system in order to uh um not in a

1971
01:19:20,070 --> 01:19:16,239
technological way but to have the effect

1972
01:19:22,310 --> 01:19:20,080
of uh of uh increasing the uh the rate

1973
01:19:23,910 --> 01:19:22,320
at which a planetary system approaches

1974
01:19:25,750 --> 01:19:23,920
equilibrium

1975
01:19:27,990 --> 01:19:25,760
yeah thank you

1976
01:19:30,550 --> 01:19:28,000
okay thanks thank you nigel i think we

1977
01:19:33,750 --> 01:19:30,560
need to move on to simone so let's thank

1978
01:19:36,709 --> 01:19:35,270
and see if we can fire seymour looks

1979
01:19:40,709 --> 01:19:36,719
like this is a host so we should be able

1980
01:19:40,719 --> 01:19:45,910
just need a screen sharing

1981
01:19:50,630 --> 01:19:48,830
is is my screen sharing turned off

1982
01:19:52,390 --> 01:19:50,640
um yours is

1983
01:19:54,229 --> 01:19:52,400
but the hosts need to re-enable

1984
01:19:56,470 --> 01:19:54,239
participant screen sharing because

1985
01:19:57,990 --> 01:19:56,480
before i can do anything i guess

1986
01:19:59,270 --> 01:19:58,000
yeah marco should hopefully be able to

1987
01:20:01,510 --> 01:19:59,280
take care of that i've made you a

1988
01:20:03,030 --> 01:20:01,520

co-host so you should be able to share

1989
01:20:05,430 --> 01:20:03,040
now simon

1990
01:20:08,310 --> 01:20:05,440
um it doesn't want me to i get the same

1991
01:20:32,070 --> 01:20:08,320
error message that i had host disabled

1992
01:20:32,080 --> 01:20:35,430
okay try now

1993
01:20:39,830 --> 01:20:36,709
yeah

1994
01:21:23,320 --> 01:20:42,790
okay i can bring up your slides

1995
01:21:23,330 --> 01:21:31,910
[Music]

1996
01:21:31,920 --> 01:21:58,550
so

1997
01:22:16,870 --> 01:22:08,830
okay

1998
01:22:18,550 --> 01:22:16,880
thank you very much

1999
01:22:21,430 --> 01:22:18,560
um

2000
01:22:24,070 --> 01:22:21,440
so uh it's actually a pretty good segue

2001
01:22:25,270 --> 01:22:24,080
from what um evelyn presented earlier

2002
01:22:27,270 --> 01:22:25,280
and then nigel and then actually were

2003
01:22:30,629 --> 01:22:27,280
represented yesterday

2004
01:22:32,229 --> 01:22:30,639
what i will talk today about is mostly

2005
01:22:34,629 --> 01:22:32,239
um

2006
01:22:36,390 --> 01:22:34,639
three main things the first two are

2007
01:22:38,070 --> 01:22:36,400
maybe more technical

2008
01:22:39,590 --> 01:22:38,080
and they are mostly you know why do we

2009
01:22:41,830 --> 01:22:39,600
even care about viruses in the

2010
01:22:46,629 --> 01:22:41,840
environment and how do we

2011
01:22:48,870 --> 01:22:46,639
try to study or even find them so that's

2012
01:22:51,030 --> 01:22:48,880
mostly sections for people who might be

2013
01:22:54,149 --> 01:22:51,040

interested into getting into the virus

2014
01:22:56,709 --> 01:22:54,159
world and especially looking to

2015
01:22:58,950 --> 01:22:56,719
do virus analysis from metagenomics you

2016
01:23:01,270 --> 01:22:58,960
will have a lot of resources in there at

2017
01:23:03,990 --> 01:23:01,280
least some pointers i should say

2018
01:23:07,910 --> 01:23:04,000
to what you can and can't do and then

2019
01:23:09,110 --> 01:23:07,920
the third part is about the latest very

2020
01:23:13,669 --> 01:23:09,120
very

2021
01:23:16,070 --> 01:23:13,679
postdoc and it's about some resource

2022
01:23:18,070 --> 01:23:16,080
dynamics which will also tie into what

2023
01:23:19,990 --> 01:23:18,080
nigel just presented and what josh

2024
01:23:21,830 --> 01:23:20,000
presented yesterday and rachel so that's

2025
01:23:23,510 --> 01:23:21,840
you know resistance dynamics and

2026
01:23:24,950 --> 01:23:23,520
everything that's pretty cool i i'm

2027
01:23:27,430 --> 01:23:24,960
pleasantly surprised how well this all

2028
01:23:28,950 --> 01:23:27,440
tied together

2029
01:23:30,950 --> 01:23:28,960
um

2030
01:23:33,350 --> 01:23:30,960
i don't think i need to delve too much

2031
01:23:35,189 --> 01:23:33,360
on this that's my typical intro slides

2032
01:23:37,510 --> 01:23:35,199
microbes are important microbes are

2033
01:23:39,590 --> 01:23:37,520
everywhere microbes are very

2034
01:23:41,910 --> 01:23:39,600
numerous and very abundant and microbes

2035
01:23:43,750 --> 01:23:41,920
are pretty much controlling

2036
01:23:46,470 --> 01:23:43,760
everything that happens on this planet

2037
01:23:48,310 --> 01:23:46,480
at scale almost

2038
01:23:50,070 --> 01:23:48,320

if microbes are so important of course

2039
01:23:52,229 --> 01:23:50,080
viruses of microbes are

2040
01:23:54,550 --> 01:23:52,239
consequently very important

2041
01:23:56,709 --> 01:23:54,560
uh we tend to focus on these tiny bits

2042
01:23:58,390 --> 01:23:56,719
of viruses especially human viruses but

2043
01:24:00,629 --> 01:23:58,400
various of microbes this

2044
01:24:01,750 --> 01:24:00,639
gigantic world outside of this and and

2045
01:24:03,110 --> 01:24:01,760
completely

2046
01:24:04,390 --> 01:24:03,120
dwarfing them

2047
01:24:06,629 --> 01:24:04,400
um

2048
01:24:08,470 --> 01:24:06,639
evelyn already shot this same type of

2049
01:24:10,070 --> 01:24:08,480
picture if you want to have more of a

2050
01:24:10,950 --> 01:24:10,080
visual representation of this various

2051
01:24:12,790 --> 01:24:10,960
world

2052
01:24:14,950 --> 01:24:12,800
if you are not here for her tone this is

2053
01:24:17,430 --> 01:24:14,960
one drop of sea water

2054
01:24:19,910 --> 01:24:17,440
concentrated and then stained with cyber

2055
01:24:22,550 --> 01:24:19,920
green which will rebuild dna

2056
01:24:24,950 --> 01:24:22,560
big dots like this and i guess you can

2057
01:24:25,830 --> 01:24:24,960
see my mouth right

2058
01:24:27,830 --> 01:24:25,840
yep

2059
01:24:29,750 --> 01:24:27,840
perfect so these big dots here are

2060
01:24:31,669 --> 01:24:29,760
microbial cells all the small dots

2061
01:24:33,430 --> 01:24:31,679
behind are viruses

2062
01:24:35,590 --> 01:24:33,440
and so you can easily just guess from

2063
01:24:37,830 --> 01:24:35,600

here how many different viruses there is

2064
01:24:39,510 --> 01:24:37,840
in just one drop of sea water if you're

2065
01:24:41,430 --> 01:24:39,520
a number person that's the kind of

2066
01:24:43,189 --> 01:24:41,440
numbers we are

2067
01:24:45,189 --> 01:24:43,199
it's not very well constrained which is

2068
01:24:47,350 --> 01:24:45,199
kind of a common theme across all of our

2069
01:24:49,750 --> 01:24:47,360
psychology none of our numbers is

2070
01:24:51,590 --> 01:24:49,760
extremely well constrained but it gives

2071
01:24:53,270 --> 01:24:51,600
you another magnitude right

2072
01:24:55,270 --> 01:24:53,280
so that's the number of virus like

2073
01:24:57,750 --> 01:24:55,280
particles a number of

2074
01:24:59,669 --> 01:24:57,760
things which look like a virus and uh

2075
01:25:01,270 --> 01:24:59,679
most likely a virus

2076
01:25:02,950 --> 01:25:01,280
that we can observe in one milliliter of

2077
01:25:03,910 --> 01:25:02,960
sea water in one milliliter of fresh

2078
01:25:06,310 --> 01:25:03,920
water

2079
01:25:09,430 --> 01:25:06,320
or in one gram of soil you know we are

2080
01:25:11,030 --> 01:25:09,440
in the tens of millions if not billions

2081
01:25:12,870 --> 01:25:11,040
so that's kind of the scale at which we

2082
01:25:13,990 --> 01:25:12,880
are we are looking at all the kind of

2083
01:25:15,590 --> 01:25:14,000
scale we are looking at we are looking

2084
01:25:18,070 --> 01:25:15,600
at when we are talking about this virus

2085
01:25:21,669 --> 01:25:19,830
now um

2086
01:25:22,550 --> 01:25:21,679
what do viruses do if they are so you

2087
01:25:23,990 --> 01:25:22,560
know they are very abundant and

2088
01:25:25,510 --> 01:25:24,000

everything that's cool but what do they

2089
01:25:27,270 --> 01:25:25,520
actually do there

2090
01:25:29,189 --> 01:25:27,280
fortunately for me most of this has

2091
01:25:31,990 --> 01:25:29,199
already been covered just in the last

2092
01:25:33,110 --> 01:25:32,000
talk so that's great but of course

2093
01:25:34,709 --> 01:25:33,120
microbes will influence microbial

2094
01:25:36,790 --> 01:25:34,719
community structure because micro

2095
01:25:38,550 --> 01:25:36,800
viruses kill microbes and so micro

2096
01:25:39,510 --> 01:25:38,560
community will be influenced by this

2097
01:25:41,830 --> 01:25:39,520
that's

2098
01:25:43,510 --> 01:25:41,840
another um schematic of the same log

2099
01:25:45,110 --> 01:25:43,520
double there are dynamics that was just

2100
01:25:46,229 --> 01:25:45,120
described so you start from a diverse

2101
01:25:48,550 --> 01:25:46,239
community

2102
01:25:51,350 --> 01:25:48,560
one species proliferates and then the

2103
01:25:54,310 --> 01:25:51,360
phage specifically infecting this

2104
01:25:55,510 --> 01:25:54,320
microbe or bacteria in this case this

2105
01:25:56,950 --> 01:25:55,520
green one

2106
01:25:58,550 --> 01:25:56,960
will kill most of these cells and

2107
01:26:00,950 --> 01:25:58,560
eventually this will

2108
01:26:02,950 --> 01:26:00,960
restore or return community to some kind

2109
01:26:05,189 --> 01:26:02,960
of equilibrium and at least a

2110
01:26:06,870 --> 01:26:05,199
higher diversity that's

2111
01:26:09,189 --> 01:26:06,880
an octave ultra dynamics which apply to

2112
01:26:11,510 --> 01:26:09,199
phage and microbes is usually deemed

2113
01:26:13,030 --> 01:26:11,520

keeps the winner

2114
01:26:14,390 --> 01:26:13,040
viruses also shuffled genes from one

2115
01:26:16,310 --> 01:26:14,400
cell to the next i actually don't have

2116
01:26:18,390 --> 01:26:16,320
to do any introduction to this

2117
01:26:20,229 --> 01:26:18,400
because nigel just covered it perfectly

2118
01:26:22,310 --> 01:26:20,239
but just remind yeah just remember that

2119
01:26:24,070 --> 01:26:22,320
viruses do shuttle genes from one

2120
01:26:26,229 --> 01:26:24,080
microbes index and that influence

2121
01:26:27,590 --> 01:26:26,239
microbial genome evolution over

2122
01:26:29,830 --> 01:26:27,600
large um

2123
01:26:31,110 --> 01:26:29,840
time frames

2124
01:26:33,350 --> 01:26:31,120
and then

2125
01:26:35,430 --> 01:26:33,360
when viruses enter cells they also

2126
01:26:37,110 --> 01:26:35,440
typically do more than just replicate

2127
01:26:37,990 --> 01:26:37,120
they will take over the cell and they

2128
01:26:39,990 --> 01:26:38,000
can

2129
01:26:42,629 --> 01:26:40,000
tweak the metabolism of these cells to

2130
01:26:44,149 --> 01:26:42,639
their own benefits and when many viruses

2131
01:26:46,550 --> 01:26:44,159
infect many microbial cells that they

2132
01:26:48,870 --> 01:26:46,560
have that can have

2133
01:26:51,110 --> 01:26:48,880
very large scale impacts

2134
01:26:52,870 --> 01:26:51,120
so here the illustration is is from one

2135
01:26:54,470 --> 01:26:52,880
of maya's paper on the cyanobacteria

2136
01:26:56,790 --> 01:26:54,480
cyanophage story so again nigel

2137
01:26:58,149 --> 01:26:56,800
presented this model just before

2138
01:27:00,709 --> 01:26:58,159

but uh

2139
01:27:01,750 --> 01:27:00,719
just to recap it the cyanobacteria is

2140
01:27:05,110 --> 01:27:01,760
this green

2141
01:27:07,430 --> 01:27:05,120
cell sine of ages is orange virus

2142
01:27:08,950 --> 01:27:07,440
infecting it injecting the dna and in

2143
01:27:10,470 --> 01:27:08,960
the genome of the cyanophage there are a

2144
01:27:12,310 --> 01:27:10,480
number of genes

2145
01:27:13,830 --> 01:27:12,320
involved in the

2146
01:27:15,510 --> 01:27:13,840
cellular metabolism that the phage

2147
01:27:17,830 --> 01:27:15,520
acquired from the host through

2148
01:27:19,270 --> 01:27:17,840
horizontal gene transfer antigen are

2149
01:27:21,110 --> 01:27:19,280
expressed and are

2150
01:27:22,870 --> 01:27:21,120
really used to redirect the metabolism

2151
01:27:24,550 --> 01:27:22,880
towards what the virus wants which is

2152
01:27:25,990 --> 01:27:24,560
more nucleotides to do more virus

2153
01:27:27,510 --> 01:27:26,000
progeny

2154
01:27:29,510 --> 01:27:27,520
and again it's more of an indirect

2155
01:27:31,430 --> 01:27:29,520
consequence but if a lot of viruses are

2156
01:27:34,470 --> 01:27:31,440
doing this at the same time on a lot of

2157
01:27:36,470 --> 01:27:34,480
microbes this will change the way these

2158
01:27:38,390 --> 01:27:36,480
microbes can process nutrients and

2159
01:27:41,270 --> 01:27:38,400
eventually just affect global

2160
01:27:42,870 --> 01:27:41,280
geochemical cycles around the world

2161
01:27:43,910 --> 01:27:42,880
so at this point we have two things we

2162
01:27:46,709 --> 01:27:43,920
have

2163
01:27:48,870 --> 01:27:46,719

here other viruses doing

2164
01:27:51,510 --> 01:27:48,880
or at least having a lot of

2165
01:27:53,590 --> 01:27:51,520
massive impact on microbial communities

2166
01:27:55,350 --> 01:27:53,600
so that's a good reason why we want to

2167
01:27:58,950 --> 01:27:55,360
know more about these viruses and these

2168
01:28:00,709 --> 01:27:58,960
various communities how do we do that

2169
01:28:02,629 --> 01:28:00,719
that's when it starts to become tricky

2170
01:28:04,550 --> 01:28:02,639
as gary mentioned earlier we don't have

2171
01:28:06,629 --> 01:28:04,560
a universal marker genes so if you are

2172
01:28:08,950 --> 01:28:06,639
familiar with 16s or 18s and the

2173
01:28:12,709 --> 01:28:08,960
amplicon worth and just revolutionized

2174
01:28:14,950 --> 01:28:12,719
microbial ecology in the 80s 90s 2000

2175
01:28:16,310 --> 01:28:14,960
we don't have this in viruses

2176
01:28:17,910 --> 01:28:16,320
where is the answer challenging to

2177
01:28:20,870 --> 01:28:17,920
cultivate i don't know if anyone here

2178
01:28:23,590 --> 01:28:20,880
has ever tried to cultivate new phages

2179
01:28:25,189 --> 01:28:23,600
or new viruses on a new host but it's

2180
01:28:27,270 --> 01:28:25,199
notoriously difficult and it's time

2181
01:28:28,629 --> 01:28:27,280
consuming and it's really hard just

2182
01:28:31,030 --> 01:28:28,639
plain hard

2183
01:28:33,590 --> 01:28:31,040
so what we need is is cultivation-free

2184
01:28:34,629 --> 01:28:33,600
approaches which require no prior

2185
01:28:37,430 --> 01:28:34,639
knowledge

2186
01:28:39,110 --> 01:28:37,440
on the virus we want to

2187
01:28:40,790 --> 01:28:39,120
on the very community we want to analyze

2188
01:28:42,390 --> 01:28:40,800

so basically we can't have the

2189
01:28:43,750 --> 01:28:42,400
primer-based thing we're going to be

2190
01:28:45,590 --> 01:28:43,760
cultivation-based approach we need

2191
01:28:47,910 --> 01:28:45,600
something we would just go in

2192
01:28:49,430 --> 01:28:47,920
and try to look at everything we can and

2193
01:28:50,470 --> 01:28:49,440
that's pretty much what metagenomics

2194
01:28:52,629 --> 01:28:50,480
gives you

2195
01:28:54,229 --> 01:28:52,639
so that's a very simplified view of the

2196
01:28:55,990 --> 01:28:54,239
pipeline that we

2197
01:28:58,790 --> 01:28:56,000
everyone applies for doing meta genomes

2198
01:28:59,910 --> 01:28:58,800
so meta genomes in a nutshell is taking

2199
01:29:02,070 --> 01:28:59,920
a sample

2200
01:29:03,669 --> 01:29:02,080
extracting dna or rna depending on your

2201
01:29:06,149 --> 01:29:03,679
favorite bugs

2202
01:29:09,350 --> 01:29:06,159
doing you know shotgun sequencing which

2203
01:29:11,189 --> 01:29:09,360
just is sequencing everything randomly

2204
01:29:11,990 --> 01:29:11,199
usually in very short bits at least for

2205
01:29:14,709 --> 01:29:12,000
now

2206
01:29:17,590 --> 01:29:14,719
and then we piece this short sequence

2207
01:29:19,669 --> 01:29:17,600
together and we reconstruct genomes from

2208
01:29:21,750 --> 01:29:19,679
this shotgun sequence thing

2209
01:29:23,990 --> 01:29:21,760
and the major difference between these

2210
01:29:25,350 --> 01:29:24,000
two lines here on top is you know what i

2211
01:29:26,790 --> 01:29:25,360
call microbial metagenome which is a

2212
01:29:28,629 --> 01:29:26,800
typical bulk method genomes where we

2213
01:29:30,709 --> 01:29:28,639

don't select for anything

2214
01:29:33,350 --> 01:29:30,719
the bottom one is viral metagenomes or

2215
01:29:35,030 --> 01:29:33,360
virums in this case there is a first

2216
01:29:37,030 --> 01:29:35,040
step where we try to remove all the

2217
01:29:38,550 --> 01:29:37,040
cells and only keep virus capsids and

2218
01:29:40,629 --> 01:29:38,560
then we follow the same dna extraction

2219
01:29:42,709 --> 01:29:40,639
etcetera etcetera etcetera the idea here

2220
01:29:44,950 --> 01:29:42,719
being if i can remove the cells and keep

2221
01:29:47,270 --> 01:29:44,960
some virus capsids most of my sequencing

2222
01:29:49,430 --> 01:29:47,280
will actually be on viruses and most of

2223
01:29:52,870 --> 01:29:49,440
the genomes i will reconstruct will be

2224
01:29:54,629 --> 01:29:52,880
virus genomes that's at least a theory

2225
01:29:55,750 --> 01:29:54,639
all of this end up on the same thing

2226
01:29:57,110 --> 01:29:55,760
which is

2227
01:29:59,590 --> 01:29:57,120
once you have all your reconstructed

2228
01:30:01,750 --> 01:29:59,600
genomes the game becomes

2229
01:30:04,709 --> 01:30:01,760
which of these genomes are virus

2230
01:30:09,430 --> 01:30:06,629
so just to give you a again a quick

2231
01:30:10,950 --> 01:30:09,440
sense of scale of why um

2232
01:30:12,709 --> 01:30:10,960
i'm excited about metagenomics and i've

2233
01:30:14,310 --> 01:30:12,719
been for a long time and also why this

2234
01:30:15,590 --> 01:30:14,320
has become such a big feature in the

2235
01:30:17,910 --> 01:30:15,600
field

2236
01:30:20,629 --> 01:30:17,920
uh this is just a graph of the number of

2237
01:30:23,910 --> 01:30:20,639
genomes we have in ncbi in refseq for

2238
01:30:25,830 --> 01:30:23,920

all various species from 201418 and

2239
01:30:28,470 --> 01:30:25,840
that's a log scale so you know we went

2240
01:30:30,870 --> 01:30:28,480
and that's mostly from isolation so by

2241
01:30:33,510 --> 01:30:30,880
doing this painstakingly hard process of

2242
01:30:35,910 --> 01:30:33,520
isolating new viruses we went from a

2243
01:30:37,189 --> 01:30:35,920
little more than 1 000 genomes to almost

2244
01:30:38,550 --> 01:30:37,199
10 thousand

2245
01:30:40,390 --> 01:30:38,560
in you know

2246
01:30:41,669 --> 01:30:40,400
almost 15 years

2247
01:30:43,510 --> 01:30:41,679
and now you can compare this to the

2248
01:30:46,709 --> 01:30:43,520
number of various genomes we have been

2249
01:30:48,229 --> 01:30:46,719
able to assemble from meta genomes

2250
01:30:49,990 --> 01:30:48,239
and it started very late like you can

2251
01:30:52,390 --> 01:30:50,000
see this yellow curve

2252
01:30:55,270 --> 01:30:52,400
really in the year 2004 all the way to

2253
01:30:56,950 --> 01:30:55,280
2016 we were started to understand that

2254
01:30:58,390 --> 01:30:56,960
we could do this but we could not do it

2255
01:31:00,550 --> 01:30:58,400
at scale and that's really the last two

2256
01:31:01,510 --> 01:31:00,560
to three years that we went from having

2257
01:31:03,110 --> 01:31:01,520
a few

2258
01:31:05,270 --> 01:31:03,120
hundreds of these

2259
01:31:07,350 --> 01:31:05,280
to nearly a million right now and and

2260
01:31:08,790 --> 01:31:07,360
again it's a log scale so this linear

2261
01:31:10,709 --> 01:31:08,800
growth here is actually an exponential

2262
01:31:13,350 --> 01:31:10,719
growth we are we are just assembling

2263
01:31:15,350 --> 01:31:13,360

like crazy all types of our genomes from

2264
01:31:17,030 --> 01:31:15,360
every type of sample it's just becoming

2265
01:31:19,590 --> 01:31:17,040
almost routine at this point to find new

2266
01:31:21,110 --> 01:31:19,600
viruses in every medium we sequence

2267
01:31:22,629 --> 01:31:21,120
so that's that's kind of the power of

2268
01:31:24,149 --> 01:31:22,639
metagenomics here

2269
01:31:26,629 --> 01:31:24,159
it's just this almost boundless

2270
01:31:30,229 --> 01:31:26,639
exploration of your virus genome

2271
01:31:33,830 --> 01:31:32,310
just so you know um if you want to take

2272
01:31:36,709 --> 01:31:33,840
a look at these genomes that have been

2273
01:31:38,470 --> 01:31:36,719
assembled from metagenomes there is a

2274
01:31:41,350 --> 01:31:38,480
database that we have put together at

2275
01:31:45,110 --> 01:31:41,360
jgi which is called imgvr you have the

2276
01:31:46,709 --> 01:31:45,120
address here um you would have a bunch

2277
01:31:48,709 --> 01:31:46,719
of information on you know which genomes

2278
01:31:50,629 --> 01:31:48,719
where does it come from which samples

2279
01:31:52,709 --> 01:31:50,639
was it a symbol from what we thinking

2280
01:31:55,110 --> 01:31:52,719
etc etc so that's kind of the collection

2281
01:31:58,390 --> 01:31:55,120
of these genomes and whatever we know we

2282
01:31:59,270 --> 01:31:58,400
think we know about this virus

2283
01:32:01,990 --> 01:31:59,280
okay

2284
01:32:03,830 --> 01:32:02,000
so at this point i have told you

2285
01:32:04,950 --> 01:32:03,840
why we are interested by viruses of

2286
01:32:07,270 --> 01:32:04,960
course

2287
01:32:09,189 --> 01:32:07,280
and how we do this with versus microbes

2288
01:32:11,510 --> 01:32:09,199

in the environment which is primarily

2289
01:32:14,390 --> 01:32:11,520
through metagenomics and i told you yeah

2290
01:32:15,270 --> 01:32:14,400
we can have tons of genomes that's great

2291
01:32:17,350 --> 01:32:15,280
um

2292
01:32:19,830 --> 01:32:17,360
what's next at this point again we do

2293
01:32:21,750 --> 01:32:19,840
that since 2016 so it's almost all news

2294
01:32:23,110 --> 01:32:21,760
at this point so you want to do you want

2295
01:32:24,470 --> 01:32:23,120
to have more than just a genome and

2296
01:32:26,709 --> 01:32:24,480
that's

2297
01:32:29,189 --> 01:32:26,719
where um we decided to you know sit

2298
01:32:30,870 --> 01:32:29,199
together as a community with all these

2299
01:32:33,270 --> 01:32:30,880
lovely folks here

2300
01:32:34,629 --> 01:32:33,280
and we just published uh six months ago

2301
01:32:36,950 --> 01:32:34,639
ish

2302
01:32:38,870 --> 01:32:36,960
this paper in in nature biotech in this

2303
01:32:40,950 --> 01:32:38,880
perspective which is

2304
01:32:43,189 --> 01:32:40,960
a big descriptions of

2305
01:32:45,750 --> 01:32:43,199
what we think or currently are the

2306
01:32:47,750 --> 01:32:45,760
standards way of analyzing describing

2307
01:32:49,590 --> 01:32:47,760
and reporting these ubiquits and new

2308
01:32:51,669 --> 01:32:49,600
bigs here stands for uncultivated

2309
01:32:54,629 --> 01:32:51,679
various genomes which is another way of

2310
01:32:56,310 --> 01:32:54,639
saying a virus genome assembled

2311
01:32:57,910 --> 01:32:56,320
for viruses that we don't have in

2312
01:32:59,270 --> 01:32:57,920
culture

2313
01:33:01,669 --> 01:32:59,280

so

2314
01:33:03,590 --> 01:33:01,679
primarily metagenome assembled various

2315
01:33:05,270 --> 01:33:03,600
genomes so there are a few other

2316
01:33:07,030 --> 01:33:05,280
techniques that is why we have this more

2317
01:33:09,350 --> 01:33:07,040
generic mgovics

2318
01:33:11,430 --> 01:33:09,360
if anyone is interested into going into

2319
01:33:13,590 --> 01:33:11,440
this area or have metagenomes they want

2320
01:33:15,510 --> 01:33:13,600
to look into viruses that is a like this

2321
01:33:18,229 --> 01:33:15,520
paper specifically is a great starting

2322
01:33:19,669 --> 01:33:18,239
point because it covers pretty much all

2323
01:33:21,910 --> 01:33:19,679
the range of what we can do with this

2324
01:33:23,590 --> 01:33:21,920
data and it tells you okay that's what

2325
01:33:25,189 --> 01:33:23,600
you can do that's a standard stool and

2326
01:33:26,709 --> 01:33:25,199
that's what you should be careful with

2327
01:33:28,229 --> 01:33:26,719
when you interpret this results that's

2328
01:33:30,229 --> 01:33:28,239
really an amazing

2329
01:33:33,270 --> 01:33:30,239
starting point that all of this

2330
01:33:35,590 --> 01:33:33,280
community put together to kind of help

2331
01:33:37,830 --> 01:33:35,600
anyone new to the field and even the not

2332
01:33:39,270 --> 01:33:37,840
so new you know it's always good for

2333
01:33:41,270 --> 01:33:39,280
memory refresh to just have this

2334
01:33:42,149 --> 01:33:41,280
reference

2335
01:33:43,910 --> 01:33:42,159
okay

2336
01:33:46,790 --> 01:33:43,920
what is in this

2337
01:33:51,510 --> 01:33:48,629
first we have a list of the tools which

2338
01:33:53,030 --> 01:33:51,520

are right now available and broadly used

2339
01:33:56,790 --> 01:33:53,040
pretty much that's a list of what i

2340
01:33:59,189 --> 01:33:56,800
would argue we know how to do so

2341
01:34:00,870 --> 01:33:59,199
identifying various sequences in genome

2342
01:34:02,709 --> 01:34:00,880
or meta genomes

2343
01:34:04,790 --> 01:34:02,719
we are not perfect at it but we kind of

2344
01:34:06,229 --> 01:34:04,800
know how to do this

2345
01:34:08,790 --> 01:34:06,239
having a sense of the distribution and

2346
01:34:10,149 --> 01:34:08,800
abundance of this genome so basically

2347
01:34:11,750 --> 01:34:10,159
looking at different beta genomes and

2348
01:34:14,790 --> 01:34:11,760
telling you if the virus is there or not

2349
01:34:16,550 --> 01:34:14,800
there we kind of know to do this as well

2350
01:34:18,550 --> 01:34:16,560
functional annotation we also know how

2351
01:34:20,070 --> 01:34:18,560
to do it we don't do it well but it's

2352
01:34:22,070 --> 01:34:20,080
more a question of having the reference

2353
01:34:23,750 --> 01:34:22,080
then actually the methods for how to do

2354
01:34:25,590 --> 01:34:23,760
it so again these are tools that are

2355
01:34:26,950 --> 01:34:25,600
available broadly used and we'll give

2356
01:34:29,189 --> 01:34:26,960
you the standards tool in the center's

2357
01:34:30,790 --> 01:34:29,199
way and pipeline to do this

2358
01:34:32,390 --> 01:34:30,800
if you don't have the resource to do it

2359
01:34:34,550 --> 01:34:32,400
on your own

2360
01:34:36,310 --> 01:34:34,560
you have actually free online platforms

2361
01:34:39,510 --> 01:34:36,320
with these virus dedicated tools that

2362
01:34:41,830 --> 01:34:39,520
exist i've listed them here the main one

2363
01:34:45,510 --> 01:34:41,840

right now is called ivirus you have the

2364
01:34:47,510 --> 01:34:45,520
docs here it's mostly

2365
01:34:49,350 --> 01:34:47,520
leveraging what's called cybers which is

2366
01:34:51,830 --> 01:34:49,360
a cyber infrastructure which is a fancy

2367
01:34:53,350 --> 01:34:51,840
name to say is a place where you can go

2368
01:34:54,950 --> 01:34:53,360
upload your sequence and analyze your

2369
01:34:56,310 --> 01:34:54,960
data without worrying about like where

2370
01:34:57,189 --> 01:34:56,320
the competition is going or anything

2371
01:34:59,109 --> 01:34:57,199
it's just

2372
01:35:01,830 --> 01:34:59,119
fun now it's free you just go put your

2373
01:35:04,870 --> 01:35:01,840
data analyze it and get your results

2374
01:35:06,950 --> 01:35:04,880
another flavor of cybers or this another

2375
01:35:07,910 --> 01:35:06,960
a comparable tool to cybersys called k

2376
01:35:10,390 --> 01:35:07,920
base

2377
01:35:12,070 --> 01:35:10,400
um same thing the i virus tools are

2378
01:35:16,149 --> 01:35:12,080
being

2379
01:35:18,070 --> 01:35:16,159
would have i virus in both cases uh

2380
01:35:19,270 --> 01:35:18,080
bottom line here is that if you want to

2381
01:35:20,950 --> 01:35:19,280
do this analysis and you don't really

2382
01:35:22,470 --> 01:35:20,960
know how to start or where to start this

2383
01:35:25,990 --> 01:35:22,480
is a great starting point in terms of

2384
01:35:30,470 --> 01:35:27,990
now i told you about what we could do or

2385
01:35:32,229 --> 01:35:30,480
at least what we think we know how to do

2386
01:35:34,070 --> 01:35:32,239
uh what we don't really know how to do

2387
01:35:36,390 --> 01:35:34,080
or what we are currently trying to

2388
01:35:38,629 --> 01:35:36,400

understand how to do are these three

2389
01:35:41,109 --> 01:35:38,639
critical pieces um taxonomy

2390
01:35:42,550 --> 01:35:41,119
classification having already covered it

2391
01:35:44,070 --> 01:35:42,560
and she already presented how

2392
01:35:46,070 --> 01:35:44,080
complicated this was so i don't need to

2393
01:35:47,350 --> 01:35:46,080
convince you that taxonomically

2394
01:35:50,070 --> 01:35:47,360
classifying

2395
01:35:53,669 --> 01:35:50,080
partial genomes from metagenomes

2396
01:35:55,430 --> 01:35:53,679
is not a standardized and rooting thing

2397
01:35:57,189 --> 01:35:55,440
right now

2398
01:35:59,109 --> 01:35:57,199
the second one what we call quality

2399
01:36:01,189 --> 01:35:59,119
estimation is

2400
01:36:03,669 --> 01:36:01,199
this idea of when you get a piece of

2401
01:36:04,790 --> 01:36:03,679
genome from a metal genome you know it's

2402
01:36:06,790 --> 01:36:04,800
a virus

2403
01:36:09,750 --> 01:36:06,800
how do you determine if this is

2404
01:36:11,189 --> 01:36:09,760
10 of the four genome 50 of the full

2405
01:36:12,470 --> 01:36:11,199
genome maybe it's actually a complete

2406
01:36:14,310 --> 01:36:12,480
genome

2407
01:36:15,430 --> 01:36:14,320
we wish there was an automatic tool to

2408
01:36:17,350 --> 01:36:15,440
tell you

2409
01:36:19,030 --> 01:36:17,360
uh but right now it's still in

2410
01:36:21,030 --> 01:36:19,040
development and and in development at

2411
01:36:23,669 --> 01:36:21,040
the stage of like we are looking and

2412
01:36:25,830 --> 01:36:23,679
exploring some ideas that we think could

2413
01:36:28,149 --> 01:36:25,840

work on how to do it so that's kind of

2414
01:36:29,669 --> 01:36:28,159
early stage development

2415
01:36:31,590 --> 01:36:29,679
and the last one i will develop a little

2416
01:36:34,629 --> 01:36:31,600
more in one or two slides

2417
01:36:35,430 --> 01:36:34,639
but just to give some framework already

2418
01:36:37,109 --> 01:36:35,440
um

2419
01:36:38,629 --> 01:36:37,119
of course as soon as you find a new

2420
01:36:41,270 --> 01:36:38,639
virus and a new various genome one of

2421
01:36:42,870 --> 01:36:41,280
the first questions that pops to mind is

2422
01:36:45,109 --> 01:36:42,880
which host does it infect

2423
01:36:47,910 --> 01:36:45,119
and so this uh there is a bunch of

2424
01:36:50,149 --> 01:36:47,920
methods that we kind of gather under

2425
01:36:52,149 --> 01:36:50,159
this term of inside co-host prediction

2426
01:36:53,830 --> 01:36:52,159
which just means

2427
01:36:56,149 --> 01:36:53,840
host prediction based on sequence

2428
01:36:57,910 --> 01:36:56,159
analysis so not going back to the bench

2429
01:36:59,510 --> 01:36:57,920
but likely from the sequence of the

2430
01:37:01,990 --> 01:36:59,520
various genome can we

2431
01:37:04,310 --> 01:37:02,000
guess or take a best guess at what host

2432
01:37:07,030 --> 01:37:04,320
this virus infects and we sort of can

2433
01:37:10,149 --> 01:37:07,040
but there is still a lot of work in this

2434
01:37:13,510 --> 01:37:10,159
area done and and to be done until we

2435
01:37:15,189 --> 01:37:13,520
get a nice and robust simple tool one

2436
01:37:15,990 --> 01:37:15,199
you know one-size-fits-all that will

2437
01:37:17,189 --> 01:37:16,000
just

2438
01:37:19,669 --> 01:37:17,199

you can apply to your sequence and you

2439
01:37:21,270 --> 01:37:19,679
will get an answer

2440
01:37:23,350 --> 01:37:21,280
okay so

2441
01:37:25,109 --> 01:37:23,360
not wanting to discourage anyone you can

2442
01:37:26,870 --> 01:37:25,119
definitely do this just know that some

2443
01:37:29,910 --> 01:37:26,880
things are pretty much advanced some are

2444
01:37:34,629 --> 01:37:31,830
in a perfect world the final results of

2445
01:37:36,470 --> 01:37:34,639
all this is you will get various genome

2446
01:37:38,310 --> 01:37:36,480
one on several

2447
01:37:40,790 --> 01:37:38,320
and along with this you will get a full

2448
01:37:42,950 --> 01:37:40,800
ecological and evolutionary context

2449
01:37:45,430 --> 01:37:42,960
of course this doesn't work or at least

2450
01:37:46,870 --> 01:37:45,440
it can't be that simple

2451
01:37:48,470 --> 01:37:46,880
and and i just wanted to highlight a few

2452
01:37:50,550 --> 01:37:48,480
of the major challenges because if you

2453
01:37:52,550 --> 01:37:50,560
are looking for a problem uh here are

2454
01:37:53,669 --> 01:37:52,560
the main ones for us

2455
01:37:55,350 --> 01:37:53,679
um

2456
01:37:57,189 --> 01:37:55,360
the first one is i i showed this

2457
01:37:59,109 --> 01:37:57,199
schematic of like you take your sample

2458
01:38:00,550 --> 01:37:59,119
you just extract your dna blah blah and

2459
01:38:03,350 --> 01:38:00,560
you end up in your various sequence as

2460
01:38:05,430 --> 01:38:03,360
if this was a done deal actually this is

2461
01:38:07,189 --> 01:38:05,440
not that straightforward for every type

2462
01:38:08,310 --> 01:38:07,199
of sample and especially this first step

2463
01:38:09,510 --> 01:38:08,320

if you want to enrich from various

2464
01:38:11,030 --> 01:38:09,520
particles

2465
01:38:12,870 --> 01:38:11,040
we have some environments where we know

2466
01:38:14,470 --> 01:38:12,880
how to do this very well somewhere we

2467
01:38:17,350 --> 01:38:14,480
are still figuring things out just as an

2468
01:38:19,510 --> 01:38:17,360
example gary has a paper um

2469
01:38:21,189 --> 01:38:19,520
not much more than six months ago i

2470
01:38:23,669 --> 01:38:21,199
guess i don't remember the date but

2471
01:38:25,030 --> 01:38:23,679
that's 2019 paper about trying to do

2472
01:38:27,990 --> 01:38:25,040
this kind of enrichment in various

2473
01:38:29,510 --> 01:38:28,000
particles from soil and that's already

2474
01:38:31,510 --> 01:38:29,520
very tricky because there very stick

2475
01:38:33,430 --> 01:38:31,520
everywhere etcetera etcetera so just

2476
01:38:35,030 --> 01:38:33,440
know that this this kind of pipeline

2477
01:38:36,390 --> 01:38:35,040
they look good on paper they are not

2478
01:38:39,270 --> 01:38:36,400
necessarily very straightforward to

2479
01:38:41,510 --> 01:38:39,280
apply to any type of sample

2480
01:38:43,270 --> 01:38:41,520
into the same kind of line of thoughts

2481
01:38:45,109 --> 01:38:43,280
i have this box saying viral sequencing

2482
01:38:46,229 --> 01:38:45,119
notification i told you we kind of know

2483
01:38:47,750 --> 01:38:46,239
to do it

2484
01:38:50,149 --> 01:38:47,760
we do but it's

2485
01:38:51,590 --> 01:38:50,159
not perfect and there is i just wanted

2486
01:38:54,390 --> 01:38:51,600
to flag this out there is a very nice

2487
01:38:57,510 --> 01:38:54,400
paper um by the lab of bonnie hovitz in

2488
01:38:59,510 --> 01:38:57,520

frontiers where they showed with nice

2489
01:39:00,790 --> 01:38:59,520
examples and and kind of

2490
01:39:02,790 --> 01:39:00,800
benchmarks

2491
01:39:04,629 --> 01:39:02,800
what are the typical mistakes that can

2492
01:39:06,709 --> 01:39:04,639
be made especially when using machine

2493
01:39:07,990 --> 01:39:06,719
learning to do this various sequence

2494
01:39:10,070 --> 01:39:08,000
identification so if you're interested

2495
01:39:13,430 --> 01:39:10,080
by this aspect this is definitely a must

2496
01:39:18,950 --> 01:39:15,669
in silico's prediction i told you we are

2497
01:39:21,189 --> 01:39:18,960
not really there yet but one of the main

2498
01:39:23,270 --> 01:39:21,199
problem we have actually is that

2499
01:39:25,430 --> 01:39:23,280
it doesn't cover a lot of viruses so

2500
01:39:27,430 --> 01:39:25,440
again even if you try to use all the

2501
01:39:28,629 --> 01:39:27,440
tools that have been developed so far

2502
01:39:31,830 --> 01:39:28,639
and you really try everything and

2503
01:39:33,430 --> 01:39:31,840
understand that we have in our toolkit

2504
01:39:34,709 --> 01:39:33,440
this is kind of the results you get so

2505
01:39:37,030 --> 01:39:34,719
that's the same curve as i've shown

2506
01:39:38,950 --> 01:39:37,040
earlier with the number of uncultivated

2507
01:39:40,149 --> 01:39:38,960
genomes this time it's a actually linear

2508
01:39:41,910 --> 01:39:40,159
scale that's why it looks like any

2509
01:39:43,430 --> 01:39:41,920
difference but the important part is on

2510
01:39:47,030 --> 01:39:43,440
the right

2511
01:39:48,870 --> 01:39:47,040
in this um you know nearly 8000 genomes

2512
01:39:51,270 --> 01:39:48,880
at this point we have a host predicted

2513
01:39:53,109 --> 01:39:51,280

for five percent of them so we have 95

2514
01:39:55,189 --> 01:39:53,119
of them for which it's not even that we

2515
01:39:57,750 --> 01:39:55,199
don't know if this is a right or wrong

2516
01:39:58,870 --> 01:39:57,760
we actually don't have any prediction on

2517
01:40:00,950 --> 01:39:58,880
the host

2518
01:40:02,790 --> 01:40:00,960
so that's a huge area for improvements

2519
01:40:04,870 --> 01:40:02,800
and a huge um

2520
01:40:06,470 --> 01:40:04,880
like a critical key

2521
01:40:07,350 --> 01:40:06,480
for the field to move forward will be to

2522
01:40:08,470 --> 01:40:07,360
kind of

2523
01:40:10,390 --> 01:40:08,480
figure out

2524
01:40:14,229 --> 01:40:10,400
how to get the host prediction or host

2525
01:40:15,990 --> 01:40:14,239
information for this 95 percent

2526
01:40:17,669 --> 01:40:16,000
and then um yeah if you want to know

2527
01:40:19,109 --> 01:40:17,679
more about this again just a few

2528
01:40:20,950 --> 01:40:19,119
references i forgot

2529
01:40:23,669 --> 01:40:20,960
this we talked about this in the manual

2530
01:40:26,550 --> 01:40:23,679
paper uh rob edwards and bus dutil had a

2531
01:40:28,229 --> 01:40:26,560
very very nice review a few years ago um

2532
01:40:29,350 --> 01:40:28,239
about this aspect of insidicos

2533
01:40:33,109 --> 01:40:29,360
prediction

2534
01:40:34,310 --> 01:40:33,119
it's still definitely relevant today so

2535
01:40:36,390 --> 01:40:34,320
go and check this if you're interested

2536
01:40:39,830 --> 01:40:36,400
by this question about trying to predict

2537
01:40:41,830 --> 01:40:39,840
a host based on the paris genome

2538
01:40:43,590 --> 01:40:41,840

final state of challenge

2539
01:40:45,510 --> 01:40:43,600
seeing a virus in a meta genome is not

2540
01:40:47,830 --> 01:40:45,520
the same thing as

2541
01:40:50,870 --> 01:40:47,840
actually knowing that this virus is

2542
01:40:52,310 --> 01:40:50,880
active and and infecting exhaust in this

2543
01:40:53,910 --> 01:40:52,320
environment there's kind of this

2544
01:40:55,990 --> 01:40:53,920
discrepancy between you see the dna

2545
01:40:57,350 --> 01:40:56,000
versus you actually have an active virus

2546
01:41:00,310 --> 01:40:57,360
and the latter is really what you want

2547
01:41:03,109 --> 01:41:00,320
to get at for any ecological studies

2548
01:41:04,550 --> 01:41:03,119
so say otherwise if you see a virus

2549
01:41:06,470 --> 01:41:04,560
sequence in a metagenomic doesn't mean

2550
01:41:08,790 --> 01:41:06,480
that you have an active virus and we

2551
01:41:11,030 --> 01:41:08,800
need something to go at activity

2552
01:41:13,590 --> 01:41:11,040
um fortunately for me i don't need to do

2553
01:41:14,709 --> 01:41:13,600
anything because gary will present one

2554
01:41:16,310 --> 01:41:14,719
type of

2555
01:41:18,390 --> 01:41:16,320
experiment you can do

2556
01:41:19,990 --> 01:41:18,400
to try to get at this activity level

2557
01:41:22,310 --> 01:41:20,000
and so i don't need to talk more about

2558
01:41:24,310 --> 01:41:22,320
this uh just stay tuned and i guess it's

2559
01:41:26,629 --> 01:41:24,320
not the next talk anyone it's a one

2560
01:41:28,310 --> 01:41:26,639
after this but just stay tuned in gary's

2561
01:41:30,229 --> 01:41:28,320
talk and you will have a nice

2562
01:41:35,990 --> 01:41:30,239
example of what you can use to try to

2563
01:41:40,390 --> 01:41:37,669

so that's enough technical talk and

2564
01:41:43,430 --> 01:41:40,400
that's enough kind of generalities

2565
01:41:45,270 --> 01:41:43,440
no to some biology

2566
01:41:47,910 --> 01:41:45,280
and what i will present here

2567
01:41:49,910 --> 01:41:47,920
is a work from my positive marine berg

2568
01:41:51,270 --> 01:41:49,920
um green sulfur bacteria

2569
01:41:52,950 --> 01:41:51,280
and the virus is infecting ventricular

2570
01:41:54,550 --> 01:41:52,960
bacteria of course

2571
01:41:56,310 --> 01:41:54,560
so very quick background on green

2572
01:41:58,149 --> 01:41:56,320
surface bacteria which i will call gsb

2573
01:41:59,430 --> 01:41:58,159
through the whole thing

2574
01:42:01,910 --> 01:41:59,440
they bloom

2575
01:42:03,270 --> 01:42:01,920
every summer in in most rectified lake

2576
01:42:06,310 --> 01:42:03,280
and that's what you have here these are

2577
01:42:07,030 --> 01:42:06,320
um from a lake in in the alps and that's

2578
01:42:10,709 --> 01:42:07,040
uh

2579
01:42:12,310 --> 01:42:10,719
gsb is detected by pcr or qpcr

2580
01:42:13,990 --> 01:42:12,320
and you have the month of the year and

2581
01:42:15,990 --> 01:42:14,000
you can see in the summer you have this

2582
01:42:17,990 --> 01:42:16,000
nice bloom of um

2583
01:42:18,870 --> 01:42:18,000
gsb is right at a

2584
01:42:21,189 --> 01:42:18,880
big

2585
01:42:24,870 --> 01:42:21,199
discrete depth uh and this is another

2586
01:42:26,870 --> 01:42:24,880
strain of gsb also blooming here

2587
01:42:29,109 --> 01:42:26,880
and the question we have was which

2588
01:42:30,950 --> 01:42:29,119

viruses in fact gsb gsb are notoriously

2589
01:42:33,189 --> 01:42:30,960
difficult to cultivate and difficult to

2590
01:42:35,030 --> 01:42:33,199
maintain so right now we have absolutely

2591
01:42:37,189 --> 01:42:35,040
zero phage isolates for them so we have

2592
01:42:39,590 --> 01:42:37,199
no idea which viruses infect them or if

2593
01:42:40,390 --> 01:42:39,600
even one virus infects them

2594
01:42:42,709 --> 01:42:40,400
so

2595
01:42:44,709 --> 01:42:42,719
we wanted to know first are the viruses

2596
01:42:47,030 --> 01:42:44,719
which are these viruses and we were very

2597
01:42:49,510 --> 01:42:47,040
interested i was very interested into

2598
01:42:51,590 --> 01:42:49,520
which type of infections do we see if we

2599
01:42:52,950 --> 01:42:51,600
find viruses we really find these lytic

2600
01:42:54,709 --> 01:42:52,960
viruses

2601
01:42:57,189 --> 01:42:54,719
and which you know lead to this arms

2602
01:42:59,830 --> 01:42:57,199
race and and very active infection

2603
01:43:01,590 --> 01:42:59,840
you could say yes because gsb's are very

2604
01:43:03,590 --> 01:43:01,600
abundant when they bloom and so this is

2605
01:43:04,629 --> 01:43:03,600
a you know kill the winner

2606
01:43:07,109 --> 01:43:04,639
at least

2607
01:43:09,109 --> 01:43:07,119
on series this is a good ground for kill

2608
01:43:10,629 --> 01:43:09,119
the winner instead there will be plenty

2609
01:43:12,790 --> 01:43:10,639
of hosts and so you could have a lot of

2610
01:43:16,070 --> 01:43:12,800
lytic infections going on and a lot of

2611
01:43:17,430 --> 01:43:16,080
again arms race coevolution everything

2612
01:43:19,270 --> 01:43:17,440
alternatively

2613
01:43:21,750 --> 01:43:19,280

what you could have is

2614
01:43:23,830 --> 01:43:21,760
lysogenic infection and i think josh

2615
01:43:25,430 --> 01:43:23,840
really presented this already yesterday

2616
01:43:27,590 --> 01:43:25,440
just to remind everyone lysogenic

2617
01:43:29,750 --> 01:43:27,600
meaning the virus enters the host but

2618
01:43:32,390 --> 01:43:29,760
then stay put for a while

2619
01:43:34,229 --> 01:43:32,400
divide along with the cell until it's

2620
01:43:37,270 --> 01:43:34,239
reactivated and actually then go through

2621
01:43:39,109 --> 01:43:37,280
its host takeover and and urine progeny

2622
01:43:40,550 --> 01:43:39,119
generation and eventually hostile

2623
01:43:41,990 --> 01:43:40,560
killing

2624
01:43:43,750 --> 01:43:42,000
the reason for lysogenic that is the

2625
01:43:44,870 --> 01:43:43,760
argument or the rational for lysogeny in

2626
01:43:46,149 --> 01:43:44,880
this case

2627
01:43:49,669 --> 01:43:46,159
will come from the strong seasonal

2628
01:43:50,790 --> 01:43:49,679
variation you have gsb very abundant in

2629
01:43:52,629 --> 01:43:50,800
the summer

2630
01:43:54,470 --> 01:43:52,639
almost disappearing in the winter and

2631
01:43:57,109 --> 01:43:54,480
that has been shown as artists that has

2632
01:43:59,510 --> 01:43:57,119
been hypothesized as a good driver for

2633
01:44:01,590 --> 01:43:59,520
lysogeny and if you are interested by

2634
01:44:03,830 --> 01:44:01,600
this specific and and how life history

2635
01:44:06,070 --> 01:44:03,840
traits of hosts will correlate or not

2636
01:44:07,990 --> 01:44:06,080
with the type of infection you can see a

2637
01:44:09,910 --> 01:44:08,000
very nice paper by uh

2638
01:44:12,070 --> 01:44:09,920

from from pastor here

2639
01:44:12,870 --> 01:44:12,080
for anyone interested

2640
01:44:14,550 --> 01:44:12,880
um

2641
01:44:16,310 --> 01:44:14,560
how do we do

2642
01:44:17,830 --> 01:44:16,320
how do we find viruses for gsb if we can

2643
01:44:19,669 --> 01:44:17,840
cultivate them

2644
01:44:21,109 --> 01:44:19,679
and how do we actually understand and

2645
01:44:23,430 --> 01:44:21,119
get into this

2646
01:44:25,270 --> 01:44:23,440
well the thing we did is we partnered

2647
01:44:27,270 --> 01:44:25,280
with trina mcmahon in wisconsin we

2648
01:44:29,350 --> 01:44:27,280
sampled or they sampled

2649
01:44:30,390 --> 01:44:29,360
trina's lab sample this um throat bug

2650
01:44:32,149 --> 01:44:30,400
lake

2651
01:44:33,830 --> 01:44:32,159
this is the same kind of plot i just

2652
01:44:36,390 --> 01:44:33,840
showed this time it's a heat map not

2653
01:44:39,270 --> 01:44:36,400
based on 16s pcr but on flow cytometry

2654
01:44:40,709 --> 01:44:39,280
counts of gsbs so you can see as you go

2655
01:44:41,910 --> 01:44:40,719
from june to october you have a very

2656
01:44:44,629 --> 01:44:41,920
nice bloom

2657
01:44:46,390 --> 01:44:44,639
the white bar here is the oxygen barrier

2658
01:44:48,070 --> 01:44:46,400
so above there is oxygen below there is

2659
01:44:50,149 --> 01:44:48,080
no oxygen and we have this bloom just

2660
01:44:53,430 --> 01:44:50,159
right under this barrier which is

2661
01:44:55,590 --> 01:44:53,440
exactly what you would expect from gsb

2662
01:44:58,470 --> 01:44:55,600
and now what we decided to do is what we

2663
01:45:00,229 --> 01:44:58,480

call targeted metagenomics which is a

2664
01:45:01,910 --> 01:45:00,239
fancy way of saying

2665
01:45:03,669 --> 01:45:01,920
because gsbs are green you can use

2666
01:45:05,430 --> 01:45:03,679
procytometry to separate them from the

2667
01:45:07,030 --> 01:45:05,440
rest of the cells and then if you

2668
01:45:08,870 --> 01:45:07,040
separate enough of these green cells and

2669
01:45:09,830 --> 01:45:08,880
put them in one well and the meta genome

2670
01:45:14,790 --> 01:45:09,840
there

2671
01:45:17,350 --> 01:45:14,800
consistent basis across replicates

2672
01:45:18,229 --> 01:45:17,360
should be a phage infecting gsbs because

2673
01:45:20,070 --> 01:45:18,239
you have

2674
01:45:23,030 --> 01:45:20,080
nearly 100 percent of cells in there

2675
01:45:24,629 --> 01:45:23,040
being gsb's so that's what we what we

2676
01:45:26,950 --> 01:45:24,639
did and um

2677
01:45:28,870 --> 01:45:26,960
i would call the long story short

2678
01:45:30,470 --> 01:45:28,880
but basically the idea here is in silico

2679
01:45:32,070 --> 01:45:30,480
prediction will typically tell you who

2680
01:45:34,390 --> 01:45:32,080
could infect whom you know who have the

2681
01:45:35,430 --> 01:45:34,400
potential to infect womb or in case of

2682
01:45:38,390 --> 01:45:35,440
crispr

2683
01:45:40,070 --> 01:45:38,400
who used to infect whom in you know the

2684
01:45:43,030 --> 01:45:40,080
more or less recent past

2685
01:45:45,030 --> 01:45:43,040
this will tell you um which virus

2686
01:45:47,750 --> 01:45:45,040
were in these cells at the exact moment

2687
01:45:48,709 --> 01:45:47,760
you sampled that's a slightly different

2688
01:45:50,870 --> 01:45:48,719

view

2689
01:45:53,990 --> 01:45:50,880
on the host cell interaction that's

2690
01:45:55,590 --> 01:45:54,000
that's that are going on in your system

2691
01:45:56,550 --> 01:45:55,600
now what did we find

2692
01:45:58,229 --> 01:45:56,560
again

2693
01:46:00,470 --> 01:45:58,239
going pretty quickly here because the

2694
01:46:01,590 --> 01:46:00,480
interesting part comes after

2695
01:46:05,109 --> 01:46:01,600
but we have two distinct house

2696
01:46:07,990 --> 01:46:05,119
populations so two hosts two gsbs these

2697
01:46:10,629 --> 01:46:08,000
are two species pretty closely related

2698
01:46:11,910 --> 01:46:10,639
and then we have uh historical data

2699
01:46:14,470 --> 01:46:11,920
because we have metagenomes for the same

2700
01:46:16,629 --> 01:46:14,480
like 1505

2701
01:46:17,910 --> 01:46:16,639
and we found two viruses consistently

2702
01:46:19,830 --> 01:46:17,920
associated with each of us we actually

2703
01:46:21,669 --> 01:46:19,840
found more viruses than this and they're

2704
01:46:23,910 --> 01:46:21,679
all normal of course because no one you

2705
01:46:24,790 --> 01:46:23,920
know could isolate a phage on gsb before

2706
01:46:26,629 --> 01:46:24,800
so

2707
01:46:28,390 --> 01:46:26,639
that's kind of regular news for various

2708
01:46:29,750 --> 01:46:28,400
gender makes everything is novel but

2709
01:46:31,430 --> 01:46:29,760
what really was interesting is like

2710
01:46:34,070 --> 01:46:31,440
these four viruses in total two four

2711
01:46:35,830 --> 01:46:34,080
hours which we are found in 2005 and

2712
01:46:37,350 --> 01:46:35,840
again in 2018.

2713
01:46:38,550 --> 01:46:37,360

so of course when you have a time series

2714
01:46:42,149 --> 01:46:38,560
you want to look at what the dynamics

2715
01:46:44,470 --> 01:46:42,159
look like and the first one looks um

2716
01:46:45,910 --> 01:46:44,480
actually surprising to me

2717
01:46:48,550 --> 01:46:45,920
but i will just yeah guide you through

2718
01:46:49,910 --> 01:46:48,560
this plot so these are

2719
01:46:53,510 --> 01:46:49,920
all the difference here that were

2720
01:46:55,990 --> 01:46:53,520
sampled from 2005 2018 the y-axis is

2721
01:46:59,109 --> 01:46:56,000
relative abundance you know analog scale

2722
01:47:00,950 --> 01:46:59,119
and then you have three um curves one

2723
01:47:03,270 --> 01:47:00,960
for the host in blue and then the shades

2724
01:47:04,950 --> 01:47:03,280
of beings for the two viruses associated

2725
01:47:06,790 --> 01:47:04,960
with this first horse which we call

2726
01:47:09,109 --> 01:47:06,800
genome a because we are not very

2727
01:47:10,310 --> 01:47:09,119
original so far

2728
01:47:12,149 --> 01:47:10,320
and you can see that

2729
01:47:15,030 --> 01:47:12,159
especially after 2008

2730
01:47:16,950 --> 01:47:15,040
the curves are literally on each other

2731
01:47:19,109 --> 01:47:16,960
so it's the relative abundance of the

2732
01:47:21,270 --> 01:47:19,119
horse is exactly the same as relatively

2733
01:47:23,590 --> 01:47:21,280
balanced virus and it just go on for

2734
01:47:25,109 --> 01:47:23,600
more than 10 years until 2018 where you

2735
01:47:27,030 --> 01:47:25,119
start to have a little separation at the

2736
01:47:29,030 --> 01:47:27,040
end of 2018 with one virus becoming

2737
01:47:30,310 --> 01:47:29,040
slightly more abundant

2738
01:47:33,030 --> 01:47:30,320

so

2739
01:47:35,430 --> 01:47:33,040
that was very surprising and then we you

2740
01:47:37,430 --> 01:47:35,440
know we just filed it on the side and

2741
01:47:39,510 --> 01:47:37,440
looked at the second host and that's

2742
01:47:41,590 --> 01:47:39,520
where things become became both

2743
01:47:43,189 --> 01:47:41,600
interesting and actually more familiar

2744
01:47:44,470 --> 01:47:43,199
the second host which is here and

2745
01:47:47,669 --> 01:47:44,480
actually there is a mistake here this

2746
01:47:49,030 --> 01:47:47,679
should not be gsb genome ace genome b

2747
01:47:51,109 --> 01:47:49,040
yeah that's the title is right the

2748
01:47:53,430 --> 01:47:51,119
legend is wrong sorry

2749
01:47:56,229 --> 01:47:53,440
so again this is relative abundance the

2750
01:47:58,229 --> 01:47:56,239
genome is in red host genomes the virus

2751
01:47:59,910 --> 01:47:58,239
are in green and these times we don't

2752
01:48:01,350 --> 01:47:59,920
have a perfect correlation between

2753
01:48:03,510 --> 01:48:01,360
genome and host we actually have a lot

2754
01:48:05,270 --> 01:48:03,520
of variation we have viruses that are

2755
01:48:07,510 --> 01:48:05,280
um

2756
01:48:09,430 --> 01:48:07,520
here and then not observed anymore like

2757
01:48:11,109 --> 01:48:09,440
you can detect them into n7 this light

2758
01:48:12,310 --> 01:48:11,119
green you don't really see it in 2008

2759
01:48:13,510 --> 01:48:12,320
even if the host is here so you have

2760
01:48:15,590 --> 01:48:13,520
something happening but if you have some

2761
01:48:17,030 --> 01:48:15,600
dynamics you have some some changes

2762
01:48:19,189 --> 01:48:17,040
happening and that's what you would

2763
01:48:21,270 --> 01:48:19,199

expect for a virus and a host on such a

2764
01:48:22,629 --> 01:48:21,280
long time scale like that's that's 13

2765
01:48:24,070 --> 01:48:22,639
years

2766
01:48:26,149 --> 01:48:24,080
so this this looked a little more

2767
01:48:27,030 --> 01:48:26,159
familiar so we started to think about

2768
01:48:28,629 --> 01:48:27,040
okay

2769
01:48:31,189 --> 01:48:28,639
on one side we

2770
01:48:33,030 --> 01:48:31,199
seem to see stable association between

2771
01:48:34,310 --> 01:48:33,040
and actually happy co-existence between

2772
01:48:35,990 --> 01:48:34,320
virus and host

2773
01:48:37,669 --> 01:48:36,000
on the right side this seems to be more

2774
01:48:40,229 --> 01:48:37,679
consistent with arms race

2775
01:48:41,830 --> 01:48:40,239
um can we show this a little more we can

2776
01:48:43,669 --> 01:48:41,840
because we can assemble genomes from all

2777
01:48:45,510 --> 01:48:43,679
of these meta genomes and compare them

2778
01:48:48,070 --> 01:48:45,520
across time points

2779
01:48:49,350 --> 01:48:48,080
and these all start you know

2780
01:48:51,430 --> 01:48:49,360
kept being consistent with our first

2781
01:48:53,270 --> 01:48:51,440
hypothesis so if you assemble genome

2782
01:48:55,030 --> 01:48:53,280
from detail seven here and i'm talking

2783
01:48:56,390 --> 01:48:55,040
about the virus genome here

2784
01:48:59,189 --> 01:48:56,400
if you compare the genome symbol from

2785
01:49:00,950 --> 01:48:59,199
2007 on top to the genome compared to

2786
01:49:02,629 --> 01:49:00,960
assembling an eight here so i'm

2787
01:49:04,629 --> 01:49:02,639
basically taking the genome from the

2788
01:49:05,910 --> 01:49:04,639

virus here and the genome of the virus

2789
01:49:11,669 --> 01:49:05,920
here

2790
01:49:14,149 --> 01:49:11,679
any snip like it's it's 99 points

2791
01:49:15,669 --> 01:49:14,159
something percent identity identical at

2792
01:49:17,350 --> 01:49:15,679
the nucleotide level so basically

2793
01:49:19,109 --> 01:49:17,360
nothing change they just stay with the

2794
01:49:20,550 --> 01:49:19,119
host and and they are happy to stay with

2795
01:49:22,229 --> 01:49:20,560
the host

2796
01:49:23,910 --> 01:49:22,239
on the other end

2797
01:49:25,669 --> 01:49:23,920
as you would expect things are much more

2798
01:49:30,149 --> 01:49:25,679
complicated on the uh

2799
01:49:32,390 --> 01:49:30,159
genome b and its viruses this is this um

2800
01:49:33,990 --> 01:49:32,400
light green virus and you can see we

2801
01:49:37,270 --> 01:49:34,000
have some older genome or a variant of

2802
01:49:38,709 --> 01:49:37,280
this genome in 2007 in 2017's and 18.

2803
01:49:40,629 --> 01:49:38,719
we never see any of this variant

2804
01:49:42,950 --> 01:49:40,639
coexisting we see them replacing each

2805
01:49:44,950 --> 01:49:42,960
other but we see clear differences from

2806
01:49:46,870 --> 01:49:44,960
one ear to the next you have some parts

2807
01:49:49,589 --> 01:49:46,880
of the genome here on the three prime

2808
01:49:51,030 --> 01:49:49,599
right side they're clearly conserved

2809
01:49:52,709 --> 01:49:51,040
and these are the genes that are

2810
01:49:54,310 --> 01:49:52,719
involved in capsid and you know the

2811
01:49:55,510 --> 01:49:54,320
structural proteins i saw the gene that

2812
01:49:57,830 --> 01:49:55,520
you would expect

2813
01:49:59,350 --> 01:49:57,840

you know to be conserved um

2814
01:50:00,870 --> 01:49:59,360
at least the most conserved but then all

2815
01:50:03,030 --> 01:50:00,880
of these genes on the left on the five

2816
01:50:04,070 --> 01:50:03,040
prime end of the genome as it is um

2817
01:50:05,750 --> 01:50:04,080
right now

2818
01:50:07,830 --> 01:50:05,760
you have tons of small genes

2819
01:50:09,510 --> 01:50:07,840
hypothetical proteins pretty much thing

2820
01:50:11,510 --> 01:50:09,520
that we don't really know what they do

2821
01:50:12,790 --> 01:50:11,520
and this seems to be turned over every

2822
01:50:15,270 --> 01:50:12,800
year

2823
01:50:17,189 --> 01:50:15,280
most likely this recombination

2824
01:50:18,070 --> 01:50:17,199
so we end up with this weird situation

2825
01:50:19,589 --> 01:50:18,080
where

2826
01:50:21,669 --> 01:50:19,599
we have two hosts

2827
01:50:23,669 --> 01:50:21,679
two species of gsb we looked at the

2828
01:50:26,149 --> 01:50:23,679
genomes they looked you know having the

2829
01:50:27,510 --> 01:50:26,159
same metabolic profiles same capacity to

2830
01:50:30,310 --> 01:50:27,520
do an oxygenic photosynthesis and

2831
01:50:33,030 --> 01:50:30,320
everything so they are you know re

2832
01:50:34,470 --> 01:50:33,040
really close and similar hosts they are

2833
01:50:36,229 --> 01:50:34,480
under the exact same on horizontal

2834
01:50:38,229 --> 01:50:36,239
constraints they are in the same lake

2835
01:50:41,189 --> 01:50:38,239
for the last 15 years and they have to

2836
01:50:42,070 --> 01:50:41,199
drastically different infection types

2837
01:50:44,550 --> 01:50:42,080
so

2838
01:50:46,390 --> 01:50:44,560

the last pretty much nine months we try

2839
01:50:47,589 --> 01:50:46,400
to get at this question of like why

2840
01:50:48,790 --> 01:50:47,599
would they have different infection

2841
01:50:50,709 --> 01:50:48,800
types

2842
01:50:51,510 --> 01:50:50,719
and we ended up on one hypothesis which

2843
01:50:53,270 --> 01:50:51,520
is

2844
01:50:55,350 --> 01:50:53,280
microdiversity of the host and we think

2845
01:50:56,470 --> 01:50:55,360
this is something that has been

2846
01:50:58,790 --> 01:50:56,480
under

2847
01:51:01,270 --> 01:50:58,800
studied in this field of

2848
01:51:02,950 --> 01:51:01,280
lytic versus lysogenic cycles

2849
01:51:04,709 --> 01:51:02,960
this is my last graph

2850
01:51:07,350 --> 01:51:04,719
it's a little complicated so let me just

2851
01:51:09,990 --> 01:51:07,360
walk you through it and bear with me

2852
01:51:11,830 --> 01:51:10,000
x-axis is a coverage so in this case

2853
01:51:14,149 --> 01:51:11,840
it's it's abundance right so if you're

2854
01:51:15,990 --> 01:51:14,159
on the left you are your genome is not

2855
01:51:17,510 --> 01:51:16,000
really abundant on the right it's highly

2856
01:51:20,550 --> 01:51:17,520
abundant

2857
01:51:21,589 --> 01:51:20,560
y-axis is nucleotide diversity so it's a

2858
01:51:23,750 --> 01:51:21,599
sense of

2859
01:51:25,350 --> 01:51:23,760
how many different populat populations

2860
01:51:27,510 --> 01:51:25,360
or how many different strengths do you

2861
01:51:29,589 --> 01:51:27,520
have within your population another way

2862
01:51:30,870 --> 01:51:29,599
of saying this is how many snips do you

2863
01:51:33,350 --> 01:51:30,880

have how many different alleles do you

2864
01:51:35,189 --> 01:51:33,360
have per snips and how do you have a

2865
01:51:36,950 --> 01:51:35,199
dominant allele or do you have pretty

2866
01:51:38,790 --> 01:51:36,960
much evenly distributed allele across

2867
01:51:40,709 --> 01:51:38,800
your strains

2868
01:51:43,350 --> 01:51:40,719
for each genome and i'm talking about

2869
01:51:45,430 --> 01:51:43,360
the host genome here there is one point

2870
01:51:46,709 --> 01:51:45,440
per for each year and then you have like

2871
01:51:48,790 --> 01:51:46,719
this

2872
01:51:51,430 --> 01:51:48,800
bar that just represents a spread during

2873
01:51:54,390 --> 01:51:51,440
the year of each value so for example

2874
01:51:56,629 --> 01:51:54,400
here for genome a in blue in seven the

2875
01:51:59,669 --> 01:51:56,639
code rate range from here to here and

2876
01:52:01,109 --> 01:51:59,679
the diversity range from here to here

2877
01:52:02,629 --> 01:52:01,119
the reason why i'm presenting abundant

2878
01:52:03,910 --> 01:52:02,639
versus diversity

2879
01:52:05,430 --> 01:52:03,920
is because you expect a correlation

2880
01:52:07,589 --> 01:52:05,440
between them and that's just a metal

2881
01:52:11,030 --> 01:52:07,599
thing it's like if you don't have a lot

2882
01:52:13,510 --> 01:52:11,040
of reads for your population you just

2883
01:52:15,990 --> 01:52:13,520
can't see any rare variant

2884
01:52:17,430 --> 01:52:16,000
so as you increase your coverage and as

2885
01:52:18,629 --> 01:52:17,440
you have more reads and more rather than

2886
01:52:19,830 --> 01:52:18,639
you sequence more deeply into your

2887
01:52:21,990 --> 01:52:19,840
population

2888
01:52:23,910 --> 01:52:22,000

you should see more and more nucleotide

2889
01:52:25,910 --> 01:52:23,920
diversity and that's pretty much what we

2890
01:52:27,270 --> 01:52:25,920
see for genome b again remember genome b

2891
01:52:29,830 --> 01:52:27,280
is this genome where we have this

2892
01:52:30,870 --> 01:52:29,840
expected arms race dynamic with the

2893
01:52:32,629 --> 01:52:30,880
virus

2894
01:52:34,470 --> 01:52:32,639
we have cases where it's low coverage

2895
01:52:36,149 --> 01:52:34,480
low diversity and the two years where we

2896
01:52:37,910 --> 01:52:36,159
actually have a lot of diversity a lot

2897
01:52:41,430 --> 01:52:37,920
of coverage we actually see

2898
01:52:43,350 --> 01:52:41,440
a good amount of diversity as expected

2899
01:52:46,070 --> 01:52:43,360
genome a is different

2900
01:52:48,550 --> 01:52:46,080
because genome has this we are very very

2901
01:52:50,629 --> 01:52:48,560
high coverage low diversity region here

2902
01:52:52,470 --> 01:52:50,639
and these are the years where we found

2903
01:52:57,350 --> 01:52:52,480
the exact correlation between virus and

2904
01:52:59,030 --> 01:52:57,360
host abundance 2008 9 12 13.

2905
01:53:01,350 --> 01:52:59,040
and it's actually something that has

2906
01:53:03,669 --> 01:53:01,360
been reported before by a paper by uh

2907
01:53:06,390 --> 01:53:03,679
matthew bendel from jgi when they were

2908
01:53:07,910 --> 01:53:06,400
looking only at the

2909
01:53:09,910 --> 01:53:07,920
time series meta genomes and the

2910
01:53:11,510 --> 01:53:09,920
microbial fraction and that what they

2911
01:53:13,910 --> 01:53:11,520
called this was a genome-wide sweep

2912
01:53:16,149 --> 01:53:13,920
which is another way of saying

2913
01:53:18,870 --> 01:53:16,159

this population this genome a become

2914
01:53:21,750 --> 01:53:18,880
clonal from 2008-13

2915
01:53:24,550 --> 01:53:21,760
no snip anymore it's like a single

2916
01:53:26,709 --> 01:53:24,560
very very low diversity population

2917
01:53:29,030 --> 01:53:26,719
and that's where we are start we started

2918
01:53:31,270 --> 01:53:29,040
to say well that may be a key drivers of

2919
01:53:33,510 --> 01:53:31,280
why we see latent infections in this

2920
01:53:34,790 --> 01:53:33,520
lysogenic long-term stable coexistence

2921
01:53:36,709 --> 01:53:34,800
with viruses

2922
01:53:39,030 --> 01:53:36,719
and so my model will go somewhere

2923
01:53:40,629 --> 01:53:39,040
something like this

2924
01:53:42,790 --> 01:53:40,639
this is a picture we always think of

2925
01:53:44,229 --> 01:53:42,800
when we think of our resource right arms

2926
01:53:47,189 --> 01:53:44,239
race

2927
01:53:48,870 --> 01:53:47,199
super quick dynamics in in slightly more

2928
01:53:52,470 --> 01:53:48,880
scientific terms

2929
01:53:53,990 --> 01:53:52,480
what we see is diverse population

2930
01:53:55,750 --> 01:53:54,000
we think arms race because we think

2931
01:53:57,109 --> 01:53:55,760
there is variation in invasive stability

2932
01:53:58,550 --> 01:53:57,119
between population members so some

2933
01:54:00,149 --> 01:53:58,560
population members are more or less

2934
01:54:02,550 --> 01:54:00,159
susceptible which means you can have

2935
01:54:03,910 --> 01:54:02,560
some resistance arising and then if

2936
01:54:06,310 --> 01:54:03,920
there are resistance arising the

2937
01:54:08,149 --> 01:54:06,320
resistance strains will will develop and

2938
01:54:10,070 --> 01:54:08,159

propagate and then and proliferate and

2939
01:54:11,510 --> 01:54:10,080
then the fade we likely have a counter

2940
01:54:13,669 --> 01:54:11,520
resistance etcetera etcetera so it's

2941
01:54:15,510 --> 01:54:13,679
electron new phage variants and that's

2942
01:54:17,910 --> 01:54:15,520
what gives you the kills the winner

2943
01:54:19,350 --> 01:54:17,920
dynamic and answers and everything

2944
01:54:21,109 --> 01:54:19,360
and that's dynamic associated with lytic

2945
01:54:22,709 --> 01:54:21,119
phages or short latency that's what we

2946
01:54:24,310 --> 01:54:22,719
see in genome b which has a diverse

2947
01:54:26,149 --> 01:54:24,320
population

2948
01:54:28,950 --> 01:54:26,159
but it seems like

2949
01:54:30,709 --> 01:54:28,960
things can also look something like this

2950
01:54:32,390 --> 01:54:30,719
and again to my surprise that's really

2951
01:54:33,270 --> 01:54:32,400
not what i expected but

2952
01:54:34,550 --> 01:54:33,280
um

2953
01:54:37,030 --> 01:54:34,560
it seems like if you have a low

2954
01:54:38,629 --> 01:54:37,040
diversity host population with most of

2955
01:54:41,669 --> 01:54:38,639
your population members having the same

2956
01:54:43,830 --> 01:54:41,679
susceptibility to the phage

2957
01:54:45,510 --> 01:54:43,840
it's way more likely that the phage is

2958
01:54:47,430 --> 01:54:45,520
able to infect every member of your

2959
01:54:50,310 --> 01:54:47,440
population before

2960
01:54:51,910 --> 01:54:50,320
any mutant can arise at least a strong

2961
01:54:53,510 --> 01:54:51,920
resistant mutant can arise and that's

2962
01:54:55,589 --> 01:54:53,520
also tied into what brit was talking

2963
01:54:57,270 --> 01:54:55,599

about yesterday in this

2964
01:54:59,830 --> 01:54:57,280
possibility of misunderstanding or

2965
01:55:01,510 --> 01:54:59,840
resistance rising in nature compared to

2966
01:55:03,589 --> 01:55:01,520
in test tube it seems to be more

2967
01:55:06,149 --> 01:55:03,599
complicated for this kind of resistance

2968
01:55:08,070 --> 01:55:06,159
to actually occur in nature and in our

2969
01:55:09,669 --> 01:55:08,080
case in plantar

2970
01:55:12,470 --> 01:55:09,679
and that's what we see here like there

2971
01:55:14,229 --> 01:55:12,480
is no resistance and it seems like the

2972
01:55:15,750 --> 01:55:14,239
phages infect every every member of the

2973
01:55:18,070 --> 01:55:15,760
community and at this point this kind of

2974
01:55:21,030 --> 01:55:18,080
situation will be selecting for temp

2975
01:55:23,589 --> 01:55:21,040
rate and long latency phases um pretty

2976
01:55:26,149 --> 01:55:23,599
much a phage infecting 100 of its host

2977
01:55:27,830 --> 01:55:26,159
in a given environment has won the game

2978
01:55:30,629 --> 01:55:27,840
and and it's you know doesn't need to do

2979
01:55:34,870 --> 01:55:32,870
so we have some model working

2980
01:55:36,629 --> 01:55:34,880
in progress the idea would be like is

2981
01:55:38,149 --> 01:55:36,639
this even plausible and can we can we

2982
01:55:40,470 --> 01:55:38,159
show the switch between selection for

2983
01:55:43,669 --> 01:55:40,480
lytic versus solution for lysogenic

2984
01:55:45,669 --> 01:55:43,679
just based on host population diversity

2985
01:55:48,550 --> 01:55:45,679
but there is a key message i want to i

2986
01:55:49,750 --> 01:55:48,560
want to kind of convene here is

2987
01:55:51,189 --> 01:55:49,760
this

2988
01:55:52,470 --> 01:55:51,199

feature host population genetic

2989
01:55:53,990 --> 01:55:52,480
diversity

2990
01:55:56,629 --> 01:55:54,000
has not really been taken into

2991
01:55:58,070 --> 01:55:56,639
consideration into litigation decision

2992
01:55:59,510 --> 01:55:58,080
and i think it should be because it

2993
01:56:01,589 --> 01:55:59,520
might be a critical driver of our

2994
01:56:03,669 --> 01:56:01,599
resource dynamics

2995
01:56:05,030 --> 01:56:03,679
and with this i'm sorry i'm slightly

2996
01:56:07,270 --> 01:56:05,040
over time

2997
01:56:08,709 --> 01:56:07,280
there's a very broad conclusion

2998
01:56:10,470 --> 01:56:08,719
we can use metagenomic for various

2999
01:56:11,910 --> 01:56:10,480
discovery we have a framework for this

3000
01:56:13,350 --> 01:56:11,920
now again we have standards paper that

3001
01:56:14,870 --> 01:56:13,360
just came out

3002
01:56:16,470 --> 01:56:14,880
go and read this if you want to start in

3003
01:56:17,830 --> 01:56:16,480
this field

3004
01:56:19,910 --> 01:56:17,840
house interactions are key to understand

3005
01:56:22,070 --> 01:56:19,920
viruses host linkage is a major

3006
01:56:23,750 --> 01:56:22,080
challenge and just remember that various

3007
01:56:25,830 --> 01:56:23,760
interactions come in way more flavor

3008
01:56:28,470 --> 01:56:25,840
than just a very center kill the cell

3009
01:56:30,550 --> 01:56:28,480
and then the large

3010
01:56:32,470 --> 01:56:30,560
finally just want to make you know

3011
01:56:34,149 --> 01:56:32,480
everyone

3012
01:56:35,589 --> 01:56:34,159
as excited as i am about the virus world

3013
01:56:37,189 --> 01:56:35,599

but basically we have so many

3014
01:56:39,430 --> 01:56:37,199
discoveries to be made in all of these

3015
01:56:41,589 --> 01:56:39,440
unknown viruses i've shown you this 95

3016
01:56:43,189 --> 01:56:41,599
of viruses for which we have no host id

3017
01:56:45,270 --> 01:56:43,199
but we have very little idea about their

3018
01:56:47,189 --> 01:56:45,280
um genome itself in terms of gene

3019
01:56:49,189 --> 01:56:47,199
content we have tons of genes without

3020
01:56:51,270 --> 01:56:49,199
any functions that are still highly

3021
01:56:52,950 --> 01:56:51,280
conserved so we think they are important

3022
01:56:54,390 --> 01:56:52,960
so yeah if you're looking for a question

3023
01:56:58,149 --> 01:56:54,400
or if you're looking for an open field

3024
01:57:00,470 --> 01:56:58,159
or a black box this is pretty much it

3025
01:57:02,229 --> 01:57:00,480
if you want to know more or to do more a

3026
01:57:04,470 --> 01:57:02,239
few

3027
01:57:06,870 --> 01:57:04,480
advertisements plugging in what we are

3028
01:57:08,629 --> 01:57:06,880
doing next uh if you want to know more

3029
01:57:10,790 --> 01:57:08,639
about viruses and hear from this very

3030
01:57:12,709 --> 01:57:10,800
very great speakers on the right you can

3031
01:57:16,790 --> 01:57:12,719
join us in oakland in 2020 we have this

3032
01:57:19,189 --> 01:57:16,800
bigger symposium march 22 and march 23.

3033
01:57:21,350 --> 01:57:19,199
uh registration will open soon so um you

3034
01:57:22,870 --> 01:57:21,360
know stay tuned i will probably retweet

3035
01:57:24,709 --> 01:57:22,880
this heavily on on twitter and

3036
01:57:26,310 --> 01:57:24,719
everything so just know that this vegas

3037
01:57:27,589 --> 01:57:26,320
impossible is coming and this should be

3038
01:57:29,510 --> 01:57:27,599

pretty good

3039
01:57:30,870 --> 01:57:29,520
if you want to um

3040
01:57:33,109 --> 01:57:30,880
learn how to lose by informatics

3041
01:57:35,030 --> 01:57:33,119
yourself we are hosting what we call mgm

3042
01:57:36,950 --> 01:57:35,040
workshops microbial genome and

3043
01:57:38,870 --> 01:57:36,960
microbiomes there is a small portion

3044
01:57:41,109 --> 01:57:38,880
about this about viruses so you would

3045
01:57:42,550 --> 01:57:41,119
also learn how to do virus analysis from

3046
01:57:44,550 --> 01:57:42,560
metagenomes

3047
01:57:46,229 --> 01:57:44,560
um these are very cool workshops we have

3048
01:57:48,709 --> 01:57:46,239
them every six months so if you want

3049
01:57:50,229 --> 01:57:48,719
some bioinformatics training in

3050
01:57:52,070 --> 01:57:50,239
microbiomes

3051
01:57:54,709 --> 01:57:52,080
just think about this and and keep keep

3052
01:57:56,070 --> 01:57:54,719
this address in mind and lately if you

3053
01:57:57,990 --> 01:57:56,080
have very very cool samples and you

3054
01:57:59,430 --> 01:57:58,000
actually want someone to sequence meta

3055
01:58:01,910 --> 01:57:59,440
genomes for you

3056
01:58:04,070 --> 01:58:01,920
that's what gti is here for so we have a

3057
01:58:06,629 --> 01:58:04,080
next our next call is uh for set number

3058
01:58:07,669 --> 01:58:06,639
26. uh i won't go into any detail but if

3059
01:58:09,430 --> 01:58:07,679
anyone

3060
01:58:10,870 --> 01:58:09,440
has some samples i want to look into

3061
01:58:12,310 --> 01:58:10,880
viruses or they want even to look at

3062
01:58:14,229 --> 01:58:12,320
microbes

3063
01:58:15,910 --> 01:58:14,239

you know just send me an email hit me on

3064
01:58:17,990 --> 01:58:15,920
twitter and i can i can you know you

3065
01:58:20,070 --> 01:58:18,000
know either connect to the right person

3066
01:58:21,750 --> 01:58:20,080
or just tell you you know what what is

3067
01:58:23,109 --> 01:58:21,760
the framework in which you can apply for

3068
01:58:24,709 --> 01:58:23,119
this kind of grants

3069
01:58:26,790 --> 01:58:24,719
and with that i just want to thank

3070
01:58:29,270 --> 01:58:26,800
everyone especially involved in this um

3071
01:58:31,510 --> 01:58:29,280
gsb story which has all been led by

3072
01:58:33,350 --> 01:58:31,520
maureen my postdoc with a great help of

3073
01:58:35,030 --> 01:58:33,360
danielle for the first centometry and

3074
01:58:43,270 --> 01:58:35,040
the lab of trina mcmahon in wisconsin

3075
01:58:43,280 --> 01:58:48,149
where is my chat

3076
01:58:55,430 --> 01:58:51,109
okay i think we've got time for maybe a

3077
01:58:56,870 --> 01:58:55,440
question um if we can do it quickly

3078
01:58:59,270 --> 01:58:56,880
if not i'm sure that simon would be more

3079
01:59:07,589 --> 01:58:59,280
than happy to uh answer questions

3080
01:59:13,270 --> 01:59:10,229
just really quickly simon if i may did

3081
01:59:15,750 --> 01:59:13,280
you notice anything specific about those

3082
01:59:17,270 --> 01:59:15,760
two populations of the gsbs was there

3083
01:59:18,629 --> 01:59:17,280
something about

3084
01:59:22,709 --> 01:59:18,639
the

3085
01:59:24,310 --> 01:59:22,719
noticed oh crispers or something else

3086
01:59:25,669 --> 01:59:24,320
that was you know variable or less

3087
01:59:27,830 --> 01:59:25,679
variable

3088
01:59:29,669 --> 01:59:27,840

so that's a good question

3089
01:59:31,189 --> 01:59:29,679
we didn't find any key feature that

3090
01:59:33,990 --> 01:59:31,199
would explain the difference between the

3091
01:59:37,030 --> 01:59:34,000
viruses one issue we have is that they

3092
01:59:39,430 --> 01:59:37,040
actually are so closely related

3093
01:59:41,350 --> 01:59:39,440
they both have a crispr array

3094
01:59:43,350 --> 01:59:41,360
and these crispr arrays have the exact

3095
01:59:45,669 --> 01:59:43,360
same repeat

3096
01:59:47,750 --> 01:59:45,679
you do energetic assembly you can't tell

3097
01:59:50,470 --> 01:59:47,760
if the space or link to repeat comes

3098
01:59:53,189 --> 01:59:50,480
from genome a or genome b basically got

3099
01:59:55,830 --> 01:59:53,199
it that's how closely related they are

3100
01:59:57,589 --> 01:59:55,840
so at this point um i mean we looked

3101
01:59:59,830 --> 01:59:57,599
into these genomes and and we couldn't

3102
02:00:02,470 --> 01:59:59,840
find any good explanation until we

3103
02:00:06,070 --> 02:00:02,480
looked into this population diversity

3104
02:00:09,350 --> 02:00:07,669
okay so catherine you're up you're up

3105
02:00:11,589 --> 02:00:09,360
next

3106
02:00:15,750 --> 02:00:11,599
slightly good program

3107
02:00:15,760 --> 02:00:21,589
see if i can

3108
02:00:26,310 --> 02:00:23,830
okay so first i have to apologize i am

3109
02:00:29,669 --> 02:00:26,320
not gary

3110
02:00:31,669 --> 02:00:29,679
i had a conflict that arose yesterday i

3111
02:00:35,030 --> 02:00:31,679
and of course out of all of the hours of

3112
02:00:37,589 --> 02:00:35,040
all of the day it came up that it was

3113
02:00:40,470 --> 02:00:37,599

going to happen from 3 30 to 4 of course

3114
02:00:42,070 --> 02:00:40,480
so uh gary was happy enough to switch

3115
02:00:45,270 --> 02:00:42,080
with me thank you

3116
02:00:47,669 --> 02:00:45,280
so i do not work on viruses but after

3117
02:00:50,229 --> 02:00:47,679
this i'm convinced that i should be

3118
02:00:52,229 --> 02:00:50,239
and i'm gonna talk gonna switch gears a

3119
02:00:54,550 --> 02:00:52,239
little bit and talk about a framework

3120
02:00:57,109 --> 02:00:54,560
for thinking about life detection

3121
02:00:59,589 --> 02:00:57,119
and then viruses as bio signatures and

3122
02:01:00,629 --> 02:00:59,599
go into instrumentation

3123
02:01:01,910 --> 02:01:00,639
of

3124
02:01:03,830 --> 02:01:01,920
how you might

3125
02:01:05,669 --> 02:01:03,840
kind of give give people i don't know

3126
02:01:08,790 --> 02:01:05,679
the background of how many people have

3127
02:01:10,790 --> 02:01:08,800
actually worked on proposals for flight

3128
02:01:14,070 --> 02:01:10,800
instruments but kind of give the

3129
02:01:15,750 --> 02:01:14,080
audience a idea of some of the avenues

3130
02:01:17,430 --> 02:01:15,760
and the requirements

3131
02:01:19,589 --> 02:01:17,440
and then talk about the technology that

3132
02:01:21,510 --> 02:01:19,599
i'm currently using

3133
02:01:25,589 --> 02:01:21,520
so

3134
02:01:27,510 --> 02:01:25,599
if we start looking at uh astrobiology

3135
02:01:29,910 --> 02:01:27,520
it's the search for life's origins

3136
02:01:32,470 --> 02:01:29,920
evolution distribution and future in the

3137
02:01:36,550 --> 02:01:32,480
universe and especially after sitting in

3138
02:01:38,950 --> 02:01:36,560

in the last day or so on this viruses

3139
02:01:41,189 --> 02:01:38,960
i've learned really play a role in every

3140
02:01:43,990 --> 02:01:41,199
single aspect of this question and it

3141
02:01:46,070 --> 02:01:44,000
seems so integral to this question it's

3142
02:01:48,229 --> 02:01:46,080
i'm really happy this meeting's taking

3143
02:01:50,070 --> 02:01:48,239
place and it's a subject i think needs

3144
02:01:53,510 --> 02:01:50,080
to be pushed and we need to be looking

3145
02:01:55,430 --> 02:01:53,520
for viruses everywhere it sounds like uh

3146
02:01:59,510 --> 02:01:55,440
so with that we'll need tools and i'll

3147
02:02:02,550 --> 02:01:59,520
get to that in the technology part uh so

3148
02:02:04,229 --> 02:02:02,560
to understand uh the distribution if we

3149
02:02:06,870 --> 02:02:04,239
go out looking for life for life

3150
02:02:09,430 --> 02:02:06,880
detection i it's kind of nice to know

3151
02:02:12,149 --> 02:02:09,440
what you're looking for so it does pull

3152
02:02:13,350 --> 02:02:12,159
us back to that question of what is life

3153
02:02:15,430 --> 02:02:13,360
and

3154
02:02:18,390 --> 02:02:15,440
you the the

3155
02:02:20,070 --> 02:02:18,400
way our theory for life right now is

3156
02:02:22,390 --> 02:02:20,080
we're defining it based on its

3157
02:02:25,109 --> 02:02:22,400
properties and you can kind of think

3158
02:02:27,510 --> 02:02:25,119
about this of how we used to define

3159
02:02:30,390 --> 02:02:27,520
water before we really knew what water

3160
02:02:33,430 --> 02:02:30,400
was uh before we knew it was h2o we said

3161
02:02:35,350 --> 02:02:33,440
it was clear colorless wet you know it

3162
02:02:37,990 --> 02:02:35,360
had we defined it based on its

3163
02:02:39,910 --> 02:02:38,000

properties and we do the same thing

3164
02:02:42,070 --> 02:02:39,920
with life now it's self-enclosed

3165
02:02:44,950 --> 02:02:42,080
self-sustained chemical system capable

3166
02:02:46,709 --> 02:02:44,960
of undergoing darwinian evolution so we

3167
02:02:49,350 --> 02:02:46,719
really are

3168
02:02:52,550 --> 02:02:49,360
when we talk about life detection

3169
02:02:54,470 --> 02:02:52,560
we talk about biosignatures biomolecules

3170
02:02:57,750 --> 02:02:54,480
because we're looking for these

3171
02:02:59,350 --> 02:02:57,760
properties of life um

3172
02:03:00,470 --> 02:02:59,360
and that's why if you're here for the

3173
02:03:02,470 --> 02:03:00,480
first day

3174
02:03:06,550 --> 02:03:02,480
uh benner this kind of goes back to

3175
02:03:09,189 --> 02:03:06,560
benner's talk about the um

3176
02:03:11,430 --> 02:03:09,199
the universal signature for life

3177
02:03:14,709 --> 02:03:11,440
uh and i'll get into that in in just a

3178
02:03:17,830 --> 02:03:14,719
little bit announce it has put out a

3179
02:03:20,709 --> 02:03:17,840
life detection ladder uh it's a nice

3180
02:03:22,950 --> 02:03:20,719
beginning a framework uh to think about

3181
02:03:24,550 --> 02:03:22,960
how we might go about looking for these

3182
02:03:26,390 --> 02:03:24,560
properties of life

3183
02:03:28,149 --> 02:03:26,400
if you haven't checked it out i would i

3184
02:03:30,790 --> 02:03:28,159
would recommend just familiarizing

3185
02:03:33,510 --> 02:03:30,800
yourself with it uh so there are rungs

3186
02:03:36,070 --> 02:03:33,520
on this life detection ladder and the

3187
02:03:38,149 --> 02:03:36,080
top of the rung is the highest evidence

3188
02:03:39,830 --> 02:03:38,159

for life and that would be darwinian

3189
02:03:41,109 --> 02:03:39,840
evolution right because that's in our

3190
02:03:43,750 --> 02:03:41,119
definition

3191
02:03:46,790 --> 02:03:43,760
uh the problem with that is if you go to

3192
02:03:49,669 --> 02:03:46,800
mars or europa enceladus or titan

3193
02:03:51,370 --> 02:03:49,679
observing darwinian evolution

3194
02:03:53,510 --> 02:03:51,380
there we go there we go

3195
02:03:55,990 --> 02:03:53,520
[Laughter]

3196
02:03:58,709 --> 02:03:56,000
might be mildly difficult and if you do

3197
02:04:01,109 --> 02:03:58,719
look at the life detection ladder uh

3198
02:04:03,030 --> 02:04:01,119
there is you know what measurements you

3199
02:04:05,189 --> 02:04:03,040
would make for all of these proposed

3200
02:04:07,510 --> 02:04:05,199
instruments uh another one would be

3201
02:04:09,669 --> 02:04:07,520
growth and reproduction

3202
02:04:12,149 --> 02:04:09,679
metabolism

3203
02:04:15,350 --> 02:04:12,159
molecules and structures conferring

3204
02:04:19,350 --> 02:04:15,360
function i really think viruses fit into

3205
02:04:21,750 --> 02:04:19,360
this role i and that is

3206
02:04:23,510 --> 02:04:21,760
you know if you go into

3207
02:04:25,589 --> 02:04:23,520
molecules and structure conferring

3208
02:04:27,750 --> 02:04:25,599
function the polymers that support

3209
02:04:30,229 --> 02:04:27,760
information storage and transfer at

3210
02:04:34,629 --> 02:04:30,239
least here on unitarian life

3211
02:04:36,310 --> 02:04:34,639
is dna and rna and viruses are little

3212
02:04:40,149 --> 02:04:36,320
storage units

3213
02:04:43,109 --> 02:04:40,159

in my my non-virus terms i just think of

3214
02:04:46,390 --> 02:04:43,119
them as you know packages of these

3215
02:04:48,310 --> 02:04:46,400
molecules that confer confer function so

3216
02:04:49,910 --> 02:04:48,320
of course there's something

3217
02:04:53,030 --> 02:04:49,920
based on this i would think there'd be

3218
02:04:56,149 --> 02:04:53,040
an interest in looking for uh even if

3219
02:04:59,189 --> 02:04:56,159
and the nasa astrobiology strategic plan

3220
02:05:01,990 --> 02:04:59,199
in 2015 did comment on weird life and

3221
02:05:04,229 --> 02:05:02,000
recognize that alien biochemistry might

3222
02:05:07,669 --> 02:05:04,239
not be the same chemistry that we see

3223
02:05:10,470 --> 02:05:07,679
here on earth but you would still need

3224
02:05:13,270 --> 02:05:10,480
to have a

3225
02:05:16,629 --> 02:05:13,280
molecule that stores information be this

3226
02:05:18,550 --> 02:05:16,639
rna dna or something completely

3227
02:05:20,229 --> 02:05:18,560
different than these molecules they'll

3228
02:05:22,310 --> 02:05:20,239
still need to be functional molecules

3229
02:05:23,510 --> 02:05:22,320
and even if viruses aren't as we know

3230
02:05:26,950 --> 02:05:23,520
them

3231
02:05:29,510 --> 02:05:26,960
we would could assume or begin to think

3232
02:05:32,310 --> 02:05:29,520
about them still encoding for

3233
02:05:33,669 --> 02:05:32,320
or storing whatever this information

3234
02:05:34,629 --> 02:05:33,679
storage is

3235
02:05:36,790 --> 02:05:34,639
so

3236
02:05:39,109 --> 02:05:36,800
uh here we have better for the universal

3237
02:05:42,629 --> 02:05:39,119
feature of life uh benner

3238
02:05:43,830 --> 02:05:42,639

based uh going back to his presentation

3239
02:05:45,990 --> 02:05:43,840
you know these are the molecules

3240
02:05:48,629 --> 02:05:46,000
encoding genetic data necessary for

3241
02:05:49,350 --> 02:05:48,639
functioning and replication of life

3242
02:05:51,109 --> 02:05:49,360
and

3243
02:05:57,189 --> 02:05:51,119
there he makes the argument for those

3244
02:05:57,199 --> 02:06:00,470
so

3245
02:06:04,310 --> 02:06:02,069
that just is kind of to give you the

3246
02:06:06,229 --> 02:06:04,320
framework about how

3247
02:06:08,950 --> 02:06:06,239
life if you're thinking about life

3248
02:06:11,270 --> 02:06:08,960
detection and proposing life detection

3249
02:06:13,910 --> 02:06:11,280
instruments or being on a life detection

3250
02:06:14,629 --> 02:06:13,920
mission kind of need to relate it back

3251
02:06:17,109 --> 02:06:14,639
to

3252
02:06:19,270 --> 02:06:17,119
sort of these benchmarks and these goals

3253
02:06:21,109 --> 02:06:19,280
that nasa has put forward

3254
02:06:23,430 --> 02:06:21,119
and

3255
02:06:26,149 --> 02:06:23,440
for life detection missions especially

3256
02:06:28,950 --> 02:06:26,159
for viruses i'm sure there's a way to

3257
02:06:31,750 --> 02:06:28,960
look at them in exoplanets but i focus

3258
02:06:35,350 --> 02:06:31,760
on instrumentation that would work in

3259
02:06:37,270 --> 02:06:35,360
this solar system i would like to see a

3260
02:06:39,750 --> 02:06:37,280
life detection mission within the solar

3261
02:06:42,069 --> 02:06:39,760
system before i dice or

3262
02:06:44,550 --> 02:06:42,079
like to touch a mission before i die so

3263
02:06:47,350 --> 02:06:44,560

i focus on this solar system

3264
02:06:49,189 --> 02:06:47,360
i with that you could look for intact

3265
02:06:51,350 --> 02:06:49,199
detection which would be

3266
02:06:52,709 --> 02:06:51,360
high resolution microscopy and these are

3267
02:06:54,950 --> 02:06:52,719
just some

3268
02:06:57,270 --> 02:06:54,960
um this is not

3269
02:06:58,870 --> 02:06:57,280
all inclusive this is not encompassing

3270
02:07:01,750 --> 02:06:58,880
of all the different technologies that

3271
02:07:03,589 --> 02:07:01,760
you could use to look for viruses this

3272
02:07:05,830 --> 02:07:03,599
is just me throwing some some

3273
02:07:09,270 --> 02:07:05,840
suggestions out there

3274
02:07:11,189 --> 02:07:09,280
so intact intact uh detection you have

3275
02:07:14,310 --> 02:07:11,199
microscopy you could look for protein

3276
02:07:16,790 --> 02:07:14,320
protein fluorescence uh and then

3277
02:07:18,069 --> 02:07:16,800
nanopore based electrical sensing and

3278
02:07:19,910 --> 02:07:18,079
i'll get that's what the rest of the

3279
02:07:20,870 --> 02:07:19,920
talk will be about

3280
02:07:23,430 --> 02:07:20,880
and then

3281
02:07:27,350 --> 02:07:23,440
if you're wanting to you know break open

3282
02:07:29,109 --> 02:07:27,360
the capsids oil extraction and look for

3283
02:07:31,030 --> 02:07:29,119
functional molecules

3284
02:07:33,589 --> 02:07:31,040
benner proposed an instrument in his

3285
02:07:35,830 --> 02:07:33,599
talk and then nano based uh electrical

3286
02:07:37,910 --> 02:07:35,840
sensing as well so

3287
02:07:40,470 --> 02:07:37,920
just hope that the community starts

3288
02:07:43,830 --> 02:07:40,480

thinking about we need to do life

3289
02:07:46,310 --> 02:07:43,840
detection missions that include viruses

3290
02:07:48,870 --> 02:07:46,320
but we can't just say that we have to

3291
02:07:50,870 --> 02:07:48,880
start thinking about and developing

3292
02:07:52,709 --> 02:07:50,880
instruments to do that

3293
02:07:55,750 --> 02:07:52,719
because i think that is a community

3294
02:07:57,270 --> 02:07:55,760
responsibility if we decide that this is

3295
02:07:59,589 --> 02:07:57,280
something we should be doing we also

3296
02:08:05,350 --> 02:07:59,599
have to help develop and provide the

3297
02:08:08,790 --> 02:08:07,350
oh sorry

3298
02:08:11,589 --> 02:08:08,800
so i

3299
02:08:15,030 --> 02:08:11,599
the instrument development

3300
02:08:17,189 --> 02:08:15,040
framework is can be confusing uh there

3301
02:08:19,990 --> 02:08:17,199
are all of these trl's which is

3302
02:08:22,069 --> 02:08:20,000
technology readiness level and if you

3303
02:08:24,310 --> 02:08:22,079
just read the definition and you start

3304
02:08:26,790 --> 02:08:24,320
trying to apply it to your technology

3305
02:08:28,709 --> 02:08:26,800
there comes some sort of ambiguity and

3306
02:08:30,550 --> 02:08:28,719
there's weird justifications that you

3307
02:08:33,430 --> 02:08:30,560
see for different instruments being at

3308
02:08:37,030 --> 02:08:33,440
different technology readiness levels

3309
02:08:39,750 --> 02:08:37,040
and this is also another reason if you

3310
02:08:40,870 --> 02:08:39,760
are tired of seeing a lot of the same

3311
02:08:43,669 --> 02:08:40,880
instruments

3312
02:08:45,669 --> 02:08:43,679
uh go on to flight

3313
02:08:47,750 --> 02:08:45,679

it's because there are very few

3314
02:08:50,950 --> 02:08:47,760
instruments that have made it through

3315
02:08:51,750 --> 02:08:50,960
this technology readiness ladder if you

3316
02:08:53,350 --> 02:08:51,760
will

3317
02:08:55,430 --> 02:08:53,360
um

3318
02:08:58,229 --> 02:08:55,440
and just because it works great here on

3319
02:09:00,790 --> 02:08:58,239
earth we know how to do it does not

3320
02:09:03,030 --> 02:09:00,800
translate directly into having an

3321
02:09:06,709 --> 02:09:03,040
instrument that can do it for flight

3322
02:09:09,669 --> 02:09:06,719
so your technology readiness level one

3323
02:09:12,550 --> 02:09:09,679
is you know the basic principles and

3324
02:09:15,510 --> 02:09:12,560
then two is the concept and application

3325
02:09:17,030 --> 02:09:15,520
formulated uh three is when you really

3326
02:09:19,109 --> 02:09:17,040
start getting your your critical

3327
02:09:21,189 --> 02:09:19,119
components down but when you hit four

3328
02:09:23,030 --> 02:09:21,199
and five that's when you start doing

3329
02:09:24,870 --> 02:09:23,040
breadboard validation you do an

3330
02:09:26,149 --> 02:09:24,880
integration of all of your different

3331
02:09:29,750 --> 02:09:26,159
components

3332
02:09:31,109 --> 02:09:29,760
uh six your subsystems in a relevant

3333
02:09:33,030 --> 02:09:31,119
environment

3334
02:09:34,790 --> 02:09:33,040
and i should back up a little bit and

3335
02:09:37,430 --> 02:09:34,800
say that

3336
02:09:40,310 --> 02:09:37,440
each component you are only at the trl

3337
02:09:43,189 --> 02:09:40,320
level of your weakest component so each

3338
02:09:45,750 --> 02:09:43,199

component within your flight instrument

3339
02:09:47,750 --> 02:09:45,760
has to have made it to that trl and not

3340
02:09:50,310 --> 02:09:47,760
just each component but in

3341
02:09:51,270 --> 02:09:50,320
in the specific integrated

3342
02:09:53,510 --> 02:09:51,280
um

3343
02:09:56,629 --> 02:09:53,520
schematic that you are proposing for

3344
02:10:01,270 --> 02:09:56,639
flight so as you can see this gets very

3345
02:10:04,149 --> 02:10:01,280
engineering this gets very um regulated

3346
02:10:06,629 --> 02:10:04,159
and it's not necessarily as biologists

3347
02:10:07,669 --> 02:10:06,639
or virologists the ways we're used to

3348
02:10:10,069 --> 02:10:07,679
thinking

3349
02:10:12,229 --> 02:10:10,079
so if you do go off to write these

3350
02:10:13,990 --> 02:10:12,239
instrument development proposals i would

3351
02:10:15,109 --> 02:10:14,000
definitely recommend teaming with an

3352
02:10:16,790 --> 02:10:15,119
engineer

3353
02:10:18,709 --> 02:10:16,800
because they start asking questions

3354
02:10:20,310 --> 02:10:18,719
especially if you get higher up in the

3355
02:10:22,790 --> 02:10:20,320
in in the trl's you know you need to

3356
02:10:24,390 --> 02:10:22,800
list every component down to the screws

3357
02:10:26,470 --> 02:10:24,400
how they've been tested what are they

3358
02:10:28,790 --> 02:10:26,480
made of what are your power requirements

3359
02:10:30,390 --> 02:10:28,800
your weight requirements how much does

3360
02:10:32,629 --> 02:10:30,400
everything weigh

3361
02:10:35,589 --> 02:10:32,639
what materials have they been proven in

3362
02:10:38,390 --> 02:10:35,599
flight so it gets rather complicated so

3363
02:10:40,470 --> 02:10:38,400

just because we can do it here on earth

3364
02:10:42,709 --> 02:10:40,480
doesn't necessarily translate into an

3365
02:10:44,709 --> 02:10:42,719
easy way to do it in space or on a

3366
02:10:46,310 --> 02:10:44,719
different planetary body not to say that

3367
02:10:47,430 --> 02:10:46,320
it can't be done it just needs to go

3368
02:10:49,750 --> 02:10:47,440
through through this sort of

3369
02:10:51,430 --> 02:10:49,760
hierarchical system i and there are

3370
02:10:53,830 --> 02:10:51,440
several different proposals for doing

3371
02:10:56,550 --> 02:10:53,840
this there's the niacc which if you have

3372
02:10:59,109 --> 02:10:56,560
a crazy concept and you want to sort of

3373
02:11:02,390 --> 02:10:59,119
explore if it's feasible that takes you

3374
02:11:04,629 --> 02:11:02,400
from material one to two i'm very new at

3375
02:11:06,229 --> 02:11:04,639
this proposal penny is the queen of

3376
02:11:08,229 --> 02:11:06,239
nyacks

3377
02:11:10,870 --> 02:11:08,239
but then there's also the picassos and

3378
02:11:13,510 --> 02:11:10,880
matisses picasso will get you from a one

3379
02:11:16,470 --> 02:11:13,520
to a three somewhere in there and then

3380
02:11:18,870 --> 02:11:16,480
four and above is a matisse uh so there

3381
02:11:20,550 --> 02:11:18,880
i i'm sure there's other proposals these

3382
02:11:23,189 --> 02:11:20,560
are just the ones that i personally am

3383
02:11:25,030 --> 02:11:23,199
familiar with um

3384
02:11:26,870 --> 02:11:25,040
so now that we've kind of talked about

3385
02:11:29,910 --> 02:11:26,880
instrument development

3386
02:11:31,350 --> 02:11:29,920
uh get on to so this is what i currently

3387
02:11:34,870 --> 02:11:31,360
am working with i'm working with

3388
02:11:35,990 --> 02:11:34,880

nanopore technology and i specifically

3389
02:11:38,310 --> 02:11:36,000
am looking for

3390
02:11:41,350 --> 02:11:38,320
long chain charged polymers not

3391
02:11:42,229 --> 02:11:41,360
necessarily dna and rna but anything

3392
02:11:46,550 --> 02:11:42,239
that

3393
02:11:49,750 --> 02:11:46,560
functional molecule uh nanopore

3394
02:11:51,910 --> 02:11:49,760
technology the concept is pretty simple

3395
02:11:53,990 --> 02:11:51,920
i it can you know there's more much more

3396
02:11:56,629 --> 02:11:54,000
layers of complication that be added on

3397
02:11:58,470 --> 02:11:56,639
as you progress but the basic concept is

3398
02:12:01,030 --> 02:11:58,480
you have a nanopore that spans a

3399
02:12:03,270 --> 02:12:01,040
membrane electrolyte solution is on

3400
02:12:06,069 --> 02:12:03,280
either side of that membrane a voltage

3401
02:12:07,510 --> 02:12:06,079
is applied an ion current established

3402
02:12:09,990 --> 02:12:07,520
through the pore

3403
02:12:12,310 --> 02:12:10,000
as a particle translocates through that

3404
02:12:14,870 --> 02:12:12,320
pore you get a blockage for a short

3405
02:12:17,589 --> 02:12:14,880
period of time and that current drops so

3406
02:12:19,669 --> 02:12:17,599
that little current dropping right there

3407
02:12:22,870 --> 02:12:19,679
in the schematic you can see you have

3408
02:12:25,430 --> 02:12:22,880
your depth and your duration so your

3409
02:12:26,870 --> 02:12:25,440
depth is actually telling you the radius

3410
02:12:28,149 --> 02:12:26,880
of the molecule

3411
02:12:30,629 --> 02:12:28,159
going through

3412
02:12:33,510 --> 02:12:30,639
based on the size of the pore as it

3413
02:12:35,589 --> 02:12:33,520

translocates if it's closer to the size

3414
02:12:37,910 --> 02:12:35,599
of the pore it's blocking more current

3415
02:12:41,030 --> 02:12:37,920
and you're going to have a a larger

3416
02:12:43,350 --> 02:12:41,040
depth where your translocation time is

3417
02:12:45,669 --> 02:12:43,360
telling you like the length of your

3418
02:12:47,030 --> 02:12:45,679
particle going through how long it takes

3419
02:12:50,870 --> 02:12:47,040
to go through

3420
02:12:52,310 --> 02:12:50,880
um so the concept there is fairly simple

3421
02:12:54,629 --> 02:12:52,320
and

3422
02:12:56,629 --> 02:12:54,639
right now there are two different there

3423
02:12:59,030 --> 02:12:56,639
are two different um ways of doing this

3424
02:13:00,310 --> 02:12:59,040
nanopore technology of the biological

3425
02:13:02,229 --> 02:13:00,320
nanopore

3426
02:13:04,310 --> 02:13:02,239
that's commercially available it's

3427
02:13:06,950 --> 02:13:04,320
operated uh it's through oxford nano

3428
02:13:09,990 --> 02:13:06,960
ports operated on the iss which was

3429
02:13:13,270 --> 02:13:10,000
really cool uh so they did sequencing in

3430
02:13:14,950 --> 02:13:13,280
situ on the iss and it worked great uh

3431
02:13:18,069 --> 02:13:14,960
actually with the

3432
02:13:21,109 --> 02:13:18,079
way the flow cells work the the um

3433
02:13:22,390 --> 02:13:21,119
min ion that's the biological one uh

3434
02:13:24,550 --> 02:13:22,400
actually worked better in the space

3435
02:13:27,270 --> 02:13:24,560
environment than here on earth and not

3436
02:13:30,069 --> 02:13:27,280
quite sure why as to yet but for the

3437
02:13:33,270 --> 02:13:30,079
biological one just to explain there is

3438
02:13:36,550 --> 02:13:33,280

actually a biological protein that

3439
02:13:39,030 --> 02:13:36,560
creates that nanopore in the membrane

3440
02:13:40,870 --> 02:13:39,040
and it's very specific to translocating

3441
02:13:43,510 --> 02:13:40,880
dna and rna

3442
02:13:46,550 --> 02:13:43,520
through through cell membranes so it is

3443
02:13:48,470 --> 02:13:46,560
very specific to dna and rna and again

3444
02:13:50,310 --> 02:13:48,480
it works great on the iss but when we

3445
02:13:52,470 --> 02:13:50,320
start thinking about longer duration

3446
02:13:54,069 --> 02:13:52,480
missions i again want to remember one of

3447
02:13:55,669 --> 02:13:54,079
the things that i talked about for

3448
02:13:57,910 --> 02:13:55,679
instruments is you have to start looking

3449
02:14:00,149 --> 02:13:57,920
at the material that you're made out of

3450
02:14:02,470 --> 02:14:00,159
and a biological membrane is not going

3451
02:14:04,229 --> 02:14:02,480
to be robust enough for space flight if

3452
02:14:06,149 --> 02:14:04,239
you think about you know your

3453
02:14:07,910 --> 02:14:06,159
temperature swings your radiation

3454
02:14:09,910 --> 02:14:07,920
environment and just over time you're

3455
02:14:11,830 --> 02:14:09,920
going to get degradation of that protein

3456
02:14:13,350 --> 02:14:11,840
so it's not robust enough for a long

3457
02:14:16,470 --> 02:14:13,360
duration mission

3458
02:14:18,870 --> 02:14:16,480
uh also versatility if we go back to

3459
02:14:21,830 --> 02:14:18,880
thinking about life might not look like

3460
02:14:24,470 --> 02:14:21,840
what it looks like here uh the

3461
02:14:25,669 --> 02:14:24,480
biological port is very specific for dna

3462
02:14:27,990 --> 02:14:25,679
and rna so you don't have the

3463
02:14:31,430 --> 02:14:28,000

versatility to look for other molecules

3464
02:14:33,990 --> 02:14:31,440
of interest whereas with the solid state

3465
02:14:35,910 --> 02:14:34,000
you can change that pore size you can

3466
02:14:37,750 --> 02:14:35,920
have it smaller larger you can change

3467
02:14:38,870 --> 02:14:37,760
the shape of the pore you can coat the

3468
02:14:40,629 --> 02:14:38,880
pore

3469
02:14:44,149 --> 02:14:40,639
you can reverse polarities make things

3470
02:14:46,550 --> 02:14:44,159
go backwards uh so here in the picture

3471
02:14:47,910 --> 02:14:46,560
is um

3472
02:14:49,910 --> 02:14:47,920
if you start at the very top there are

3473
02:14:51,270 --> 02:14:49,920
two little electrodes going into the

3474
02:14:55,350 --> 02:14:51,280
reservoirs

3475
02:14:58,550 --> 02:14:55,360
golden things as you can see this is a

3476
02:15:00,470 --> 02:14:58,560
very small very lightweight device um

3477
02:15:02,229 --> 02:15:00,480
it's made of silicon nitride which has

3478
02:15:04,629 --> 02:15:02,239
flight heritage we fly things made out

3479
02:15:06,950 --> 02:15:04,639
of silicon nitride all the time so we

3480
02:15:09,990 --> 02:15:06,960
know the material works it's a small

3481
02:15:12,310 --> 02:15:10,000
weight a small power requirement

3482
02:15:15,350 --> 02:15:12,320
and in between those reservoirs you can

3483
02:15:17,990 --> 02:15:15,360
see in the image on the bottom left is a

3484
02:15:20,470 --> 02:15:18,000
little micro fluidic channel

3485
02:15:23,109 --> 02:15:20,480
in that channel there's a reservoir with

3486
02:15:25,109 --> 02:15:23,119
a tiny little hole so

3487
02:15:27,990 --> 02:15:25,119
this is just a this is just a single

3488
02:15:30,149 --> 02:15:28,000

hole chip we you can do multiple arrays

3489
02:15:31,589 --> 02:15:30,159
there are various different schematics

3490
02:15:34,149 --> 02:15:31,599
um

3491
02:15:37,350 --> 02:15:34,159
and you would introduce your sample into

3492
02:15:39,990 --> 02:15:37,360
that square reservoir which is on top of

3493
02:15:40,790 --> 02:15:40,000
this gold reservoir sitting on top of

3494
02:15:44,149 --> 02:15:40,800
that

3495
02:15:47,350 --> 02:15:44,159
and you would wait and watch the signal

3496
02:15:50,709 --> 02:15:47,360
as your in bioma biological molecule of

3497
02:15:52,390 --> 02:15:50,719
interest translocates

3498
02:15:55,830 --> 02:15:52,400
uh so

3499
02:15:57,990 --> 02:15:55,840
uh again i don't work with viruses but i

3500
02:16:00,310 --> 02:15:58,000
thought this would be a good way to

3501
02:16:03,669 --> 02:16:00,320
introduce how this

3502
02:16:06,470 --> 02:16:03,679
technology is currently being used to

3503
02:16:10,149 --> 02:16:06,480
look at viruses so here's a paper from

3504
02:16:13,270 --> 02:16:10,159
2014 it looks at the stiff filamentous

3505
02:16:14,310 --> 02:16:13,280
virus fd you can see a translocation

3506
02:16:19,270 --> 02:16:14,320
there

3507
02:16:21,430 --> 02:16:19,280
these the in the folded if the virus is

3508
02:16:23,430 --> 02:16:21,440
folded it's too large to fit through

3509
02:16:25,669 --> 02:16:23,440
that nanopore so you have a lot of

3510
02:16:27,350 --> 02:16:25,679
collisions and as you can imagine those

3511
02:16:28,470 --> 02:16:27,360
collisions as the

3512
02:16:31,270 --> 02:16:28,480
viral

3513
02:16:33,030 --> 02:16:31,280

uh particles bump up against that hole

3514
02:16:36,230 --> 02:16:33,040
they're blocking the current for very

3515
02:16:38,790 --> 02:16:36,240
intermittent uh amounts of time so you

3516
02:16:41,030 --> 02:16:38,800
look there on the graph and you can see

3517
02:16:44,150 --> 02:16:41,040
those collisions happening right before

3518
02:16:45,190 --> 02:16:44,160
the translocation so the

3519
02:16:50,629 --> 02:16:45,200
viral

3520
02:16:53,349 --> 02:16:50,639
nanopore multiple times until it finally

3521
02:16:55,270 --> 02:16:53,359
gets the right geometry to translocate

3522
02:16:58,070 --> 02:16:55,280
and then if you look at sort of that

3523
02:16:59,750 --> 02:16:58,080
heat map it's just showing you the

3524
02:17:02,230 --> 02:16:59,760
um

3525
02:17:04,309 --> 02:17:02,240
the the grouping where you can see the

3526
02:17:07,030 --> 02:17:04,319
collisions fall and a much broader

3527
02:17:10,870 --> 02:17:07,040
spectrum but the translocations uh

3528
02:17:18,389 --> 02:17:15,589
so in 2016 we did a study looking at the

3529
02:17:22,309 --> 02:17:18,399
tobacco mosaic virus and found that

3530
02:17:25,750 --> 02:17:22,319
there is actually a very clear pattern

3531
02:17:26,950 --> 02:17:25,760
as the tran as the tobacco mosaic virus

3532
02:17:29,190 --> 02:17:26,960
goes through

3533
02:17:31,110 --> 02:17:29,200
the nanopore because it's very rigid it

3534
02:17:34,070 --> 02:17:31,120
doesn't have that flexibility that we

3535
02:17:36,629 --> 02:17:34,080
saw previously so there are these three

3536
02:17:40,070 --> 02:17:36,639
steps in the translocation pattern that

3537
02:17:43,190 --> 02:17:40,080
are very very indicative of this rod

3538
02:17:46,070 --> 02:17:43,200

rigid rod shaped structure and you have

3539
02:17:48,070 --> 02:17:46,080
the first part in blue there where

3540
02:17:50,389 --> 02:17:48,080
the virus is trying to enter the

3541
02:17:52,150 --> 02:17:50,399
nanopore wiggling around a bit

3542
02:17:54,070 --> 02:17:52,160
red where it's finally trans it's

3543
02:17:56,469 --> 02:17:54,080
starting to translocate and then green

3544
02:17:58,950 --> 02:17:56,479
where you get the actual translocation

3545
02:18:01,669 --> 02:17:58,960
so as you can imagine uh you could look

3546
02:18:03,669 --> 02:18:01,679
for any virus because you can change the

3547
02:18:06,230 --> 02:18:03,679
pore size right you could make it

3548
02:18:08,309 --> 02:18:06,240
smaller or larger depending on what

3549
02:18:09,270 --> 02:18:08,319
virus you're looking for and once you

3550
02:18:10,709 --> 02:18:09,280
know the

3551
02:18:12,629 --> 02:18:10,719
the um

3552
02:18:15,830 --> 02:18:12,639
the indicative pattern

3553
02:18:17,990 --> 02:18:15,840
you might not be able to say for sure

3554
02:18:19,669 --> 02:18:18,000
that this is tobacco mosaic virus but

3555
02:18:22,389 --> 02:18:19,679
you would be able to say that this is a

3556
02:18:23,429 --> 02:18:22,399
rod shaped virus i and you would know

3557
02:18:26,150 --> 02:18:23,439
you could

3558
02:18:29,990 --> 02:18:26,160
differ from that the radius and the

3559
02:18:30,000 --> 02:18:35,429
uh and then whoa in 2011

3560
02:18:44,309 --> 02:18:40,150
so again these are looking at um

3561
02:18:49,030 --> 02:18:44,319
the hpv capsids uh t3 and t4 as you can

3562
02:18:52,230 --> 02:18:49,040
see their uh radius is 31 to and 36

3563
02:18:54,469 --> 02:18:52,240

nanometers right so they are they're a

3564
02:18:57,270 --> 02:18:54,479
different size and you see that in the

3565
02:18:59,589 --> 02:18:57,280
translocation data up there on the top

3566
02:19:01,830 --> 02:18:59,599
uh you have your baseline signal and

3567
02:19:03,830 --> 02:19:01,840
then your t3 which has your smaller

3568
02:19:06,389 --> 02:19:03,840
radius you see the translocations

3569
02:19:08,709 --> 02:19:06,399
happening there and again remember the

3570
02:19:10,549 --> 02:19:08,719
depth of your translocation signal is

3571
02:19:12,070 --> 02:19:10,559
how much current is being blocked so if

3572
02:19:15,509 --> 02:19:12,080
it's smaller you're not going to be

3573
02:19:17,830 --> 02:19:15,519
blocking as much current uh and then

3574
02:19:20,230 --> 02:19:17,840
your t4 which is larger you see that

3575
02:19:21,750 --> 02:19:20,240
jump and current blockade there and

3576
02:19:23,589 --> 02:19:21,760
that's because you're blocking more of

3577
02:19:26,870 --> 02:19:23,599
the current as it translocates through

3578
02:19:30,629 --> 02:19:29,349
and then uh last one on this i'm just

3579
02:19:32,870 --> 02:19:30,639
hoping to

3580
02:19:34,549 --> 02:19:32,880
show you the ability and the diversity

3581
02:19:37,190 --> 02:19:34,559
of the system sort of some of the things

3582
02:19:39,270 --> 02:19:37,200
that you can do for it do with it uh

3583
02:19:43,270 --> 02:19:39,280
there also is a way to get at the

3584
02:19:45,830 --> 02:19:43,280
density so you can measure the mass of

3585
02:19:49,830 --> 02:19:45,840
the nanoparticles or viruses and their

3586
02:19:52,950 --> 02:19:49,840
sedimentation so looking at the time it

3587
02:19:55,030 --> 02:19:52,960
takes for them you know their their

3588
02:19:56,870 --> 02:19:55,040

radius and their length and then you

3589
02:19:59,670 --> 02:19:56,880
know the time at which it takes them to

3590
02:20:01,270 --> 02:19:59,680
settle um

3591
02:20:04,389 --> 02:20:01,280
this might be different for spaceflight

3592
02:20:06,230 --> 02:20:04,399
but this works great here on earth

3593
02:20:07,830 --> 02:20:06,240
but one of the other things that is is

3594
02:20:10,070 --> 02:20:07,840
good that this paper points out is you

3595
02:20:12,950 --> 02:20:10,080
are working very very small small

3596
02:20:15,110 --> 02:20:12,960
volumes and small concentrations i and

3597
02:20:16,630 --> 02:20:15,120
this is very important to have

3598
02:20:18,710 --> 02:20:16,640
technology that can work with small

3599
02:20:21,030 --> 02:20:18,720
volumes because if you're thinking about

3600
02:20:24,550 --> 02:20:21,040
going through say the enceladus plume

3601
02:20:27,830 --> 02:20:24,560
and capturing uh some of the samples so

3602
02:20:30,870 --> 02:20:27,840
enceladus is uh icy moon that is spewing

3603
02:20:33,190 --> 02:20:30,880
sort of like a geyser particles of water

3604
02:20:35,030 --> 02:20:33,200
out into space and if you wanted to fly

3605
02:20:37,030 --> 02:20:35,040
through that and collect that you might

3606
02:20:39,590 --> 02:20:37,040
be getting a milliliter i think is right

3607
02:20:42,150 --> 02:20:39,600
now the projected volume that they're

3608
02:20:44,309 --> 02:20:42,160
thinking of collecting so

3609
02:20:45,750 --> 02:20:44,319
again you are looking for a needle in a

3610
02:20:47,830 --> 02:20:45,760
haystack so you have to have something

3611
02:20:50,790 --> 02:20:47,840
that's very sensitive and can work with

3612
02:20:53,510 --> 02:20:50,800
these low concentrations uh and and

3613
02:20:56,870 --> 02:20:53,520

small volumes

3614
02:20:59,429 --> 02:20:56,880
one more thing before i run out of time

3615
02:21:01,910 --> 02:20:59,439
um alexander couldn't be here but i did

3616
02:21:04,710 --> 02:21:01,920
want to point to a paper that was

3617
02:21:06,710 --> 02:21:04,720
published in 2018 the need for include

3618
02:21:08,870 --> 02:21:06,720
including virus detection methods in

3619
02:21:11,830 --> 02:21:08,880
future mars missions

3620
02:21:13,510 --> 02:21:11,840
and this goes over sort of the argument

3621
02:21:15,429 --> 02:21:13,520
or um

3622
02:21:18,870 --> 02:21:15,439
reasons why to look for

3623
02:21:21,190 --> 02:21:18,880
for viruses on mars and it's that we

3624
02:21:23,990 --> 02:21:21,200
don't really know how and when uh viral

3625
02:21:26,630 --> 02:21:24,000
units were created here on earth

3626
02:21:30,150 --> 02:21:26,640
the virus first hypothesis is very

3627
02:21:32,389 --> 02:21:30,160
controversial so exploration of mars you

3628
02:21:34,790 --> 02:21:32,399
could kind of finally close this gap if

3629
02:21:35,830 --> 02:21:34,800
viral units were actually arose before

3630
02:21:37,510 --> 02:21:35,840
cells

3631
02:21:39,750 --> 02:21:37,520
um

3632
02:21:43,510 --> 02:21:39,760
we can find viral if we find viable

3633
02:21:45,110 --> 02:21:43,520
units on mars but no cells then that can

3634
02:21:46,950 --> 02:21:45,120
only be explained by the virus first

3635
02:21:48,469 --> 02:21:46,960
hypothesis

3636
02:21:51,270 --> 02:21:48,479
and um

3637
02:21:53,750 --> 02:21:51,280
you know if they don't exist on mars

3638
02:21:55,590 --> 02:21:53,760

uh falsifies the viral first hypothesis

3639
02:21:58,309 --> 02:21:55,600
again this isn't my work i just wanted

3640
02:21:59,910 --> 02:21:58,319
to point put this up here and uh point

3641
02:22:02,469 --> 02:21:59,920
to the paper because unfortunately he

3642
02:22:03,830 --> 02:22:02,479
couldn't be here with us today

3643
02:22:05,590 --> 02:22:03,840
and

3644
02:22:08,630 --> 02:22:05,600
with that i just want to say build it

3645
02:22:11,110 --> 02:22:08,640
and they will come um you know there are

3646
02:22:12,070 --> 02:22:11,120
so few instruments that are flight ready

3647
02:22:15,670 --> 02:22:12,080
so

3648
02:22:17,110 --> 02:22:15,680
get get your instruments out there and

3649
02:22:20,070 --> 02:22:17,120
work on them as much as possible and

3650
02:22:22,070 --> 02:22:20,080
often and you'll have a very good chance

3651
02:22:24,270 --> 02:22:22,080
um for spaceflight and that's what i

3652
02:22:31,349 --> 02:22:24,280
have

3653
02:22:41,110 --> 02:22:32,870
i guess we'll open up for any questions

3654
02:22:45,429 --> 02:22:43,830
so i'll start it off um catherine so

3655
02:22:46,389 --> 02:22:45,439
as we've heard from a number of people

3656
02:22:47,510 --> 02:22:46,399
here particularly evelyn right at the

3657
02:22:50,550 --> 02:22:47,520
beginning

3658
02:22:51,510 --> 02:22:50,560
many bacteriophage have these sort of

3659
02:22:53,349 --> 02:22:51,520
mixed

3660
02:22:55,990 --> 02:22:53,359
structures and i'll show some of the

3661
02:22:59,030 --> 02:22:56,000
other crazy archaeal virus structures in

3662
02:23:01,670 --> 02:22:59,040
just a couple of minutes um

3663
02:23:03,590 --> 02:23:01,680

how easy is it for these nanopores to

3664
02:23:05,270 --> 02:23:03,600
sort of look at a head and tail phage

3665
02:23:06,790 --> 02:23:05,280
for instance or something like that it

3666
02:23:09,190 --> 02:23:06,800
seems that that might be

3667
02:23:10,389 --> 02:23:09,200
quite a challenge

3668
02:23:12,870 --> 02:23:10,399
you know

3669
02:23:15,030 --> 02:23:12,880
i actually think the more crazy the

3670
02:23:17,830 --> 02:23:15,040
structure the better it will be at

3671
02:23:19,830 --> 02:23:17,840
detecting it because as it passes

3672
02:23:23,270 --> 02:23:19,840
through the nanopore you're going to get

3673
02:23:26,630 --> 02:23:23,280
a very indicative signal so just as we

3674
02:23:28,870 --> 02:23:26,640
saw the tobacco mosaic virus translocate

3675
02:23:31,270 --> 02:23:28,880
and you had that three stepwise

3676
02:23:33,750 --> 02:23:31,280
translocation phase that happened almost

3677
02:23:36,630 --> 02:23:33,760
every single time it translocates

3678
02:23:39,349 --> 02:23:36,640
the crazier the structure the more

3679
02:23:41,990 --> 02:23:39,359
specific the signature so

3680
02:23:50,469 --> 02:23:42,000
it actually i think would be a bonus

3681
02:23:53,910 --> 02:23:51,750
yes

3682
02:23:55,510 --> 02:23:53,920
if someone happened to have a crazy

3683
02:23:57,670 --> 02:23:55,520
virus and wanted to put it through a

3684
02:23:59,429 --> 02:23:57,680
nanopore who would one talk to

3685
02:24:01,830 --> 02:23:59,439
[Laughter]

3686
02:24:03,510 --> 02:24:01,840
i would say send me an email because i

3687
02:24:06,389 --> 02:24:03,520
love putting things through the nano

3688
02:24:13,349 --> 02:24:09,510

i will take your samples

3689
02:24:13,359 --> 02:24:17,429
we got some crazy ones

3690
02:24:22,790 --> 02:24:20,309
i also just wanted to point out there uh

3691
02:24:24,710 --> 02:24:22,800
the benefits we get from the long read

3692
02:24:26,150 --> 02:24:24,720
sequencing not only are we getting the

3693
02:24:28,870 --> 02:24:26,160
in-situ

3694
02:24:31,990 --> 02:24:28,880
sequencing done but also potentially a

3695
02:24:33,510 --> 02:24:32,000
whole virus which is phenomenal

3696
02:24:37,510 --> 02:24:33,520
yes the

3697
02:24:39,190 --> 02:24:37,520
the uh earth based biological pores by

3698
02:24:42,870 --> 02:24:39,200
oxford nanopore

3699
02:24:44,790 --> 02:24:42,880
minion and the flongol flangoll

3700
02:24:46,790 --> 02:24:44,800
don't know their naming um

3701
02:24:48,710 --> 02:24:46,800
but uh they're

3702
02:24:51,270 --> 02:24:48,720
they're great i've actually taken it uh

3703
02:24:53,670 --> 02:24:51,280
to the atacama and did in situ dna and

3704
02:24:57,910 --> 02:24:53,680
rna sequencing there i took it to the

3705
02:25:00,710 --> 02:24:57,920
arctic it functions great um it in situ

3706
02:25:02,550 --> 02:25:00,720
to kind of look at what you have kind of

3707
02:25:05,190 --> 02:25:02,560
on the spot right there

3708
02:25:07,429 --> 02:25:05,200
um you know and it's commercial off the

3709
02:25:09,750 --> 02:25:07,439
shelf available and it's

3710
02:25:11,830 --> 02:25:09,760
it's kind of revolutionizing the way

3711
02:25:13,510 --> 02:25:11,840
we're able to do biology and the way

3712
02:25:16,630 --> 02:25:13,520
that it's open to multiple different

3713
02:25:18,870 --> 02:25:16,640

people because the protocols for using

3714
02:25:21,110 --> 02:25:18,880
it are so very simple you just need to

3715
02:25:23,110 --> 02:25:21,120
extract your dna and then they even will

3716
02:25:25,670 --> 02:25:23,120
do a library prep for you on an

3717
02:25:27,990 --> 02:25:25,680
automated voltrex and then the minion

3718
02:25:31,429 --> 02:25:28,000
you just need to know how to pipette uh

3719
02:25:36,309 --> 02:25:31,439
onto it and it takes care of the rest so

3720
02:25:43,429 --> 02:25:38,550
have someone pipette for you you're good

3721
02:25:47,110 --> 02:25:45,670
all right it's not been done in case

3722
02:25:49,190 --> 02:25:47,120
it is not

3723
02:25:51,190 --> 02:25:49,200
next place it needs to be done i got a

3724
02:25:52,710 --> 02:25:51,200
list

3725
02:25:54,870 --> 02:25:52,720
i just wanted to bring up some of the

3726
02:25:57,190 --> 02:25:54,880
discussions going on in the chat is that

3727
02:25:59,910 --> 02:25:57,200
with the long read there are some error

3728
02:26:00,870 --> 02:25:59,920
rates yes that's very true so i don't i

3729
02:26:02,630 --> 02:26:00,880
don't know if you want to say anything

3730
02:26:04,070 --> 02:26:02,640
about that i have something else yet but

3731
02:26:06,950 --> 02:26:04,080
go ahead um

3732
02:26:09,590 --> 02:26:06,960
you know i would say that

3733
02:26:12,870 --> 02:26:09,600
you know using dual methods

3734
02:26:15,110 --> 02:26:12,880
uh to try to overcome the error rates

3735
02:26:16,710 --> 02:26:15,120
is the most successful strategy that

3736
02:26:17,750 --> 02:26:16,720
i've seen so far

3737
02:26:19,990 --> 02:26:17,760
so

3738
02:26:22,150 --> 02:26:20,000

the minion is great for taking it in the

3739
02:26:24,389 --> 02:26:22,160
field and doing stuff in situ and for

3740
02:26:26,389 --> 02:26:24,399
getting long reads but then

3741
02:26:28,469 --> 02:26:26,399
it's a good practice to always collect

3742
02:26:30,309 --> 02:26:28,479
and process samples using traditional

3743
02:26:34,630 --> 02:26:30,319
techniques so that you can kind of get

3744
02:26:38,710 --> 02:26:36,950
that's harder on the spacecraft right

3745
02:26:41,590 --> 02:26:38,720
but first for space

3746
02:26:44,150 --> 02:26:41,600
for testing and for space flights

3747
02:26:45,830 --> 02:26:44,160
you know i

3748
02:26:47,270 --> 02:26:45,840
depends on what you want to do but most

3749
02:26:49,030 --> 02:26:47,280
cases you don't need that level of

3750
02:26:50,469 --> 02:26:49,040
resolution

3751
02:26:53,110 --> 02:26:50,479
but you can definitely imagine the

3752
02:26:55,830 --> 02:26:53,120
promise of it because like what our vend

3753
02:26:57,750 --> 02:26:55,840
was talking about segmented genomes we

3754
02:26:59,830 --> 02:26:57,760
have so many reads that we also throw

3755
02:27:00,950 --> 02:26:59,840
assembling viral convicts can be

3756
02:27:03,190 --> 02:27:00,960
difficult

3757
02:27:05,429 --> 02:27:03,200
so if you had a short read version and

3758
02:27:07,670 --> 02:27:05,439
you had the long read to kind of be the

3759
02:27:09,590 --> 02:27:07,680
reference genome and then just fix some

3760
02:27:12,150 --> 02:27:09,600
error rates that would be beneficial

3761
02:27:14,950 --> 02:27:12,160
well even for looking at stuff on iss or

3762
02:27:16,630 --> 02:27:14,960
the lunar surface if you wanted to see

3763
02:27:18,070 --> 02:27:16,640

how things are responding to their

3764
02:27:20,950 --> 02:27:18,080
environment and you don't want to

3765
02:27:23,590 --> 02:27:20,960
necessarily think send everything back

3766
02:27:26,070 --> 02:27:23,600
to earth you can switch

3767
02:27:28,550 --> 02:27:26,080
sequence in situ and yes you might have

3768
02:27:31,429 --> 02:27:28,560
these error rates but you'll get

3769
02:27:32,950 --> 02:27:31,439
data in situ which is something that

3770
02:27:34,870 --> 02:27:32,960
we have you know we haven't really done

3771
02:27:37,110 --> 02:27:34,880
before yeah

3772
02:27:38,630 --> 02:27:37,120
there's another question here

3773
02:27:40,469 --> 02:27:38,640
um

3774
02:27:42,550 --> 02:27:40,479
felipe

3775
02:27:44,870 --> 02:27:42,560
i wonder about the minimum quantity of

3776
02:27:47,110 --> 02:27:44,880
the target sample to be able to detect

3777
02:27:49,750 --> 02:27:47,120
the molecules

3778
02:27:51,750 --> 02:27:49,760
in martian soil desert soil etc so your

3779
02:27:55,349 --> 02:27:51,760
level of detection

3780
02:27:59,590 --> 02:27:55,359
for the solid state nano port is really

3781
02:28:01,910 --> 02:27:59,600
based on your level of patience

3782
02:28:04,710 --> 02:28:01,920
in theory you could do single molecule

3783
02:28:06,070 --> 02:28:04,720
detection so you are if it's a charged

3784
02:28:08,309 --> 02:28:06,080
poly

3785
02:28:10,710 --> 02:28:08,319
charged particle and you are applying

3786
02:28:12,790 --> 02:28:10,720
that voltage it will eventually

3787
02:28:16,550 --> 02:28:12,800
find its way through the nanopore and

3788
02:28:19,030 --> 02:28:16,560

translocate in theory uh so

3789
02:28:22,469 --> 02:28:19,040
the level of detection

3790
02:28:25,190 --> 02:28:22,479
really just is based on time um

3791
02:28:26,790 --> 02:28:25,200
in lab with samples that are of course

3792
02:28:29,510 --> 02:28:26,800
uh you know we took some lake vita

3793
02:28:30,469 --> 02:28:29,520
samples and it's fairly sparse

3794
02:28:32,790 --> 02:28:30,479
after

3795
02:28:34,630 --> 02:28:32,800
30 minutes we were seeing translocations

3796
02:28:37,349 --> 02:28:34,640
we're not going to see everything unless

3797
02:28:42,230 --> 02:28:37,359
we've run it for several days

3798
02:28:44,230 --> 02:28:42,240
but you also have um you know it's

3799
02:28:46,070 --> 02:28:44,240
the smaller you're we're also working on

3800
02:28:48,230 --> 02:28:46,080
a front end uh

3801
02:28:51,030 --> 02:28:48,240
sample processor for

3802
02:28:52,469 --> 02:28:51,040
mars or an ocean world where it takes a

3803
02:28:53,750 --> 02:28:52,479
bulk sample

3804
02:28:56,950 --> 02:28:53,760
um

3805
02:28:58,469 --> 02:28:56,960
lyses extracts the polymers and then

3806
02:29:00,389 --> 02:28:58,479
goes through a concentration and

3807
02:29:02,469 --> 02:29:00,399
desalination phase because in a lot of

3808
02:29:05,190 --> 02:29:02,479
these environments as you concentrate

3809
02:29:07,429 --> 02:29:05,200
your salts increase and you want to be

3810
02:29:10,150 --> 02:29:07,439
working within an electrolyte buffer

3811
02:29:12,230 --> 02:29:10,160
range right so you have to work on

3812
02:29:13,750 --> 02:29:12,240
desalination and that's that's one of

3813
02:29:15,670 --> 02:29:13,760

the the questions that we're trying to

3814
02:29:17,590 --> 02:29:15,680
tackle is with this front-end sample

3815
02:29:20,389 --> 02:29:17,600
processing system where we're

3816
02:29:22,870 --> 02:29:20,399
concentrating um what what sort of

3817
02:29:24,630 --> 02:29:22,880
levels of detection can we get in to

3818
02:29:26,309 --> 02:29:24,640
with functioning within within a

3819
02:29:27,910 --> 02:29:26,319
reasonable limit of time

3820
02:29:29,990 --> 02:29:27,920
so

3821
02:29:32,070 --> 02:29:30,000
that is a currently funded ongoing

3822
02:29:33,830 --> 02:29:32,080
project

3823
02:29:36,630 --> 02:29:33,840
i also want to point out there for

3824
02:29:38,309 --> 02:29:36,640
people who maybe are intimidated by

3825
02:29:39,910 --> 02:29:38,319
bioinformatics

3826
02:29:41,670 --> 02:29:39,920
because short reads been around for so

3827
02:29:43,270 --> 02:29:41,680
long we have all these app based systems

3828
02:29:45,349 --> 02:29:43,280
to help you and stuff

3829
02:29:47,270 --> 02:29:45,359
now a lot of these apps don't work for

3830
02:29:48,630 --> 02:29:47,280
the long read so this is an ongoing

3831
02:29:50,950 --> 02:29:48,640
process and there's definitely going to

3832
02:29:53,590 --> 02:29:50,960
be a lot of

3833
02:29:55,830 --> 02:29:53,600
increases in efficiency and workflows

3834
02:29:58,309 --> 02:29:55,840
and stuff for bioinformatics i think

3835
02:30:00,150 --> 02:29:58,319
within the next one to two even

3836
02:30:02,469 --> 02:30:00,160
beyond just within the next year we'll

3837
02:30:04,070 --> 02:30:02,479
definitely see a lot more on this and we

3838
02:30:06,710 --> 02:30:04,080

talk about promise but again this is

3839
02:30:08,309 --> 02:30:06,720
something that is already being used at

3840
02:30:09,910 --> 02:30:08,319
large in the community it's not

3841
02:30:11,830 --> 02:30:09,920
something that might be working in the

3842
02:30:13,190 --> 02:30:11,840
future i mean it's used to track ebola

3843
02:30:15,750 --> 02:30:13,200
outbreaks in

3844
02:30:18,550 --> 02:30:15,760
africa it's used you know pretty much

3845
02:30:20,950 --> 02:30:18,560
all over there's a pilot program in at

3846
02:30:22,870 --> 02:30:20,960
the hospitals in san francisco

3847
02:30:24,630 --> 02:30:22,880
if you have come down with an illness

3848
02:30:26,309 --> 02:30:24,640
and they can't diagnose it they send it

3849
02:30:28,469 --> 02:30:26,319
away for culturing it comes back

3850
02:30:30,550 --> 02:30:28,479
negative you have to send it away again

3851
02:30:32,710 --> 02:30:30,560
it comes back i think like 95 percent of

3852
02:30:35,270 --> 02:30:32,720
those patients die if it's a it's a

3853
02:30:37,830 --> 02:30:35,280
serious disease so now if the culture

3854
02:30:39,429 --> 02:30:37,840
comes back negative the first go around

3855
02:30:42,150 --> 02:30:39,439
they go ahead and use the menon to try

3856
02:30:45,110 --> 02:30:42,160
and identify and get get information

3857
02:30:47,349 --> 02:30:45,120
right away uh yeah there's there's you

3858
02:30:49,030 --> 02:30:47,359
know i could go on and on about

3859
02:30:51,830 --> 02:30:49,040
you know crime scenes and their

3860
02:30:53,910 --> 02:30:51,840
application for blood spatter and

3861
02:30:58,070 --> 02:30:53,920
identifying they're they're working in

3862
02:31:04,870 --> 02:30:58,080
every aspect so it is forensics so it is

3863
02:31:09,349 --> 02:31:07,670

i think i'm out of time

3864
02:31:10,550 --> 02:31:09,359
yeah i think we'll go ahead and start

3865
02:31:18,790 --> 02:31:10,560
our break

3866
02:31:23,510 --> 02:31:20,950
i wanted to talk about using stable

3867
02:31:25,429 --> 02:31:23,520
isotopes generally to track

3868
02:31:27,270 --> 02:31:25,439
viruses and soils and i'm focusing on

3869
02:31:29,110 --> 02:31:27,280
soils because

3870
02:31:31,670 --> 02:31:29,120
particularly phage from soils are what i

3871
02:31:33,750 --> 02:31:31,680
care about most but this can be applied

3872
02:31:36,309 --> 02:31:33,760
to any environment and this is not the

3873
02:31:38,070 --> 02:31:36,319
only tool that can be used to track

3874
02:31:41,990 --> 02:31:38,080
viruses or look at the activity of

3875
02:31:46,469 --> 02:31:44,150
so starting off that soils are

3876
02:31:48,790 --> 02:31:46,479
complex ecosystems

3877
02:31:50,870 --> 02:31:48,800
in this cartoon here you can look at

3878
02:31:53,349 --> 02:31:50,880
about one centimeter squared area and

3879
02:31:54,309 --> 02:31:53,359
see that's teeming with organisms and

3880
02:31:57,110 --> 02:31:54,319
this is

3881
02:31:59,030 --> 02:31:57,120
many range sizes we have our plants we

3882
02:32:01,590 --> 02:31:59,040
have our microphone our mites these

3883
02:32:03,830 --> 02:32:01,600
nematodes amoebas if we really want to

3884
02:32:06,870 --> 02:32:03,840
look at microbes and viruses we have to

3885
02:32:09,670 --> 02:32:06,880
zoom in on this one soil particle

3886
02:32:11,750 --> 02:32:09,680
now we can see our bacteria in some poor

3887
02:32:13,830 --> 02:32:11,760
water and then if we zoom in further

3888
02:32:15,670 --> 02:32:13,840

then we can see our viruses

3889
02:32:17,750 --> 02:32:15,680
so another point is out there because

3890
02:32:19,910 --> 02:32:17,760
they're just here with all these other

3891
02:32:21,990 --> 02:32:19,920
organisms and they can be random in

3892
02:32:24,550 --> 02:32:22,000
random spots in the soil

3893
02:32:26,790 --> 02:32:24,560
and most of what we know about viruses

3894
02:32:29,030 --> 02:32:26,800
comes from isolates which are extremely

3895
02:32:32,309 --> 02:32:29,040
hard to isolate as we've heard earlier

3896
02:32:34,630 --> 02:32:32,319
and then from metagenomics

3897
02:32:36,710 --> 02:32:34,640
because there's so many organisms with

3898
02:32:39,830 --> 02:32:36,720
their larger genomes when we do

3899
02:32:41,990 --> 02:32:39,840
metagenomes of a soil environment we get

3900
02:32:44,469 --> 02:32:42,000
low resolution on viruses typically less

3901
02:32:48,230 --> 02:32:44,479
than two percent of our information goes

3902
02:32:52,070 --> 02:32:49,670
from this we've learned that there's

3903
02:32:54,710 --> 02:32:52,080
about 10 million to a billion viruses

3904
02:32:57,590 --> 02:32:54,720
per gram of soil and because soil is

3905
02:33:00,630 --> 02:32:57,600
structured and has micro heterogeneity

3906
02:33:02,309 --> 02:33:00,640
the amount of viruses per house can vary

3907
02:33:04,790 --> 02:33:02,319
from about one to a thousand or even

3908
02:33:08,950 --> 02:33:06,630
we've seen that soul viruses are

3909
02:33:11,190 --> 02:33:08,960
morphologically diverse we have our

3910
02:33:13,510 --> 02:33:11,200
classic phage or double-stranded dna

3911
02:33:15,190 --> 02:33:13,520
viruses we have our smaller

3912
02:33:17,590 --> 02:33:15,200
single-stranded dna viruses we can have

3913
02:33:20,309 --> 02:33:17,600

some filamentous viruses and some pill

3914
02:33:22,070 --> 02:33:20,319
leaking viruses

3915
02:33:23,990 --> 02:33:22,080
but with recent metagenomics and

3916
02:33:26,150 --> 02:33:24,000
viromics which is separating the viral

3917
02:33:28,550 --> 02:33:26,160
particles before sequencing we've been

3918
02:33:30,630 --> 02:33:28,560
able to start learning a lot

3919
02:33:32,830 --> 02:33:30,640
the main thing is that we've been able

3920
02:33:34,710 --> 02:33:32,840
to recover thousands of viral

3921
02:33:36,870 --> 02:33:34,720
populations and these are different

3922
02:33:39,110 --> 02:33:36,880
flavors of viruses that are similar to

3923
02:33:41,670 --> 02:33:39,120
group into these populations and we use

3924
02:33:43,670 --> 02:33:41,680
a 10 kb threshold to get really robust

3925
02:33:46,230 --> 02:33:43,680
populations

3926
02:33:48,389 --> 02:33:46,240
by looking at their genomes we've been

3927
02:33:50,389 --> 02:33:48,399
able to identify that viruses can have

3928
02:33:52,710 --> 02:33:50,399
direct impacts on microbial body

3929
02:33:55,270 --> 02:33:52,720
geochemistry not only from lysine of

3930
02:33:57,110 --> 02:33:55,280
dominant microbial hosts but also

3931
02:33:59,190 --> 02:33:57,120
from having the host express the

3932
02:34:00,389 --> 02:33:59,200
auxiliary metabolic genes that they

3933
02:34:02,469 --> 02:34:00,399
carry

3934
02:34:03,910 --> 02:34:02,479
in this example we have viruses killing

3935
02:34:05,950 --> 02:34:03,920
some dominant hosts

3936
02:34:08,389 --> 02:34:05,960
they have an amg which is called a

3937
02:34:09,910 --> 02:34:08,399
glycohydrolase in this case

3938
02:34:11,349 --> 02:34:09,920

and these can help break down these

3939
02:34:13,590 --> 02:34:11,359
complex carbohydrates these

3940
02:34:15,590 --> 02:34:13,600
polysaccharides into these small

3941
02:34:19,030 --> 02:34:15,600
monomers that are digestible and can

3942
02:34:21,349 --> 02:34:19,040
feed in a range of metabolisms

3943
02:34:22,550 --> 02:34:21,359
a kind of example i like to give is

3944
02:34:24,309 --> 02:34:22,560
imagine an environment where there's

3945
02:34:26,469 --> 02:34:24,319
pineapples everywhere and you have to

3946
02:34:27,910 --> 02:34:26,479
eat the rind and everything this is

3947
02:34:30,150 --> 02:34:27,920
really hard to do and not everyone can

3948
02:34:31,910 --> 02:34:30,160
do this and people don't want to do this

3949
02:34:34,070 --> 02:34:31,920
so this is broken down and let's say you

3950
02:34:35,590 --> 02:34:34,080
have cheeseburgers and salads around

3951
02:34:37,110 --> 02:34:35,600
this is great and a lot of different

3952
02:34:39,429 --> 02:34:37,120
people a lot of different organisms can

3953
02:34:42,309 --> 02:34:39,439
actually digest this and function off

3954
02:34:45,750 --> 02:34:43,510
now

3955
02:34:46,950 --> 02:34:45,760
i've told you from metagenomics we've

3956
02:34:49,270 --> 02:34:46,960
learned a lot of this and it's from

3957
02:34:51,429 --> 02:34:49,280
these recent vat um advancements along

3958
02:34:53,990 --> 02:34:51,439
with these laborious virus but if we

3959
02:34:56,950 --> 02:34:54,000
think back there's another issue looking

3960
02:34:58,870 --> 02:34:56,960
back at the soil

3961
02:35:00,950 --> 02:34:58,880
not all of the microbes here and the

3962
02:35:02,230 --> 02:35:00,960
viruses are active at any point so when

3963
02:35:03,910 --> 02:35:02,240

you take a metagenome you're just

3964
02:35:07,590 --> 02:35:03,920
getting everything that's there and we

3965
02:35:09,030 --> 02:35:07,600
know there can be relic dna in soils

3966
02:35:12,389 --> 02:35:09,040
so what do we actually care about we

3967
02:35:13,830 --> 02:35:12,399
care about the microbes that are active

3968
02:35:15,110 --> 02:35:13,840
and to help explain this these are ones

3969
02:35:16,469 --> 02:35:15,120
that are growing and actively

3970
02:35:17,910 --> 02:35:16,479
interacting contributing to the

3971
02:35:19,830 --> 02:35:17,920
environment not ones that are

3972
02:35:21,510 --> 02:35:19,840
necessarily dormant or deceased which

3973
02:35:23,110 --> 02:35:21,520
when we take medicine and we are seeing

3974
02:35:27,270 --> 02:35:23,120
all of this which is biasing what we

3975
02:35:31,030 --> 02:35:29,510
so the goals here are to increase

3976
02:35:33,030 --> 02:35:31,040
resolution on viruses so we can

3977
02:35:34,950 --> 02:35:33,040
understand them in their ecology and

3978
02:35:38,710 --> 02:35:34,960
then target the active microbes and the

3979
02:35:41,349 --> 02:35:40,469
so normally how would you track an

3980
02:35:44,309 --> 02:35:41,359
animal

3981
02:35:46,630 --> 02:35:44,319
well for a cat here we have a collar

3982
02:35:49,510 --> 02:35:46,640
for butterfly sharks we can put a tag on

3983
02:35:50,389 --> 02:35:49,520
them but viruses are way too small for

3984
02:35:52,389 --> 02:35:50,399
this

3985
02:35:54,870 --> 02:35:52,399
and like i said there's many methods but

3986
02:35:57,429 --> 02:35:54,880
the one that i use are stable isotopes

3987
02:35:59,110 --> 02:35:57,439
it's also called stable isotope probate

3988
02:36:01,510 --> 02:35:59,120

and stable isotopes are atoms that

3989
02:36:03,590 --> 02:36:01,520
contain the same number of protons but

3990
02:36:05,990 --> 02:36:03,600
differ in the number of neutrons so an

3991
02:36:07,990 --> 02:36:06,000
example i have here we have hydrogen

3992
02:36:11,030 --> 02:36:08,000
which is stable and exists in the

3993
02:36:13,349 --> 02:36:11,040
environment and the most um

3994
02:36:16,309 --> 02:36:13,359
the one that we most likely see is that

3995
02:36:19,590 --> 02:36:16,319
it has one proton and one electron but

3996
02:36:21,429 --> 02:36:19,600
it also exists in the form of one proton

3997
02:36:22,950 --> 02:36:21,439
and one neutron

3998
02:36:24,630 --> 02:36:22,960
and the difference is we have a little

3999
02:36:27,510 --> 02:36:24,640
bit more mass and we can use that to our

4000
02:36:34,150 --> 02:36:31,190
so here we can incorporate atoms

4001
02:36:36,550 --> 02:36:34,160
that these elements into compounds that

4002
02:36:39,270 --> 02:36:36,560
can help us track biological processes

4003
02:36:41,270 --> 02:36:39,280
so here we have water h2o two hydrogens

4004
02:36:43,030 --> 02:36:41,280
one oxygen there's eight protons and

4005
02:36:44,230 --> 02:36:43,040
eight neutrons let's call this the

4006
02:36:45,830 --> 02:36:44,240
tootsie roll

4007
02:36:47,750 --> 02:36:45,840
now we can have heavy water and this is

4008
02:36:49,830 --> 02:36:47,760
not with deuterium this is two hydrogens

4009
02:36:52,469 --> 02:36:49,840
and one oxygen but it's 18 0 so it's

4010
02:36:54,309 --> 02:36:52,479
eight protons and 10 neutrons so it has

4011
02:36:55,990 --> 02:36:54,319
a higher mass so these are different

4012
02:36:59,270 --> 02:36:56,000
flavors of our water these are like our

4013
02:37:02,469 --> 02:37:00,790

so i'm going to give you two examples of

4014
02:37:04,710 --> 02:37:02,479
how i use these in experiments the first

4015
02:37:06,710 --> 02:37:04,720
one is we have this heavy water and then

4016
02:37:08,469 --> 02:37:06,720
regular abundance water and we incubate

4017
02:37:11,190 --> 02:37:08,479
that in soil

4018
02:37:13,750 --> 02:37:11,200
we can also be more specific here we

4019
02:37:18,070 --> 02:37:13,760
have plant biomass and we can grow

4020
02:37:19,990 --> 02:37:18,080
plants with enriched carbon with 13 co2

4021
02:37:23,750 --> 02:37:20,000
so the whole plant is enriched and feed

4022
02:37:27,830 --> 02:37:26,070
so with the heavy water we can get at

4023
02:37:30,630 --> 02:37:27,840
all the active organisms because all

4024
02:37:33,429 --> 02:37:30,640
organisms use water and then any viruses

4025
02:37:35,910 --> 02:37:33,439
that infect them will also get labeled

4026
02:37:38,070 --> 02:37:35,920
now for the plant biomass this is

4027
02:37:40,309 --> 02:37:38,080
specific organisms we get a higher level

4028
02:37:42,630 --> 02:37:40,319
of detail here we're looking at microbes

4029
02:37:44,550 --> 02:37:42,640
that are pre breaking down this plant

4030
02:37:47,429 --> 02:37:44,560
biomass and those viruses are infecting

4031
02:37:49,270 --> 02:37:47,439
those specific microbes

4032
02:37:51,630 --> 02:37:49,280
we can extract the dna like we would do

4033
02:37:53,750 --> 02:37:51,640
with a normal metagenome we can do

4034
02:37:56,150 --> 02:37:53,760
ultrasonification with a cesium chloride

4035
02:37:59,270 --> 02:37:56,160
density gradient in order to separate

4036
02:38:01,270 --> 02:37:59,280
the light dna and the heavy dna so we

4037
02:38:03,190 --> 02:38:01,280
can have a dormant and deceased microbes

4038
02:38:05,510 --> 02:38:03,200

in a different area than we have our

4039
02:38:08,230 --> 02:38:05,520
active microbes that we want

4040
02:38:10,550 --> 02:38:08,240
so by being able to separate this dna we

4041
02:38:12,150 --> 02:38:10,560
can sequence it separately and we can do

4042
02:38:14,469 --> 02:38:12,160
comparative bioinformatics if we want to

4043
02:38:16,469 --> 02:38:14,479
compare that but by separating it we can

4044
02:38:18,230 --> 02:38:16,479
have more sequencing power going towards

4045
02:38:20,550 --> 02:38:18,240
the relevant microbes and viruses that

4046
02:38:24,790 --> 02:38:20,560
we care about and this helps reduce

4047
02:38:26,870 --> 02:38:24,800
complexity and increase our resolution

4048
02:38:28,389 --> 02:38:26,880
so our goals were to increase resolution

4049
02:38:29,990 --> 02:38:28,399
and it's going to do that and then we

4050
02:38:31,510 --> 02:38:30,000
can target our active microbes and their

4051
02:38:34,550 --> 02:38:31,520
viruses and it's hopefully going to do

4052
02:38:38,389 --> 02:38:35,830
so when we were conducting this

4053
02:38:39,830 --> 02:38:38,399
experiment our two main goals was to try

4054
02:38:41,190 --> 02:38:39,840
on two different types of soils that

4055
02:38:43,510 --> 02:38:41,200
were extremely different so different

4056
02:38:45,110 --> 02:38:43,520
soil biomes and then see if we can if it

4057
02:38:47,349 --> 02:38:45,120
actually works if we can identify active

4058
02:38:49,349 --> 02:38:47,359
viruses and their microbes so the two

4059
02:38:51,510 --> 02:38:49,359
field site which two field sites that we

4060
02:38:53,349 --> 02:38:51,520
chose were two long-term ecological

4061
02:38:54,870 --> 02:38:53,359
research sites the first one was in

4062
02:38:57,270 --> 02:38:54,880
alaska it's a partially thawed

4063
02:38:59,349 --> 02:38:57,280

permafrost bog habitat so it's frigid

4064
02:39:02,070 --> 02:38:59,359
it's cold we don't expect life to be

4065
02:39:04,070 --> 02:39:02,080
there below freezing our other one is in

4066
02:39:06,070 --> 02:39:04,080
puerto rico it's in a tropical rain

4067
02:39:07,990 --> 02:39:06,080
forest in the la cuyo experimental

4068
02:39:10,870 --> 02:39:08,000
forest and it's a highly dynamic

4069
02:39:13,830 --> 02:39:10,880
tropical rainforest

4070
02:39:15,750 --> 02:39:13,840
so zooming in on the alaska site i first

4071
02:39:17,750 --> 02:39:15,760
wanted to point out that everything you

4072
02:39:19,429 --> 02:39:17,760
see in purple is permafrost and

4073
02:39:20,950 --> 02:39:19,439
discontinuous permafrost which is ground

4074
02:39:23,270 --> 02:39:20,960
that's frozen for two or more

4075
02:39:25,429 --> 02:39:23,280
consecutive years

4076
02:39:27,429 --> 02:39:25,439
so zooming in more at our bonanza creek

4077
02:39:32,710 --> 02:39:27,439
site where the average temperature is

4078
02:39:38,710 --> 02:39:35,270
sorry we have some delay um we collected

4079
02:39:39,830 --> 02:39:38,720
some soil cores from this field site

4080
02:39:41,270 --> 02:39:39,840
and i'm going to quickly go through the

4081
02:39:43,750 --> 02:39:41,280
methods so you can get an idea of what

4082
02:39:46,389 --> 02:39:43,760
we actually have to do we took two grams

4083
02:39:48,710 --> 02:39:46,399
of soil from the top 10 centimeters of

4084
02:39:50,630 --> 02:39:48,720
our cores and we incubated that and we

4085
02:39:52,550 --> 02:39:50,640
pressurized it to remove water and put

4086
02:39:54,950 --> 02:39:52,560
in either our

4087
02:39:56,630 --> 02:39:54,960
natural abundance water or a heavy water

4088
02:39:59,030 --> 02:39:56,640

and then we incubated that for half a

4089
02:40:00,950 --> 02:39:59,040
year and for a full year

4090
02:40:03,030 --> 02:40:00,960
and from this we collected a lot of

4091
02:40:04,790 --> 02:40:03,040
information but i'm only going to talk

4092
02:40:07,030 --> 02:40:04,800
about the metagenomic information that

4093
02:40:10,790 --> 02:40:07,040
we gathered

4094
02:40:13,030 --> 02:40:10,800
so looking at the 23 meta genomes

4095
02:40:14,150 --> 02:40:13,040
we were first able to identify and

4096
02:40:19,190 --> 02:40:14,160
assemble

4097
02:40:22,630 --> 02:40:19,200
genomes these are microbial highly wrote

4098
02:40:23,830 --> 02:40:22,640
um these are almost whole medellin are

4099
02:40:27,030 --> 02:40:23,840
genomes that we assembled from the

4100
02:40:29,269 --> 02:40:27,040
metagenomic data so we found 30 microbes

4101
02:40:31,510 --> 02:40:29,279
that are active below freezing they're

4102
02:40:34,070 --> 02:40:31,520
interacting with the environment below

4103
02:40:35,750 --> 02:40:34,080
freezing

4104
02:40:39,670 --> 02:40:35,760
this bacterial host spanned three

4105
02:40:42,790 --> 02:40:40,790
and then we were

4106
02:40:44,230 --> 02:40:42,800
again sorry so back to the bacteria

4107
02:40:46,309 --> 02:40:44,240
active i really wanted to drive that

4108
02:40:48,070 --> 02:40:46,319
home this is an important concept that

4109
02:40:52,230 --> 02:40:48,080
even though the ground is frozen they

4110
02:40:56,230 --> 02:40:54,230
from this we are able to detect about 4

4111
02:40:57,830 --> 02:40:56,240
000 viruses and we use two different

4112
02:41:00,230 --> 02:40:57,840
viral detection methods the first is

4113
02:41:02,630 --> 02:41:00,240

virus order and the second is deep veer

4114
02:41:04,389 --> 02:41:02,640
finder

4115
02:41:06,230 --> 02:41:04,399
now a lot of these viruses can be very

4116
02:41:07,830 --> 02:41:06,240
similar so we group them into viral

4117
02:41:09,429 --> 02:41:07,840
populations

4118
02:41:10,870 --> 02:41:09,439
so it's like having different flavors of

4119
02:41:12,550 --> 02:41:10,880
the same virus or different strains of

4120
02:41:14,070 --> 02:41:12,560
this virus

4121
02:41:15,990 --> 02:41:14,080
so i put up here the different methods

4122
02:41:19,830 --> 02:41:16,000
you can go back and look at how we group

4123
02:41:23,510 --> 02:41:21,750
our next thing is to link these viruses

4124
02:41:25,429 --> 02:41:23,520
to host now these these are preliminary

4125
02:41:27,110 --> 02:41:25,439
results we still have more to go through

4126
02:41:29,510 --> 02:41:27,120
but at first we did just nucleotide

4127
02:41:32,630 --> 02:41:29,520
identity which we took the genome of the

4128
02:41:38,230 --> 02:41:32,640
virus and we look for similarity with uh

4129
02:41:42,070 --> 02:41:39,670
so the first thing i'm going to show you

4130
02:41:44,389 --> 02:41:42,080
is this blue heat map here so to help

4131
02:41:47,670 --> 02:41:44,399
you digest it on the x-axis we have our

4132
02:41:49,910 --> 02:41:47,680
332 viral populations and on the y-axis

4133
02:41:51,990 --> 02:41:49,920
we have our different sample treatments

4134
02:41:54,389 --> 02:41:52,000
the two blue ones at the top are our

4135
02:41:56,870 --> 02:41:54,399
natural abundance water the two red on

4136
02:41:58,309 --> 02:41:56,880
the bottom are our heavy water and you

4137
02:41:59,670 --> 02:41:58,319
can see we have it for a half a year and

4138
02:42:01,590 --> 02:41:59,680

for a full year

4139
02:42:04,469 --> 02:42:01,600
now when you look at the heat map the

4140
02:42:06,950 --> 02:42:04,479
darker the color the more abundance more

4141
02:42:09,750 --> 02:42:06,960
abundant that viral population is so

4142
02:42:11,910 --> 02:42:09,760
first we detected a lot of viruses below

4143
02:42:13,910 --> 02:42:11,920
freezing this is incredible we have

4144
02:42:17,269 --> 02:42:13,920
active microbes and viruses below

4145
02:42:19,910 --> 02:42:18,790
so we were able to

4146
02:42:21,110 --> 02:42:19,920
like i said we could do comparative

4147
02:42:22,550 --> 02:42:21,120
bioinformatics we're able to separate

4148
02:42:24,230 --> 02:42:22,560
active and non-active so here i pull it

4149
02:42:25,750 --> 02:42:24,240
together for one figure so we have our

4150
02:42:29,110 --> 02:42:25,760
non-active at top

4151
02:42:32,150 --> 02:42:29,120
and our active at bottom and we have 256

4152
02:42:33,110 --> 02:42:32,160
active viral populations

4153
02:42:34,870 --> 02:42:33,120
now the first thing i want to show you

4154
02:42:37,190 --> 02:42:34,880
is if we look at these viruses that were

4155
02:42:38,950 --> 02:42:37,200
there at a half a year to a full year we

4156
02:42:42,230 --> 02:42:38,960
can see that a lot of viruses actually

4157
02:42:46,309 --> 02:42:44,389
68 of these viruses that persist

4158
02:42:49,510 --> 02:42:46,319
actually increase their abundance and

4159
02:42:51,190 --> 02:42:49,520
this can be for many different reasons

4160
02:42:53,349 --> 02:42:51,200
these could be temperate viruses

4161
02:42:55,670 --> 02:42:53,359
propagating microbial hosts these could

4162
02:42:57,590 --> 02:42:55,680
be virions persisting in the soils

4163
02:42:59,269 --> 02:42:57,600

themselves or this could be that they

4164
02:43:01,110 --> 02:42:59,279
burst open they got degraded but the

4165
02:43:05,910 --> 02:43:01,120
viral dna itself is accumulating the

4166
02:43:09,590 --> 02:43:07,750
if you also look to the far left you'll

4167
02:43:11,269 --> 02:43:09,600
see that some viruses are actually gone

4168
02:43:12,550 --> 02:43:11,279
so we have viral populations here that

4169
02:43:14,469 --> 02:43:12,560
were abundant

4170
02:43:17,190 --> 02:43:14,479
and six months later we no longer see

4171
02:43:21,990 --> 02:43:19,830
we've also seen new viruses emerge so we

4172
02:43:26,790 --> 02:43:22,000
have a dynamic viral community that is

4173
02:43:30,710 --> 02:43:28,790
so just hitting home active viruses

4174
02:43:32,469 --> 02:43:30,720
below freezing before i look into our

4175
02:43:35,349 --> 02:43:32,479
other samples

4176
02:43:37,830 --> 02:43:35,359
so now moving on to our highly dynamic

4177
02:43:40,469 --> 02:43:37,840
tropical rainforest

4178
02:43:45,429 --> 02:43:40,479
this is in puerto rico i put up some

4179
02:43:50,469 --> 02:43:48,469
it's in the northeast of puerto rico oh

4180
02:43:53,429 --> 02:43:50,479
sorry we got some lag going on okay so

4181
02:43:55,670 --> 02:43:53,439
you can see from here it's a highly lush

4182
02:43:59,670 --> 02:43:55,680
beautiful environment totally different

4183
02:44:02,230 --> 02:44:01,110
so here we approach things a little

4184
02:44:04,630 --> 02:44:02,240
differently

4185
02:44:06,950 --> 02:44:04,640
we took enriched biomass and natural

4186
02:44:10,150 --> 02:44:06,960
abundant biomass and we added that to 20

4187
02:44:11,269 --> 02:44:10,160
grams of soil and we let that incubate

4188
02:44:13,349 --> 02:44:11,279

now we didn't want to just do that we

4189
02:44:15,830 --> 02:44:13,359
wanted to mimic the environment and this

4190
02:44:17,910 --> 02:44:15,840
environment has heavy rains that occur

4191
02:44:19,990 --> 02:44:17,920
often so these soils go from oxic to

4192
02:44:21,590 --> 02:44:20,000
anoxic frequently and how does this

4193
02:44:23,670 --> 02:44:21,600
affect the microbial communities and the

4194
02:44:26,389 --> 02:44:23,680
viral communities

4195
02:44:28,070 --> 02:44:26,399
so our four treatments were static oxic

4196
02:44:30,070 --> 02:44:28,080
and then we did a high and low frequency

4197
02:44:32,230 --> 02:44:30,080
of oxic and anoxic and then we did

4198
02:44:36,469 --> 02:44:32,240
totally anoxic which is all controlled

4199
02:44:41,830 --> 02:44:39,429
so from this we we incubated for 44 days

4200
02:44:43,590 --> 02:44:41,840
and we generated 85 sip fraction

4201
02:44:46,070 --> 02:44:43,600
metagenomes and this is where we divided

4202
02:44:47,510 --> 02:44:46,080
the dna on the cesium chloride density

4203
02:44:49,750 --> 02:44:47,520
gradient and then we sequence those

4204
02:44:51,510 --> 02:44:49,760
separately but then we also took all the

4205
02:44:53,670 --> 02:44:51,520
dna together and just sampled and

4206
02:44:55,750 --> 02:44:53,680
sequenced what we call the bulk dna this

4207
02:44:58,710 --> 02:44:55,760
is a microbial metagenome as simon

4208
02:45:03,670 --> 02:45:00,870
we able to we were able to get 95

4209
02:45:05,830 --> 02:45:03,680
metagenomes from this

4210
02:45:08,070 --> 02:45:05,840
which formed into 214 different

4211
02:45:09,510 --> 02:45:08,080
microbial assembled genomes so these are

4212
02:45:11,910 --> 02:45:09,520
genomes from the microbe from the

4213
02:45:13,750 --> 02:45:11,920

metagenomic dataset

4214
02:45:17,830 --> 02:45:13,760
and these span four different phy left

4215
02:45:21,349 --> 02:45:19,750
again we want to detect our viruses so

4216
02:45:26,550 --> 02:45:21,359
we used our same detection methods and

4217
02:45:31,990 --> 02:45:29,510
we're able to link 11 of these viruses

4218
02:45:33,349 --> 02:45:32,000
to hosts from these same soils and more

4219
02:45:34,550 --> 02:45:33,359
work is going on with this so you'll

4220
02:45:37,030 --> 02:45:34,560
definitely see something in the near

4221
02:45:41,590 --> 02:45:39,510
so bringing it back to a heat map here i

4222
02:45:43,429 --> 02:45:41,600
now use colors to remind you this is no

4223
02:45:45,590 --> 02:45:43,439
longer this cold frigid environment but

4224
02:45:48,389 --> 02:45:45,600
this is this lush

4225
02:45:50,230 --> 02:45:48,399
fluctuating anoxic toxic environment so

4226
02:45:52,870 --> 02:45:50,240
at the top here in black we have our

4227
02:45:54,870 --> 02:45:52,880
bulk metagenome samples and then in blue

4228
02:45:57,030 --> 02:45:54,880
and red we have our sip fractionated

4229
02:45:59,670 --> 02:45:57,040
meta genomes in light in the light blue

4230
02:46:02,150 --> 02:45:59,680
we have our 12c our natural abundance

4231
02:46:05,190 --> 02:46:02,160
and in red we have our enriched samples

4232
02:46:07,030 --> 02:46:05,200
so these are microbes that ate plant

4233
02:46:11,349 --> 02:46:07,040
biomass and these are the viruses that

4234
02:46:15,190 --> 02:46:13,110
so the first thing i want to point out

4235
02:46:17,510 --> 02:46:15,200
is that by comparing the bulk to sip

4236
02:46:20,389 --> 02:46:17,520
fraction and metagenomes we're able to

4237
02:46:22,630 --> 02:46:20,399
get eight percent more viral populations

4238
02:46:24,790 --> 02:46:22,640

so by doing sip fractionation we did

4239
02:46:26,469 --> 02:46:24,800
decrease complexity increased resolution

4240
02:46:27,910 --> 02:46:26,479
on viruses we're getting more viral

4241
02:46:32,389 --> 02:46:27,920
diversity from doing the sip

4242
02:46:36,950 --> 02:46:34,389
now if we zoom in just on the active

4243
02:46:38,870 --> 02:46:36,960
viruses and we the way we did this is

4244
02:46:40,950 --> 02:46:38,880
that they're active in the natural

4245
02:46:42,870 --> 02:46:40,960
abundance and in the enriched samples we

4246
02:46:45,670 --> 02:46:42,880
took them out so these are only ones

4247
02:46:48,230 --> 02:46:45,680
that were recorded in the active sorry

4248
02:46:51,349 --> 02:46:48,240
in the 13 scene rich plant biomass

4249
02:46:55,510 --> 02:46:53,030
the first thing you should immediately

4250
02:46:57,349 --> 02:46:55,520
see is that from the oxic samples we

4251
02:47:00,389 --> 02:46:57,359
have more diversity of viruses we have

4252
02:47:02,550 --> 02:47:00,399
more viral populations so oxygen is just

4253
02:47:04,790 --> 02:47:02,560
definitely affecting our

4254
02:47:06,790 --> 02:47:04,800
correlating with our viruses

4255
02:47:09,429 --> 02:47:06,800
the other thing you should note though

4256
02:47:11,910 --> 02:47:09,439
is that we have unique viruses in our

4257
02:47:13,590 --> 02:47:11,920
anoxic samples that are not when is

4258
02:47:15,349 --> 02:47:13,600
non-oxic samples

4259
02:47:17,910 --> 02:47:15,359
so why is it that these viruses would

4260
02:47:20,469 --> 02:47:17,920
infect hosts that have these slower

4261
02:47:22,550 --> 02:47:20,479
metabolisms is this a niche

4262
02:47:24,150 --> 02:47:22,560
differentiation or what is this so

4263
02:47:27,110 --> 02:47:24,160

definitely going to be looking more into

4264
02:47:31,750 --> 02:47:28,790
so to summarize and bring it all back

4265
02:47:33,590 --> 02:47:31,760
together from our first experiment i i

4266
02:47:35,510 --> 02:47:33,600
first actually just want to say that the

4267
02:47:38,389 --> 02:47:35,520
sip process worked we were able to

4268
02:47:39,910 --> 02:47:38,399
identify active microbes and viruses and

4269
02:47:41,750 --> 02:47:39,920
this reduced our complexity so we got

4270
02:47:43,349 --> 02:47:41,760
increased resolution and these are the

4271
02:47:45,269 --> 02:47:43,359
metabolisms that we care about because

4272
02:47:47,429 --> 02:47:45,279
they're actively interacting with the

4273
02:47:48,870 --> 02:47:47,439
environment so for our first site this

4274
02:47:50,230 --> 02:47:48,880
is the bonanza creek this is our

4275
02:47:51,910 --> 02:47:50,240
permafrost

4276
02:47:54,389 --> 02:47:51,920
not only did we identify these microbes

4277
02:47:57,429 --> 02:47:54,399
and viruses below freezing so they are

4278
02:47:59,910 --> 02:47:57,439
contributing in these frozen soils but

4279
02:48:01,590 --> 02:47:59,920
we also see this temporal

4280
02:48:02,550 --> 02:48:01,600
succession we see this change in viral

4281
02:48:03,990 --> 02:48:02,560
community

4282
02:48:06,070 --> 02:48:04,000
and the evidence that viruses can

4283
02:48:10,309 --> 02:48:06,080
persist environment but more work is

4284
02:48:15,030 --> 02:48:12,150
switching over to our highly dynamic

4285
02:48:17,510 --> 02:48:15,040
soils we are able to link viruses

4286
02:48:19,750 --> 02:48:17,520
to key microbes that break down and

4287
02:48:21,590 --> 02:48:19,760
control the fate of this organic carbon

4288
02:48:23,590 --> 02:48:21,600

in the soil environments

4289
02:48:26,550 --> 02:48:23,600
and that we saw that redox strongly

4290
02:48:28,630 --> 02:48:26,560
influenced our virus activity and that

4291
02:48:31,190 --> 02:48:28,640
the sip fractions recovered more viral

4292
02:48:33,110 --> 02:48:31,200
populations

4293
02:48:34,630 --> 02:48:33,120
so to synthesize what we gained from

4294
02:48:37,269 --> 02:48:34,640
both these environments is that we got a

4295
02:48:39,030 --> 02:48:37,279
lot of cool novel viruses here and the

4296
02:48:41,349 --> 02:48:39,040
sip fraction metagenomes helped us see

4297
02:48:43,429 --> 02:48:41,359
stuff that we couldn't see otherwise

4298
02:48:45,750 --> 02:48:43,439
i also wanted to point out we put a lot

4299
02:48:48,389 --> 02:48:45,760
more effort into the tropical soils with

4300
02:48:50,309 --> 02:48:48,399
8x more sequencing but we only had about

4301
02:48:51,830 --> 02:48:50,319
double the number of viral populations

4302
02:48:53,830 --> 02:48:51,840
and there can be many reasons for this

4303
02:48:55,750 --> 02:48:53,840
you could think maybe viral diversity

4304
02:48:56,950 --> 02:48:55,760
doesn't track with microbial diversity

4305
02:48:58,950 --> 02:48:56,960
or it could be that there's less

4306
02:49:00,710 --> 02:48:58,960
organisms reducing the metagenome

4307
02:49:03,830 --> 02:49:00,720
complexity

4308
02:49:05,910 --> 02:49:03,840
finally the metabolic repertoire between

4309
02:49:07,429 --> 02:49:05,920
the active host and the dormant host was

4310
02:49:08,710 --> 02:49:07,439
completely different so we need to think

4311
02:49:10,710 --> 02:49:08,720
about this when we're looking at our

4312
02:49:12,389 --> 02:49:10,720
meta genomes unless we're actually

4313
02:49:14,230 --> 02:49:12,399

looking at the active ones we can be

4314
02:49:16,309 --> 02:49:14,240
seeing signals of stuff that's actually

4315
02:49:18,389 --> 02:49:16,319
not going on and how do we infer this if

4316
02:49:19,670 --> 02:49:18,399
we're looking at biosignatures or if

4317
02:49:21,190 --> 02:49:19,680
we're looking up gases and we're trying

4318
02:49:22,550 --> 02:49:21,200
to relate it to microbes we need to make

4319
02:49:24,830 --> 02:49:22,560
sure that they're active when we're

4320
02:49:27,110 --> 02:49:24,840
taking those

4321
02:49:29,590 --> 02:49:27,120
samples so i just want to acknowledge

4322
02:49:30,950 --> 02:49:29,600
everyone at jgi and lawrence livermore

4323
02:49:32,870 --> 02:49:30,960
national love the help especially my

4324
02:49:34,469 --> 02:49:32,880
mentor steve blazewicz and jennifer

4325
02:49:35,590 --> 02:49:34,479
pepridge

4326
02:49:51,830 --> 02:49:35,600
and with that

4327
02:49:51,840 --> 02:49:57,750
some online

4328
02:50:06,550 --> 02:49:59,670
you see that gary

4329
02:50:10,710 --> 02:50:08,309
nothing could be used for snow dwelling

4330
02:50:13,190 --> 02:50:10,720
viruses what about aquatic wires yeah so

4331
02:50:15,670 --> 02:50:13,200
this this technique can be applied in

4332
02:50:18,630 --> 02:50:15,680
any environment it's not just for soils

4333
02:50:20,309 --> 02:50:18,640
that's the one that i cared about um

4334
02:50:22,389 --> 02:50:20,319
there has been work

4335
02:50:24,469 --> 02:50:22,399
i posted a paper earlier that uses bond

4336
02:50:27,110 --> 02:50:24,479
cat and uses nano sims to look at it

4337
02:50:29,510 --> 02:50:27,120
with with uh bonk as just another way to

4338
02:50:30,630 --> 02:50:29,520

look at activity of in-situ protein

4339
02:50:32,070 --> 02:50:30,640
synthesis

4340
02:50:34,550 --> 02:50:32,080
um

4341
02:50:36,710 --> 02:50:34,560
i can also post uh

4342
02:50:38,550 --> 02:50:36,720
a book from one of my mentors that she

4343
02:50:39,990 --> 02:50:38,560
talks about how we can use stabilized

4344
02:50:42,550 --> 02:50:40,000
jennifer beverage talk about how we use

4345
02:50:44,710 --> 02:50:42,560
stable isotopes to look at microbes and

4346
02:50:46,150 --> 02:50:44,720
their viruses and gives a great review

4347
02:50:47,670 --> 02:50:46,160
on the different isotopes you can use

4348
02:50:49,910 --> 02:50:47,680
different compounds can be applied to

4349
02:50:52,469 --> 02:50:49,920
and what are outcomes and prospects of

4350
02:50:56,790 --> 02:50:55,110
okay arvin has a question

4351
02:50:57,910 --> 02:50:56,800
hi

4352
02:50:59,510 --> 02:50:57,920
thanks for the

4353
02:51:02,389 --> 02:50:59,520
talk i've got a good question and this

4354
02:51:04,790 --> 02:51:02,399
is your observation that you're noticing

4355
02:51:07,990 --> 02:51:04,800
in your say your

4356
02:51:11,030 --> 02:51:08,000
uh soils in puerto rico and because the

4357
02:51:13,190 --> 02:51:11,040
soils go through all these variations

4358
02:51:15,030 --> 02:51:13,200
and i'm wondering whether the viruses

4359
02:51:16,550 --> 02:51:15,040
you're finding in those ecosystems are

4360
02:51:18,790 --> 02:51:16,560
more generalists

4361
02:51:20,309 --> 02:51:18,800
so that it's it's a way it's a mechanism

4362
02:51:21,590 --> 02:51:20,319
for the virus to persist in that

4363
02:51:24,630 --> 02:51:21,600

environment

4364
02:51:26,389 --> 02:51:24,640
without actually leaving the gene pool

4365
02:51:28,710 --> 02:51:26,399
and so with that you might have less

4366
02:51:30,630 --> 02:51:28,720
diversity of viruses

4367
02:51:32,550 --> 02:51:30,640
which

4368
02:51:34,950 --> 02:51:32,560
is kind of what i expect to see in these

4369
02:51:36,309 --> 02:51:34,960
kind of systems which is not far from a

4370
02:51:38,870 --> 02:51:36,319
system that you would see in an

4371
02:51:40,790 --> 02:51:38,880
agricultural setup where you have crop

4372
02:51:42,469 --> 02:51:40,800
rotation or something and you have

4373
02:51:45,110 --> 02:51:42,479
reservoir species

4374
02:51:47,510 --> 02:51:45,120
that the viruses go to when a crop is

4375
02:51:49,830 --> 02:51:47,520
taken out of place and then you go back

4376
02:51:51,750 --> 02:51:49,840
into the same circle so this is kind of

4377
02:51:54,070 --> 02:51:51,760
what i think might be a reason you're

4378
02:51:56,870 --> 02:51:54,080
seeing low that lower diversity compared

4379
02:51:58,389 --> 02:51:56,880
to your other samples

4380
02:51:59,590 --> 02:51:58,399
yeah that's actually really great i took

4381
02:52:01,990 --> 02:51:59,600
note of that this is something i've been

4382
02:52:03,590 --> 02:52:02,000
thinking about as well which is why i'm

4383
02:52:05,110 --> 02:52:03,600
excited to get more into the host

4384
02:52:07,190 --> 02:52:05,120
linking i'd like to know if there are

4385
02:52:11,190 --> 02:52:07,200
these polyvalent viruses that can affect

4386
02:52:13,190 --> 02:52:11,200
a broad range of microbes uh so far i've

4387
02:52:15,349 --> 02:52:13,200
only done a nucleotide similarity via

4388
02:52:17,590 --> 02:52:15,359

blast obviously some more work with

4389
02:52:20,630 --> 02:52:17,600
crispr and maybe another application to

4390
02:52:22,230 --> 02:52:20,640
really see if i can get in on this

4391
02:52:25,990 --> 02:52:22,240
cool thank you

4392
02:52:31,990 --> 02:52:29,670
there's another comment about doing a

4393
02:52:34,790 --> 02:52:32,000
thought experiment as to how the sips

4394
02:52:37,590 --> 02:52:34,800
might work in astrovirology

4395
02:52:39,830 --> 02:52:37,600
or astrobiology context

4396
02:52:44,230 --> 02:52:39,840
yeah so this is

4397
02:52:47,510 --> 02:52:45,269
i always think of it more of a

4398
02:52:49,030 --> 02:52:47,520
biosignature aspect of when we read one

4399
02:52:50,710 --> 02:52:49,040
how can be different if a virus is

4400
02:52:52,950 --> 02:52:50,720
infecting the host but

4401
02:52:55,190 --> 02:52:52,960
if we're going to take

4402
02:52:57,990 --> 02:52:55,200
stable isotope probing

4403
02:53:00,150 --> 02:52:58,000
to another planet to label something

4404
02:53:01,990 --> 02:53:00,160
i mean that would that would be great in

4405
02:53:04,469 --> 02:53:02,000
in terms of maybe we set we're studying

4406
02:53:07,030 --> 02:53:04,479
a trap to look for life um we take

4407
02:53:09,510 --> 02:53:07,040
something it's labeled we leave it we

4408
02:53:11,670 --> 02:53:09,520
see if anything interacts with it and

4409
02:53:14,309 --> 02:53:11,680
somehow have to control for abiotic

4410
02:53:18,150 --> 02:53:16,230
um

4411
02:53:19,269 --> 02:53:18,160
i don't know if ken or any of anyone

4412
02:53:20,630 --> 02:53:19,279
else has any thoughts on that i wish

4413
02:53:24,389 --> 02:53:20,640

kathy was here right now she would be

4414
02:53:25,269 --> 02:53:24,399
able to uh talk about that for another

4415
02:53:27,510 --> 02:53:25,279
um

4416
02:53:28,710 --> 02:53:27,520
i think it's a really tough call about

4417
02:53:30,870 --> 02:53:28,720
how to do that

4418
02:53:31,670 --> 02:53:30,880
but i think it's worth figuring it out

4419
02:53:33,750 --> 02:53:31,680
yeah

4420
02:53:35,750 --> 02:53:33,760
um and i think it's mostly a tough call

4421
02:53:37,110 --> 02:53:35,760
right now because it hasn't received the

4422
02:53:38,630 --> 02:53:37,120
same attention

4423
02:53:42,469 --> 02:53:38,640
that um

4424
02:53:44,550 --> 02:53:42,479
you know observing bacterial size things

4425
02:53:48,309 --> 02:53:44,560
uh with their properties

4426
02:53:50,389 --> 02:53:48,319
has been uh had a lot of attention so

4427
02:53:53,830 --> 02:53:50,399
you know i think this could be some

4428
02:53:55,910 --> 02:53:53,840
thing to actually try to expand in a

4429
02:53:58,389 --> 02:53:55,920
journal article it's like how could this

4430
02:53:59,910 --> 02:53:58,399
be applied in that astrobiological

4431
02:54:06,309 --> 02:53:59,920
context so i think it's the next

4432
02:54:10,150 --> 02:54:08,150
gary i just have some sort of technical

4433
02:54:12,710 --> 02:54:10,160
questions uh if you remember correctly

4434
02:54:13,990 --> 02:54:12,720
you didn't actually physically separate

4435
02:54:18,469 --> 02:54:14,000
the

4436
02:54:21,269 --> 02:54:18,479
your bacteria this was just a whole

4437
02:54:24,710 --> 02:54:21,279
metagenome correct this is true so sorry

4438
02:54:26,550 --> 02:54:24,720

let me specify that for now because the

4439
02:54:29,269 --> 02:54:26,560
way our pipelines work we have to have a

4440
02:54:30,950 --> 02:54:29,279
large amount of dna in order to have the

4441
02:54:32,630 --> 02:54:30,960
the fractionation for the season card

4442
02:54:36,469 --> 02:54:32,640
density right

4443
02:54:38,790 --> 02:54:36,479
so currently these are sip metagenomes

4444
02:54:40,790 --> 02:54:38,800
i would definitely stay tuned

4445
02:54:48,150 --> 02:54:40,800
for a potential sip firearm in the near

4446
02:54:54,870 --> 02:54:51,910
and sort of a a follow up on that um

4447
02:54:57,269 --> 02:54:54,880
do you actually see bands in the cesium

4448
02:54:59,910 --> 02:54:57,279
or do you just fractionate and pull off

4449
02:55:01,590 --> 02:54:59,920
what's the right size

4450
02:55:02,950 --> 02:55:01,600
so i'm actually going to share the

4451
02:55:04,950 --> 02:55:02,960
slides back with you i had a backup

4452
02:55:07,830 --> 02:55:04,960
slide in case there's a question i make

4453
02:55:10,309 --> 02:55:07,840
it seem easy but it's not easy yeah i i

4454
02:55:13,590 --> 02:55:10,319
i've done cesium too so

4455
02:55:16,950 --> 02:55:14,469
so

4456
02:55:18,750 --> 02:55:16,960
i showed these clear distinct bands but

4457
02:55:20,150 --> 02:55:18,760
that's actually not what you get

4458
02:55:21,349 --> 02:55:20,160
[Music]

4459
02:55:23,349 --> 02:55:21,359
um

4460
02:55:24,469 --> 02:55:23,359
really what it is is that we get many

4461
02:55:25,750 --> 02:55:24,479
bands

4462
02:55:28,389 --> 02:55:25,760
and

4463
02:55:30,070 --> 02:55:28,399

this can be problematic because

4464
02:55:31,990 --> 02:55:30,080
it's all about density so if you have

4465
02:55:34,550 --> 02:55:32,000
microbes that are high gc content

4466
02:55:36,469 --> 02:55:34,560
they're going to be more dense you have

4467
02:55:37,910 --> 02:55:36,479
less gc they're going to be less dense

4468
02:55:39,190 --> 02:55:37,920
and if you put this in viruses too

4469
02:55:41,190 --> 02:55:39,200
because their genomes are small this can

4470
02:55:43,510 --> 02:55:41,200
have a larger impact and then with

4471
02:55:45,269 --> 02:55:43,520
microbes you have such large genomes

4472
02:55:46,710 --> 02:55:45,279
that you can imagine their genome is

4473
02:55:48,230 --> 02:55:46,720
like spread out

4474
02:55:49,190 --> 02:55:48,240
so

4475
02:55:50,469 --> 02:55:49,200
uh

4476
02:55:52,870 --> 02:55:50,479
this is something that we're working on

4477
02:55:55,269 --> 02:55:52,880
to try to fine tune and we take many

4478
02:56:02,150 --> 02:55:55,279
step fractions uh when we do the

4479
02:56:05,910 --> 02:56:03,670
uh now that kathy's back i actually

4480
02:56:07,910 --> 02:56:05,920
wanted to bring that question back up

4481
02:56:10,790 --> 02:56:07,920
there was a question about

4482
02:56:11,590 --> 02:56:10,800
how would we use sip if we were to send

4483
02:56:14,469 --> 02:56:11,600
it

4484
02:56:16,070 --> 02:56:14,479
on in on a rover into space and we were

4485
02:56:17,750 --> 02:56:16,080
looking elsewhere how would we be able

4486
02:56:19,590 --> 02:56:17,760
to use like a stable isotope probing

4487
02:56:21,510 --> 02:56:19,600
technique where something gets labeled

4488
02:56:23,750 --> 02:56:21,520

and we can track it

4489
02:56:24,950 --> 02:56:23,760
well um i don't know if anyone's

4490
02:56:26,469 --> 02:56:24,960
mentioned

4491
02:56:28,150 --> 02:56:26,479
viking and the labeled release

4492
02:56:29,750 --> 02:56:28,160
experiments cassie

4493
02:56:31,110 --> 02:56:29,760
is going to mention that pass it out

4494
02:56:33,429 --> 02:56:31,120
just mention that

4495
02:56:34,950 --> 02:56:33,439
although she didn't elaborate on it

4496
02:56:37,110 --> 02:56:34,960
obviously

4497
02:56:40,389 --> 02:56:37,120
i nailed it

4498
02:56:42,469 --> 02:56:40,399
one uh so one of my favorite

4499
02:56:44,870 --> 02:56:42,479
uh missions was viking and the labeled

4500
02:56:47,590 --> 02:56:44,880
release experiment because

4501
02:56:49,269 --> 02:56:47,600
uh it was really a life detection

4502
02:56:51,990 --> 02:56:49,279
mission like that's what it was there to

4503
02:56:54,550 --> 02:56:52,000
do was to detect life um you know

4504
02:56:56,710 --> 02:56:54,560
inherently it made a lot of assumptions

4505
02:56:59,110 --> 02:56:56,720
of if there was life in the martian

4506
02:57:02,070 --> 02:56:59,120
regolith the nutrient soup in which it

4507
02:57:03,910 --> 02:57:02,080
would need to be metabolically active

4508
02:57:04,790 --> 02:57:03,920
you know just all these assumptions

4509
02:57:06,630 --> 02:57:04,800
about

4510
02:57:09,349 --> 02:57:06,640
uh life they're having having these

4511
02:57:12,230 --> 02:57:09,359
sorts of requirements so

4512
02:57:14,710 --> 02:57:12,240
i think with a live with a labeled

4513
02:57:16,870 --> 02:57:14,720

release sort of experiment uh you have

4514
02:57:19,269 --> 02:57:16,880
to be very careful not to assume things

4515
02:57:21,910 --> 02:57:19,279
about metabolic activity

4516
02:57:23,510 --> 02:57:21,920
uh so that is that is sort of the crux

4517
02:57:24,550 --> 02:57:23,520
of that problem and i don't know if i

4518
02:57:25,830 --> 02:57:24,560
have a

4519
02:57:27,510 --> 02:57:25,840
a good answer

4520
02:57:29,269 --> 02:57:27,520
but i can i will definitely think about

4521
02:57:41,429 --> 02:57:29,279
it

4522
02:57:44,150 --> 02:57:42,790
and do you want to comment on cave

4523
02:57:47,510 --> 02:57:44,160
viruses penny

4524
02:57:49,110 --> 02:57:47,520
uh yeah i did i did online but um i

4525
02:57:51,349 --> 02:57:49,120
pointed out that there's less than a

4526
02:57:54,309 --> 02:57:51,359
handful of papers in english

4527
02:57:57,510 --> 02:57:54,319
in the literature on the subject

4528
02:57:59,030 --> 02:57:57,520
we do know from uh being in the field at

4529
02:58:01,670 --> 02:57:59,040
the um

4530
02:58:05,429 --> 02:58:01,680
nica caves that are very very um hot and

4531
02:58:07,670 --> 02:58:05,439
have no natural openings uh from uh just

4532
02:58:10,230 --> 02:58:07,680
some very preliminary work that curtis

4533
02:58:13,190 --> 02:58:10,240
suttle and his group did uh that there's

4534
02:58:15,269 --> 02:58:13,200
a a huge load of viruses

4535
02:58:17,269 --> 02:58:15,279
and so that was something that we have

4536
02:58:19,990 --> 02:58:17,279
used to actually

4537
02:58:21,670 --> 02:58:20,000
indicate that the extraordinary

4538
02:58:24,070 --> 02:58:21,680

microbial communities that we're finding

4539
02:58:26,630 --> 02:58:24,080
there are indigenous because where there

4540
02:58:28,230 --> 02:58:26,640
are bugs there are tiny bugs

4541
02:58:31,349 --> 02:58:28,240
that are eating them

4542
02:58:34,790 --> 02:58:31,359
bugging them or something and so

4543
02:58:37,750 --> 02:58:34,800
but no actual work on the identification

4544
02:58:39,670 --> 02:58:37,760
of those was done in that case um i

4545
02:58:40,870 --> 02:58:39,680
think that dale griffin has published

4546
02:58:44,070 --> 02:58:40,880
maybe

4547
02:58:47,030 --> 02:58:44,080
one to two papers in terms of uh cave

4548
02:58:49,110 --> 02:58:47,040
virus stuff this was work done with my

4549
02:58:51,510 --> 02:58:49,120
friend and colleague diana northam at

4550
02:58:54,150 --> 02:58:51,520
the university of new mexico

4551
02:58:55,830 --> 02:58:54,160
quite a few years ago now and

4552
02:58:58,230 --> 02:58:55,840
part of the reason

4553
02:59:01,190 --> 02:58:58,240
as i mentioned on the first day that

4554
02:59:04,550 --> 02:59:01,200
barry was interested in a um you know a

4555
02:59:07,429 --> 02:59:04,560
cave field trip for virologists was to

4556
02:59:10,950 --> 02:59:07,439
actually stimulate interest in the field

4557
02:59:13,750 --> 02:59:10,960
and if people are interested in

4558
02:59:15,269 --> 02:59:13,760
samples from caves i go into many many

4559
02:59:16,950 --> 02:59:15,279
of them so

4560
02:59:21,670 --> 02:59:16,960
i think that one of the

4561
02:59:24,550 --> 02:59:21,680
selling points for cave samples is that

4562
02:59:27,190 --> 02:59:24,560
at least if it reflects in the virology

4563
02:59:30,150 --> 02:59:27,200

what we see in the bacteriology and the

4564
02:59:32,469 --> 02:59:30,160
studies of archaea that because

4565
02:59:34,790 --> 02:59:32,479
it's a geological environment that is

4566
02:59:37,510 --> 02:59:34,800
highly partitioned physically

4567
02:59:39,990 --> 02:59:37,520
with limited modes of transmission by

4568
02:59:40,870 --> 02:59:40,000
way of air currents and fluids

4569
02:59:42,790 --> 02:59:40,880
that

4570
02:59:44,950 --> 02:59:42,800
we see tremendous

4571
02:59:46,469 --> 02:59:44,960
individual

4572
02:59:47,910 --> 02:59:46,479
evolutionary

4573
02:59:50,550 --> 02:59:47,920
experiments

4574
02:59:52,710 --> 02:59:50,560
and so the diversity is staggeringly

4575
02:59:55,349 --> 02:59:52,720
huge amongst the

4576
02:59:57,750 --> 02:59:55,359
bacterial and our keel populations and

4577
03:00:00,469 --> 02:59:57,760
it's because we believe of this you know

4578
03:00:01,670 --> 03:00:00,479
partitioning and so i would guess that

4579
03:00:03,190 --> 03:00:01,680
we would see

4580
03:00:05,750 --> 03:00:03,200
you know an equal

4581
03:00:07,590 --> 03:00:05,760
stunning diversity not that viruses

4582
03:00:09,910 --> 03:00:07,600
aren't stunningly diverse enough as it

4583
03:00:13,670 --> 03:00:09,920
is uh but you know that that would

4584
03:00:15,990 --> 03:00:13,680
reflect that same uh isolation so in a

4585
03:00:18,950 --> 03:00:16,000
way it's sort of uh an underground

4586
03:00:21,429 --> 03:00:18,960
island biogeography uh situation where

4587
03:00:23,670 --> 03:00:21,439
you've got all these isolated uh

4588
03:00:25,429 --> 03:00:23,680

geological habitats

4589
03:00:27,910 --> 03:00:25,439
so if anybody is interested in cave

4590
03:00:30,070 --> 03:00:27,920
samples uh and wants to actually do some

4591
03:00:31,429 --> 03:00:30,080
serious work on them uh you know where i

4592
03:00:33,990 --> 03:00:31,439
can be found on

4593
03:00:36,150 --> 03:00:34,000
email i bet

4594
03:00:40,070 --> 03:00:36,160
can you repeat that paper again

4595
03:00:42,469 --> 03:00:40,080
which one um it was the k virus paper

4596
03:00:43,590 --> 03:00:42,479
was there actual paper yes yes uh dale

4597
03:00:46,230 --> 03:00:43,600
griffith

4598
03:00:48,070 --> 03:00:46,240
was or is it griffin dale

4599
03:00:50,389 --> 03:00:48,080
he's at um

4600
03:00:53,590 --> 03:00:50,399
usgs in florida

4601
03:00:55,670 --> 03:00:53,600
is it griffin i think it's griffin i

4602
03:00:57,910 --> 03:00:55,680
yeah i think it's griffin okay so dale

4603
03:00:59,750 --> 03:00:57,920
but anyway he's at the usgs in saint

4604
03:01:03,269 --> 03:00:59,760
petersburg florida

4605
03:01:05,429 --> 03:01:03,279
and um so he did some work with diana

4606
03:01:07,830 --> 03:01:05,439
looking at the viruses

4607
03:01:09,190 --> 03:01:07,840
in parts i believe it was carlsbad

4608
03:01:10,469 --> 03:01:09,200
caverns

4609
03:01:11,750 --> 03:01:10,479
and

4610
03:01:14,790 --> 03:01:11,760
he also

4611
03:01:16,870 --> 03:01:14,800
has done other creative things about the

4612
03:01:20,150 --> 03:01:16,880
you know the microbial load coming

4613
03:01:23,510 --> 03:01:20,160

across on dust across the atlantic ocean

4614
03:01:25,670 --> 03:01:23,520
from the deserts of africa and stuff and

4615
03:01:27,990 --> 03:01:25,680
a lot of that material gets deposited

4616
03:01:29,349 --> 03:01:28,000
into caves is this material called terra

4617
03:01:30,309 --> 03:01:29,359
rosa

4618
03:01:32,389 --> 03:01:30,319
and so

4619
03:01:34,630 --> 03:01:32,399
that obviously carries all different

4620
03:01:36,790 --> 03:01:34,640
kinds of life forms undoubtedly there

4621
03:01:37,990 --> 03:01:36,800
are viruses coming along with that as

4622
03:01:39,030 --> 03:01:38,000
well so

4623
03:01:40,710 --> 03:01:39,040
um

4624
03:01:41,990 --> 03:01:40,720
maybe one of the things that i can do as

4625
03:01:45,429 --> 03:01:42,000
a follow-up

4626
03:01:48,070 --> 03:01:45,439
to this is assemble the feeble amount of

4627
03:01:50,630 --> 03:01:48,080
literature that there exists in terms of

4628
03:01:53,030 --> 03:01:50,640
subsurface viruses in that way and make

4629
03:01:55,590 --> 03:01:53,040
that available to to the group

4630
03:01:58,070 --> 03:01:55,600
uh that will be a very short task

4631
03:01:59,910 --> 03:01:58,080
yeah yeah unfortunately

4632
03:02:03,429 --> 03:01:59,920
i think i believe they may have found it

4633
03:02:07,429 --> 03:02:04,950
yeah

4634
03:02:09,510 --> 03:02:07,439
okay good thank you i don't know that's

4635
03:02:11,590 --> 03:02:09,520
human visitation so maybe it's not oh

4636
03:02:14,230 --> 03:02:11,600
yeah i think that's the different one

4637
03:02:16,550 --> 03:02:14,240
but yeah there are several

4638
03:02:19,110 --> 03:02:16,560

cool great

4639
03:02:21,830 --> 03:02:19,120
and um laura g says she's actually

4640
03:02:25,030 --> 03:02:21,840
beginning with some cave microbiology

4641
03:02:27,110 --> 03:02:25,040
research in mexico or at least trying so

4642
03:02:29,429 --> 03:02:27,120
go laura

4643
03:02:30,550 --> 03:02:29,439
or maybe if you're mexican it's lara

4644
03:02:31,750 --> 03:02:30,560
okay

4645
03:02:33,830 --> 03:02:31,760
i don't know

4646
03:02:39,990 --> 03:02:33,840
but you know email me if you want to

4647
03:02:45,030 --> 03:02:42,630
okay so maybe we're ready for you ken i

4648
03:02:47,590 --> 03:02:45,040
think i'm up yeah i just need to move my

4649
03:02:50,550 --> 03:02:47,600
window here sorry about this it's a

4650
03:02:51,990 --> 03:02:50,560
little confusing yeah it's always always

4651
03:02:53,990 --> 03:02:52,000
a problem to have you know three

4652
03:02:58,790 --> 03:02:54,000
monitors so that's just still one of

4653
03:03:03,750 --> 03:03:01,269
okay here we go

4654
03:03:05,349 --> 03:03:04,010
go

4655
03:03:07,510 --> 03:03:05,359
[Music]

4656
03:03:07,940 --> 03:03:07,520
and

4657
03:03:10,309 --> 03:03:07,950
share

4658
03:03:12,070 --> 03:03:10,319
[Music]

4659
03:03:17,190 --> 03:03:12,080
which one shall i share let's try

4660
03:03:22,710 --> 03:03:19,830
okay if you got my

4661
03:03:23,830 --> 03:03:22,720
got my slides there

4662
03:03:24,870 --> 03:03:23,840
good

4663
03:03:26,790 --> 03:03:24,880

okay

4664
03:03:28,309 --> 03:03:26,800
i don't see anybody you know waving or

4665
03:03:29,349 --> 03:03:28,319
complaining here

4666
03:03:31,590 --> 03:03:29,359
so

4667
03:03:33,670 --> 03:03:31,600
i just wanted to finish up um today

4668
03:03:35,429 --> 03:03:33,680
again thanks at the very first

4669
03:03:38,309 --> 03:03:35,439
absolutely everyone for participating

4670
03:03:40,230 --> 03:03:38,319
this has been absolutely fabulous um and

4671
03:03:43,429 --> 03:03:40,240
i think we're gonna get some really nice

4672
03:03:46,309 --> 03:03:43,439
products out of it um eventually as well

4673
03:03:47,830 --> 03:03:46,319
um the background picture here is a

4674
03:03:49,990 --> 03:03:47,840
place called boiling springs lake which

4675
03:03:51,269 --> 03:03:50,000
is one of our rain field sites probably

4676
03:03:53,590 --> 03:03:51,279
a little easier to get to than some of

4677
03:03:55,269 --> 03:03:53,600
the caves um but nonetheless i think a

4678
03:03:56,150 --> 03:03:55,279
really pretty fascinating place to look

4679
03:04:00,550 --> 03:03:56,160
at

4680
03:04:03,590 --> 03:04:00,560
change my slides here there we go um is

4681
03:04:05,510 --> 03:04:03,600
talk a little bit about some virus

4682
03:04:06,469 --> 03:04:05,520
definitions and we got some of this

4683
03:04:08,550 --> 03:04:06,479
already

4684
03:04:10,389 --> 03:04:08,560
with evelyn right at the beginning but

4685
03:04:11,750 --> 03:04:10,399
just a bit of a way to sort of try and

4686
03:04:13,349 --> 03:04:11,760
bring people again back together and

4687
03:04:16,150 --> 03:04:13,359
some people may have missed that talk at

4688
03:04:19,429 --> 03:04:16,160

the beginning so classical definitions

4689
03:04:21,910 --> 03:04:19,439
of viruses were you know i love this one

4690
03:04:24,309 --> 03:04:21,920
simply a piece of bad news wrapped up in

4691
03:04:26,550 --> 03:04:24,319
a protein um cerpedometer

4692
03:04:28,790 --> 03:04:26,560
and that's how the vast majority of

4693
03:04:31,110 --> 03:04:28,800
people think about viruses and thanks to

4694
03:04:32,630 --> 03:04:31,120
nigel for pointing out that you know

4695
03:04:34,710 --> 03:04:32,640
most viruses actually have a really

4696
03:04:37,190 --> 03:04:34,720
pretty bad rap and that's a lot of what

4697
03:04:38,630 --> 03:04:37,200
i do particularly in my public outreach

4698
03:04:41,190 --> 03:04:38,640
is that you know try and explain that

4699
03:04:42,870 --> 03:04:41,200
they do have a very bad rap and so if

4700
03:04:44,630 --> 03:04:42,880
you ever go into the textbooks you see

4701
03:04:46,309 --> 03:04:44,640
this you know they're very small

4702
03:04:47,750 --> 03:04:46,319
infectious obligate anticellular

4703
03:04:49,590 --> 03:04:47,760
parasites

4704
03:04:51,750 --> 03:04:49,600
again evelyn did a great job i'm

4705
03:04:53,670 --> 03:04:51,760
pointing out that very small

4706
03:04:55,750 --> 03:04:53,680
all my students know if i put something

4707
03:04:57,349 --> 03:04:55,760
in quotes and in red i don't believe it

4708
03:04:59,269 --> 03:04:57,359
um and so that gives you a bit of an

4709
03:04:59,990 --> 03:04:59,279
indication what i thought about it and

4710
03:05:02,070 --> 03:05:00,000
then

4711
03:05:03,990 --> 03:05:02,080
one of my favorite definitions actually

4712
03:05:06,469 --> 03:05:04,000
a lot like the one that penny just used

4713
03:05:09,750 --> 03:05:06,479

a virus is a bag of nucleic acid

4714
03:05:11,990 --> 03:05:09,760
um it's a very specialized bag and it's

4715
03:05:15,910 --> 03:05:12,000
usually very specialized nucleic acid

4716
03:05:16,870 --> 03:05:15,920
but that's basically it

4717
03:05:18,790 --> 03:05:16,880
but

4718
03:05:21,030 --> 03:05:18,800
i don't like really any of these

4719
03:05:22,550 --> 03:05:21,040
particular definitions the one that i

4720
03:05:24,150 --> 03:05:22,560
actually like the best has been going

4721
03:05:27,990 --> 03:05:24,160
back into the literature

4722
03:05:30,950 --> 03:05:28,000
and salvador luria had this

4723
03:05:33,269 --> 03:05:30,960
definition in 1978 which i really really

4724
03:05:34,230 --> 03:05:33,279
like and we'll get back to in just a

4725
03:05:37,110 --> 03:05:34,240
second

4726
03:05:39,429 --> 03:05:37,120
what i really like about this but

4727
03:05:40,630 --> 03:05:39,439
viruses are entities you know not saying

4728
03:05:43,110 --> 03:05:40,640
anything particularly about them whose

4729
03:05:45,910 --> 03:05:43,120
genomes are elements of nucleic acid

4730
03:05:47,910 --> 03:05:45,920
that replicate inside living cells using

4731
03:05:49,830 --> 03:05:47,920
the cellular synthetic machinery and

4732
03:05:51,830 --> 03:05:49,840
causing the synthesis of specialized

4733
03:05:54,309 --> 03:05:51,840
elements that can transfer the viral

4734
03:05:56,630 --> 03:05:54,319
genome to other cells and so that fits

4735
03:05:58,950 --> 03:05:56,640
nicely with this you know overall image

4736
03:06:01,349 --> 03:05:58,960
here this is a lytic image and you can

4737
03:06:02,790 --> 03:06:01,359
look at the lysogenic images

4738
03:06:04,469 --> 03:06:02,800

you have an

4739
03:06:06,389 --> 03:06:04,479
extracellular

4740
03:06:09,510 --> 03:06:06,399
specialized element

4741
03:06:11,990 --> 03:06:09,520
that interacts with a living cell then

4742
03:06:14,469 --> 03:06:12,000
there's the nucleic acid that's inside

4743
03:06:15,429 --> 03:06:14,479
this package that gets put inside the

4744
03:06:18,150 --> 03:06:15,439
cell

4745
03:06:19,190 --> 03:06:18,160
that gets replicated and made into phage

4746
03:06:22,710 --> 03:06:19,200
protein

4747
03:06:24,870 --> 03:06:22,720
by cellular components and so again the

4748
03:06:27,349 --> 03:06:24,880
key here is it's a cellular synthetic

4749
03:06:29,349 --> 03:06:27,359
machinery and one particular piece of

4750
03:06:32,389 --> 03:06:29,359
that machinery which is absolutely

4751
03:06:34,230 --> 03:06:32,399
critical to all viruses that we know of

4752
03:06:37,190 --> 03:06:34,240
is the ribosome

4753
03:06:39,510 --> 03:06:37,200
no viruses that have been found to date

4754
03:06:41,429 --> 03:06:39,520
have ribosomes

4755
03:06:43,030 --> 03:06:41,439
in their genomes

4756
03:06:44,790 --> 03:06:43,040
they have things that modify ribosomes

4757
03:06:47,030 --> 03:06:44,800
and modify translation but not actually

4758
03:06:49,190 --> 03:06:47,040
have ribosomes and so that i think is a

4759
03:06:50,790 --> 03:06:49,200
really nice way of defining the

4760
03:06:53,030 --> 03:06:50,800
difference between

4761
03:06:54,790 --> 03:06:53,040
cellular life and

4762
03:06:57,269 --> 03:06:54,800
viral life

4763
03:06:59,990 --> 03:06:57,279

then you get the assembly of these

4764
03:07:01,510 --> 03:07:00,000
particles then these particles are

4765
03:07:03,510 --> 03:07:01,520
released and you go through the whole

4766
03:07:05,190 --> 03:07:03,520
cycle again you could also of course

4767
03:07:06,150 --> 03:07:05,200
have a replication of these infected

4768
03:07:09,030 --> 03:07:06,160
cells

4769
03:07:10,710 --> 03:07:09,040
where things will continue along so

4770
03:07:12,309 --> 03:07:10,720
that's your again i really like this

4771
03:07:14,870 --> 03:07:12,319
definition i think it's a very useful

4772
03:07:16,710 --> 03:07:14,880
definition and we can talk more about it

4773
03:07:18,710 --> 03:07:16,720
later in terms of

4774
03:07:21,590 --> 03:07:18,720
whether people particularly like it but

4775
03:07:25,429 --> 03:07:21,600
before i go there i wanted to mention um

4776
03:07:28,070 --> 03:07:25,439
a particular piece of art um which is i

4777
03:07:30,950 --> 03:07:28,080
think very relevant to thinking about

4778
03:07:33,830 --> 03:07:30,960
viruses and particularly in the

4779
03:07:36,870 --> 03:07:33,840
discussion of you know whether viruses

4780
03:07:39,190 --> 03:07:36,880
are alive or not um this of course is

4781
03:07:42,309 --> 03:07:39,200
the famous surrealist painting by lenny

4782
03:07:46,309 --> 03:07:42,319
magritte yeah cecine peep but as patrick

4783
03:07:47,110 --> 03:07:46,319
forte always loves to say cecine villas

4784
03:07:48,870 --> 03:07:47,120
and

4785
03:07:51,110 --> 03:07:48,880
this particular

4786
03:07:54,389 --> 03:07:51,120
virion which is that

4787
03:07:55,590 --> 03:07:54,399
entity which is built by viruses

4788
03:07:57,750 --> 03:07:55,600

is

4789
03:07:59,670 --> 03:07:57,760
very different and very unique and very

4790
03:08:02,790 --> 03:07:59,680
specific to

4791
03:08:06,469 --> 03:08:02,800
viruses but it's not the whole story the

4792
03:08:09,110 --> 03:08:06,479
whole story is really much more about

4793
03:08:11,190 --> 03:08:09,120
the virus life cycle and this gets us

4794
03:08:14,550 --> 03:08:11,200
back to joshua vice's presentation

4795
03:08:18,630 --> 03:08:14,560
yesterday was really talking about

4796
03:08:21,429 --> 03:08:18,640
viruses and cells and how those viruses

4797
03:08:23,030 --> 03:08:21,439
and cells come together and once a

4798
03:08:25,670 --> 03:08:23,040
cell has been

4799
03:08:27,830 --> 03:08:25,680
infected by a virus

4800
03:08:29,510 --> 03:08:27,840
that cell and particularly in a lytic

4801
03:08:31,349 --> 03:08:29,520
case but even if it's replicating in a

4802
03:08:34,309 --> 03:08:31,359
lysogenic way

4803
03:08:36,790 --> 03:08:34,319
that's also really a virus and so

4804
03:08:38,950 --> 03:08:36,800
patrick forte and dd avalor came up with

4805
03:08:42,150 --> 03:08:38,960
this concept which they called

4806
03:08:45,110 --> 03:08:42,160
the ribose cell again because these have

4807
03:08:47,830 --> 03:08:45,120
ribosomes and all viruses need cellular

4808
03:08:50,469 --> 03:08:47,840
ribosomes to replicate and the virus

4809
03:08:52,469 --> 03:08:50,479
cell which is where you have virus

4810
03:08:53,990 --> 03:08:52,479
replication so again i love joshua

4811
03:08:56,070 --> 03:08:54,000
weitz's presentation thinking about

4812
03:08:59,269 --> 03:08:56,080
modeling because i think you really need

4813
03:09:01,110 --> 03:08:59,279

to think about viruses in a much more

4814
03:09:03,910 --> 03:09:01,120
general case

4815
03:09:06,309 --> 03:09:03,920
rather than in this sort of we're just

4816
03:09:07,910 --> 03:09:06,319
going to look at these extracellular

4817
03:09:10,150 --> 03:09:07,920
particles and the extracellular

4818
03:09:11,510 --> 03:09:10,160
particles of course are the virions

4819
03:09:14,150 --> 03:09:11,520
which we've heard about quite a bit

4820
03:09:17,349 --> 03:09:14,160
before and these are things that are

4821
03:09:20,710 --> 03:09:17,359
really really specific to

4822
03:09:22,229 --> 03:09:20,720
viruses only viruses make these things

4823
03:09:24,389 --> 03:09:22,239
and they have these incredibly

4824
03:09:26,790 --> 03:09:24,399
distinctive morphologies you know

4825
03:09:29,190 --> 03:09:26,800
kathy was just talking about these ways

4826
03:09:31,110 --> 03:09:29,200
that you can detect them nicely with

4827
03:09:33,830 --> 03:09:31,120
nanopores um particularly the tomato

4828
03:09:36,070 --> 03:09:33,840
mosaic virus this long

4829
03:09:38,469 --> 03:09:36,080
stiff rod shaped form

4830
03:09:40,309 --> 03:09:38,479
many of the bacterial viruses or this is

4831
03:09:42,790 --> 03:09:40,319
just an environmental sample here which

4832
03:09:46,630 --> 03:09:42,800
have these heads and tail structures and

4833
03:09:49,030 --> 03:09:46,640
very often with icosahedral symmetry and

4834
03:09:51,269 --> 03:09:49,040
this kind of symmetry again these are

4835
03:09:53,590 --> 03:09:51,279
small dots that evelyn was talking about

4836
03:09:56,389 --> 03:09:53,600
before i think this is really indicative

4837
03:09:59,110 --> 03:09:56,399
of what a lot of people think about the

4838
03:10:01,429 --> 03:09:59,120

viruses now i she alluded to this as

4839
03:10:03,670 --> 03:10:01,439
well but my favorite kinds of viruses

4840
03:10:06,070 --> 03:10:03,680
are the really really weird virion

4841
03:10:09,349 --> 03:10:06,080
containing ones and so these are a

4842
03:10:11,429 --> 03:10:09,359
sampling of some of the archaea viruses

4843
03:10:13,830 --> 03:10:11,439
evelyn mentioned this acidionis

4844
03:10:15,750 --> 03:10:13,840
filamentous virus with these nanoclaws

4845
03:10:17,990 --> 03:10:15,760
at the end i'll talk a little bit more

4846
03:10:19,429 --> 03:10:18,000
about our favorite viruses these ssd

4847
03:10:21,990 --> 03:10:19,439
ones that i've been looking at before

4848
03:10:24,550 --> 03:10:22,000
have this spindle or like lemon shape

4849
03:10:26,150 --> 03:10:24,560
there's an individual virion right here

4850
03:10:27,990 --> 03:10:26,160
and they're probably the most amazing of

4851
03:10:30,070 --> 03:10:28,000
these were some virions discovered by

4852
03:10:32,790 --> 03:10:30,080
david prangeshville's group um when he

4853
03:10:35,990 --> 03:10:32,800
was in regensburg where they have these

4854
03:10:38,309 --> 03:10:36,000
um appropriately named acidionis

4855
03:10:40,309 --> 03:10:38,319
bottle-shaped virus and these really do

4856
03:10:42,469 --> 03:10:40,319
look kind of like champagne bottles that

4857
03:10:45,190 --> 03:10:42,479
have you served a birthday cake and put

4858
03:10:47,269 --> 03:10:45,200
the candles at one end how the heck you

4859
03:10:49,830 --> 03:10:47,279
make one of these structures and what it

4860
03:10:52,309 --> 03:10:49,840
would look like going through a nanopore

4861
03:10:54,030 --> 03:10:52,319
i would just love to know but i think

4862
03:10:56,150 --> 03:10:54,040
it's really important from an

4863
03:10:59,190 --> 03:10:56,160

astrovirology point of view

4864
03:11:02,630 --> 03:10:59,200
is to not get fixated on

4865
03:11:04,630 --> 03:11:02,640
head and tail virion morphologies but

4866
03:11:07,190 --> 03:11:04,640
really start thinking about a lot of

4867
03:11:09,349 --> 03:11:07,200
these other ones and the nice example of

4868
03:11:10,389 --> 03:11:09,359
that and again evelyn brought this up

4869
03:11:13,590 --> 03:11:10,399
are the

4870
03:11:15,429 --> 03:11:13,600
gyruses the giant viruses and the first

4871
03:11:17,670 --> 03:11:15,439
time that i saw this picture it just

4872
03:11:20,469 --> 03:11:17,680
completely blew me away each of the

4873
03:11:23,190 --> 03:11:20,479
little dots appear in the corner each of

4874
03:11:25,590 --> 03:11:23,200
these is a virion and this is light

4875
03:11:27,030 --> 03:11:25,600
microscopy so light microscopy you can

4876
03:11:28,950 --> 03:11:27,040
actually see

4877
03:11:30,469 --> 03:11:28,960
the individual virions and you know

4878
03:11:31,590 --> 03:11:30,479
there's an electron micrograph down here

4879
03:11:33,670 --> 03:11:31,600
of one of these

4880
03:11:37,349 --> 03:11:33,680
mimi viruses as well so

4881
03:11:39,910 --> 03:11:37,359
virions are definitely not just small

4882
03:11:42,150 --> 03:11:39,920
they're not just head and tail they're

4883
03:11:44,309 --> 03:11:42,160
all kinds of different things and even

4884
03:11:47,750 --> 03:11:44,319
there are some viruses

4885
03:11:49,349 --> 03:11:47,760
which have viruses that infect them and

4886
03:11:52,070 --> 03:11:49,359
this is the first case of the so-called

4887
03:11:54,229 --> 03:11:52,080
virophage the sputniks

4888
03:11:55,590 --> 03:11:54,239

which infect the mimi viruses another

4889
03:11:57,830 --> 03:11:55,600
thing to mention about these giant

4890
03:11:58,950 --> 03:11:57,840
viruses is that they have massive

4891
03:12:01,030 --> 03:11:58,960
genomes

4892
03:12:03,269 --> 03:12:01,040
mimi virus itself has about over a

4893
03:12:05,030 --> 03:12:03,279
million base pair genome

4894
03:12:07,830 --> 03:12:05,040
and some of the larger viruses since

4895
03:12:09,670 --> 03:12:07,840
then have gotten to you know almost tens

4896
03:12:13,190 --> 03:12:09,680
of millions of base pair size genomes

4897
03:12:15,429 --> 03:12:13,200
well larger than any bacteria or not any

4898
03:12:17,590 --> 03:12:15,439
bacteria but many bacteria

4899
03:12:20,550 --> 03:12:17,600
but nonetheless none of them to date

4900
03:12:21,990 --> 03:12:20,560
have any of these ribosomal proteins in

4901
03:12:23,750 --> 03:12:22,000
them so

4902
03:12:25,990 --> 03:12:23,760
massive amounts of diversity we could

4903
03:12:28,950 --> 03:12:26,000
also from evelyn that we've got

4904
03:12:30,469 --> 03:12:28,960
these um very different kinds of genomes

4905
03:12:31,269 --> 03:12:30,479
that they could have

4906
03:12:33,830 --> 03:12:31,279
but

4907
03:12:35,990 --> 03:12:33,840
we also heard from her that they're

4908
03:12:37,670 --> 03:12:36,000
incredibly common and we also saw

4909
03:12:39,190 --> 03:12:37,680
exactly i think exactly the same picture

4910
03:12:42,150 --> 03:12:39,200
turned around a little bit in seymour's

4911
03:12:45,030 --> 03:12:42,160
talk um if you look at seawater samples

4912
03:12:47,750 --> 03:12:45,040
lots and lots of little virions here one

4913
03:12:50,469 --> 03:12:47,760

portable eukaryotic diatom and a few

4914
03:12:52,309 --> 03:12:50,479
overwhelmed bacteria and archaea and i

4915
03:12:54,870 --> 03:12:52,319
love this juxtaposition i forget who

4916
03:12:57,510 --> 03:12:54,880
first showed it to me of the hubble um

4917
03:12:59,349 --> 03:12:57,520
deep field because i think these look

4918
03:13:01,750 --> 03:12:59,359
incredibly similar to each other but we

4919
03:13:03,590 --> 03:13:01,760
may actually know more about these dots

4920
03:13:04,309 --> 03:13:03,600
than we know about these dots here as

4921
03:13:06,630 --> 03:13:04,319
well

4922
03:13:08,950 --> 03:13:06,640
so one of the things that you can do

4923
03:13:10,309 --> 03:13:08,960
with samples like this and as simo

4924
03:13:12,389 --> 03:13:10,319
mentioned these there's a lot of

4925
03:13:14,550 --> 03:13:12,399
uncertainty in many of these numbers you

4926
03:13:16,710 --> 03:13:14,560
can do some calculations and you get to

4927
03:13:19,269 --> 03:13:16,720
ridiculous numbers of virions on our

4928
03:13:21,510 --> 03:13:19,279
planet and roger hendricks kind of took

4929
03:13:23,830 --> 03:13:21,520
some of these to heart and mentioned

4930
03:13:26,469 --> 03:13:23,840
that this is what the earth looks like

4931
03:13:29,349 --> 03:13:26,479
with 10 to the 31

4932
03:13:31,190 --> 03:13:29,359
virions most of them bacteriophages

4933
03:13:32,710 --> 03:13:31,200
and jbs haldane said the creator must

4934
03:13:33,429 --> 03:13:32,720
have had inordinate fondness for beetles

4935
03:13:35,349 --> 03:13:33,439
well

4936
03:13:37,590 --> 03:13:35,359
if the creator really had as much fonus

4937
03:13:40,630 --> 03:13:37,600
for beetles as he did for phage the

4938
03:13:43,269 --> 03:13:40,640

earth would be this size so clearly the

4939
03:13:44,710 --> 03:13:43,279
creator was much more fond of phage as

4940
03:13:47,910 --> 03:13:44,720
we've heard from many people already in

4941
03:13:49,510 --> 03:13:47,920
the presentations um than she ever was

4942
03:13:50,710 --> 03:13:49,520
um with beatles

4943
03:13:53,190 --> 03:13:50,720
the other thing you can do of course is

4944
03:13:55,990 --> 03:13:53,200
line all of these phage particles up end

4945
03:13:58,389 --> 03:13:56,000
to end and they end up depending again

4946
03:13:59,910 --> 03:13:58,399
on lots of orders of magnitude of 210 to

4947
03:14:02,630 --> 03:13:59,920
the seventh light years if you're then

4948
03:14:04,630 --> 03:14:02,640
these are just the baryons on earth so

4949
03:14:06,309 --> 03:14:04,640
um that to me is sort of as as

4950
03:14:09,510 --> 03:14:06,319
astrobiological

4951
03:14:11,269 --> 03:14:09,520
as this could possibly get um but that

4952
03:14:15,349 --> 03:14:11,279
being said one of the things about these

4953
03:14:17,830 --> 03:14:15,359
very very large numbers of

4954
03:14:19,750 --> 03:14:17,840
viruses and their whole process again

4955
03:14:21,670 --> 03:14:19,760
remember this is not just virions but

4956
03:14:24,550 --> 03:14:21,680
viruses on our planet means that they

4957
03:14:26,269 --> 03:14:24,560
actually do play very important roles in

4958
03:14:28,389 --> 03:14:26,279
some big

4959
03:14:30,229 --> 03:14:28,399
biogeochemical cycles we heard a little

4960
03:14:33,190 --> 03:14:30,239
bit about that earlier but i love this

4961
03:14:34,710 --> 03:14:33,200
image admittedly almost 15 years ago a

4962
03:14:36,070 --> 03:14:34,720
really nice review article by curtis

4963
03:14:39,429 --> 03:14:36,080

suttle in nature

4964
03:14:43,750 --> 03:14:39,439
just looking at what happens with the

4965
03:14:45,510 --> 03:14:43,760
presence of viruses in oceans and what

4966
03:14:47,349 --> 03:14:45,520
they do and again we've heard about this

4967
03:14:50,550 --> 03:14:47,359
in previous talks as well

4968
03:14:53,590 --> 03:14:50,560
is they recycle a lot of the

4969
03:14:56,150 --> 03:14:53,600
dissolved organic carbon and keep it

4970
03:14:59,510 --> 03:14:56,160
actually in the upper ocean as opposed

4971
03:15:01,510 --> 03:14:59,520
to having a lot of that disappear and

4972
03:15:04,070 --> 03:15:01,520
this carbon actually sinking to the

4973
03:15:07,110 --> 03:15:04,080
bottom of the ocean so the presence of

4974
03:15:09,349 --> 03:15:07,120
viruses is really important in some

4975
03:15:11,429 --> 03:15:09,359
real on earth you know global

4976
03:15:13,429 --> 03:15:11,439
biogeochemical cycles and i think that

4977
03:15:15,510 --> 03:15:13,439
one of the things and we mentioned this

4978
03:15:18,469 --> 03:15:15,520
in our astrovirology article

4979
03:15:20,790 --> 03:15:18,479
that people should be thinking about is

4980
03:15:23,670 --> 03:15:20,800
thinking about these kinds of biogenic

4981
03:15:26,309 --> 03:15:23,680
chemical cycles on an early earth in

4982
03:15:27,910 --> 03:15:26,319
some other astrobiological context i

4983
03:15:29,750 --> 03:15:27,920
think this is something that's really

4984
03:15:31,590 --> 03:15:29,760
not been looked into at all and

4985
03:15:33,670 --> 03:15:31,600
something that i really would hope that

4986
03:15:36,070 --> 03:15:33,680
we can get people convinced

4987
03:15:37,990 --> 03:15:36,080
to start to look at but it's not just in

4988
03:15:39,429 --> 03:15:38,000

the oceans it turns out if you look at

4989
03:15:40,630 --> 03:15:39,439
our genome

4990
03:15:42,550 --> 03:15:40,640
it's also

4991
03:15:45,429 --> 03:15:42,560
massive numbers of viruses sorry the

4992
03:15:48,550 --> 03:15:45,439
animation here got lost in the process

4993
03:15:50,630 --> 03:15:48,560
this is a linear representation of our

4994
03:15:54,710 --> 03:15:50,640
human genome here

4995
03:15:56,950 --> 03:15:54,720
this gray part here is really clearly

4996
03:15:59,190 --> 03:15:56,960
viral derived

4997
03:16:01,590 --> 03:15:59,200
retroviral like elements

4998
03:16:03,830 --> 03:16:01,600
about eight percent of our human genome

4999
03:16:06,070 --> 03:16:03,840
are these retroviral like elements

5000
03:16:07,429 --> 03:16:06,080
line elements and sine elements

5001
03:16:09,349 --> 03:16:07,439
people can argue about whether

5002
03:16:11,990 --> 03:16:09,359
retroviruses came from

5003
03:16:13,429 --> 03:16:12,000
these retro-like elements some people

5004
03:16:15,670 --> 03:16:13,439
argue one way some people argue the

5005
03:16:16,710 --> 03:16:15,680
other but b as it may there were 40 of

5006
03:16:19,190 --> 03:16:16,720
our genome

5007
03:16:21,349 --> 03:16:19,200
compared to our protein coding parts of

5008
03:16:23,110 --> 03:16:21,359
the genome which are 1.5

5009
03:16:24,870 --> 03:16:23,120
so basically we're more viral than we

5010
03:16:25,590 --> 03:16:24,880
are human

5011
03:16:27,910 --> 03:16:25,600
and

5012
03:16:29,830 --> 03:16:27,920
that also is something that i think most

5013
03:16:31,750 --> 03:16:29,840

people don't think too much about they

5014
03:16:33,030 --> 03:16:31,760
think viruses are bad well then then

5015
03:16:35,349 --> 03:16:33,040
we're bad too

5016
03:16:38,469 --> 03:16:35,359
we can argue about that as well

5017
03:16:39,750 --> 03:16:38,479
but i think that we've also gotten some

5018
03:16:41,910 --> 03:16:39,760
really nice examples of some of the

5019
03:16:44,550 --> 03:16:41,920
chats we were talking about earlier um

5020
03:16:47,190 --> 03:16:44,560
talking about good viruses and how

5021
03:16:48,950 --> 03:16:47,200
viruses can be good my favorite example

5022
03:16:50,229 --> 03:16:48,960
of good viruses

5023
03:16:51,990 --> 03:16:50,239
is the

5024
03:16:53,269 --> 03:16:52,000
set of viruses the thermal tolerance

5025
03:16:54,790 --> 03:16:53,279
devices that have been studied by

5026
03:16:56,389 --> 03:16:54,800
marilyn roosy

5027
03:16:59,349 --> 03:16:56,399
and some of her colleagues that's this

5028
03:17:01,269 --> 03:16:59,359
is a nice review in 2011 um also i think

5029
03:17:03,670 --> 03:17:01,279
we got to reference later

5030
03:17:05,150 --> 03:17:03,680
this is a plant in fact something called

5031
03:17:08,790 --> 03:17:05,160
dicanthelian

5032
03:17:11,429 --> 03:17:08,800
lenuginosum um also known as panic grass

5033
03:17:12,710 --> 03:17:11,439
we actually see this around all of our

5034
03:17:15,750 --> 03:17:12,720
thermal sites when we're going and

5035
03:17:17,670 --> 03:17:15,760
sampling for the extremophilic viruses

5036
03:17:20,790 --> 03:17:17,680
this panic grass basically tells you

5037
03:17:22,150 --> 03:17:20,800
where not to step because it's at 50

5038
03:17:24,229 --> 03:17:22,160

degrees plus

5039
03:17:26,389 --> 03:17:24,239
but it turns out that 50 degrees celsius

5040
03:17:27,429 --> 03:17:26,399
the plant can only survive under these

5041
03:17:30,309 --> 03:17:27,439
conditions

5042
03:17:32,389 --> 03:17:30,319
if it's infected with a fungus and that

5043
03:17:35,110 --> 03:17:32,399
fungus is infected with one of these

5044
03:17:38,150 --> 03:17:35,120
thermal tolerance viruses so this i

5045
03:17:40,550 --> 03:17:38,160
think is probably the best example of a

5046
03:17:43,190 --> 03:17:40,560
good virus because

5047
03:17:45,910 --> 03:17:43,200
without this good virus this plant could

5048
03:17:47,590 --> 03:17:45,920
definitely not survive and there's some

5049
03:17:49,510 --> 03:17:47,600
great um there's a science article in

5050
03:17:51,110 --> 03:17:49,520
fact that marilyn wrote about this you

5051
03:17:53,269 --> 03:17:51,120
can clearly see that the plant does not

5052
03:17:54,630 --> 03:17:53,279
do well at 50 degrees unless it has both

5053
03:17:56,229 --> 03:17:54,640
the virus

5054
03:17:57,750 --> 03:17:56,239
and the fungus

5055
03:17:59,910 --> 03:17:57,760
so the last thing i wanted to talk about

5056
03:18:02,150 --> 03:17:59,920
in terms of a general

5057
03:18:04,229 --> 03:18:02,160
virology point of view is talk a little

5058
03:18:05,269 --> 03:18:04,239
about the origin of viruses and i know

5059
03:18:07,429 --> 03:18:05,279
this is

5060
03:18:09,750 --> 03:18:07,439
highly controversial but i just wanted

5061
03:18:12,229 --> 03:18:09,760
to throw it out there um mark krupavik

5062
03:18:13,190 --> 03:18:12,239
valerian dolga and eugene coonan

5063
03:18:15,110 --> 03:18:13,200

came out with an article in nature

5064
03:18:16,870 --> 03:18:15,120
reviews microbiology just came out a

5065
03:18:19,429 --> 03:18:16,880
little bit earlier this year

5066
03:18:20,950 --> 03:18:19,439
thinking about the origin of viruses and

5067
03:18:22,790 --> 03:18:20,960
whether we're ever going to know about

5068
03:18:23,670 --> 03:18:22,800
the origin of viruses where they came

5069
03:18:26,710 --> 03:18:23,680
from

5070
03:18:28,870 --> 03:18:26,720
of was there a universal viral common

5071
03:18:30,389 --> 03:18:28,880
ancestor i think these are you know very

5072
03:18:32,950 --> 03:18:30,399
open questions

5073
03:18:35,110 --> 03:18:32,960
but at least the way that

5074
03:18:37,349 --> 03:18:35,120
krubervik golgi and kunan put it they

5075
03:18:40,469 --> 03:18:37,359
thought that there were these

5076
03:18:43,429 --> 03:18:40,479
autonomous replicators which were

5077
03:18:45,670 --> 03:18:43,439
probably present at a pre-cellular stage

5078
03:18:47,510 --> 03:18:45,680
of evolution and you can see some of

5079
03:18:49,269 --> 03:18:47,520
that by looking at the phylogeny of some

5080
03:18:51,429 --> 03:18:49,279
of these genes here and i'll let you

5081
03:18:53,269 --> 03:18:51,439
read the paper take a look at it

5082
03:18:55,750 --> 03:18:53,279
one of the big questions that they tried

5083
03:18:58,550 --> 03:18:55,760
to answer is how do you get capsids how

5084
03:18:59,750 --> 03:18:58,560
do you get virions and that seems to

5085
03:19:03,190 --> 03:18:59,760
have been

5086
03:19:06,070 --> 03:19:03,200
potentially for the very first viruses

5087
03:19:08,710 --> 03:19:06,080
they picked up virions from some parts

5088
03:19:11,750 --> 03:19:08,720

of a primitive cell

5089
03:19:14,070 --> 03:19:11,760
but clearly over evolutionary time there

5090
03:19:16,150 --> 03:19:14,080
have been massive amounts of gene

5091
03:19:18,950 --> 03:19:16,160
transfer that have allowed the

5092
03:19:22,309 --> 03:19:18,960
development of all of these other

5093
03:19:24,950 --> 03:19:22,319
friendly or not so friendly viruses

5094
03:19:28,070 --> 03:19:24,960
that we know today

5095
03:19:29,670 --> 03:19:28,080
so that's the background stuff i want to

5096
03:19:32,070 --> 03:19:29,680
spend a couple of minutes talking about

5097
03:19:37,750 --> 03:19:32,080
some of the stuff that my lab is doing

5098
03:19:42,070 --> 03:19:38,950
oh

5099
03:19:43,990 --> 03:19:42,080
so i've been told the translation is

5100
03:19:46,389 --> 03:19:44,000
hell virus

5101
03:19:48,469 --> 03:19:46,399
this is chinookie goku

5102
03:19:51,269 --> 03:19:48,479
beppu japan where i was fortunate enough

5103
03:19:53,349 --> 03:19:51,279
to go before and we're pretty sure but

5104
03:19:55,590 --> 03:19:53,359
not absolutely certain that this is the

5105
03:19:57,990 --> 03:19:55,600
place that our favorite virion comes

5106
03:19:59,990 --> 03:19:58,000
from ssv1 which we've solved the

5107
03:20:02,309 --> 03:20:00,000
structure of together with mark murray

5108
03:20:04,070 --> 03:20:02,319
at the ut medical branch and um this

5109
03:20:05,910 --> 03:20:04,080
background here is actually boiling

5110
03:20:08,389 --> 03:20:05,920
springs lake so we've been studying the

5111
03:20:10,630 --> 03:20:08,399
structure of these viruses we've been

5112
03:20:12,710 --> 03:20:10,640
studying the genetics of these viruses

5113
03:20:15,190 --> 03:20:12,720

most of this work was the work done by

5114
03:20:18,309 --> 03:20:15,200
eric iverson and one of the things that

5115
03:20:21,030 --> 03:20:18,319
we found there is that this virus genome

5116
03:20:22,790 --> 03:20:21,040
is incredibly

5117
03:20:25,269 --> 03:20:22,800
tolerant to mutation we can actually

5118
03:20:29,030 --> 03:20:25,279
mutate about half the genes in this

5119
03:20:32,150 --> 03:20:29,040
viral genome and the virus is still okay

5120
03:20:33,990 --> 03:20:32,160
which is very bizarre most viral genomes

5121
03:20:37,429 --> 03:20:34,000
you start messing around with some of

5122
03:20:39,190 --> 03:20:37,439
the virus genes they become unhappy

5123
03:20:41,190 --> 03:20:39,200
very very quickly and again this is all

5124
03:20:43,670 --> 03:20:41,200
the work of eric iverson here we are in

5125
03:20:45,110 --> 03:20:43,680
yellowstone standing next to what we

5126
03:20:46,630 --> 03:20:45,120
thought was the hot spring that tac

5127
03:20:48,630 --> 03:20:46,640
polymerase came from

5128
03:20:51,750 --> 03:20:48,640
um actually it's over here about a

5129
03:20:53,429 --> 03:20:51,760
quarter of a mile so um we're close so

5130
03:20:55,110 --> 03:20:53,439
why are we studying these viruses in the

5131
03:20:56,950 --> 03:20:55,120
hot springs well one of the reasons is

5132
03:20:58,070 --> 03:20:56,960
they're just bizarre and interesting

5133
03:21:00,070 --> 03:20:58,080
because they've got these really

5134
03:21:02,550 --> 03:21:00,080
fascinating structures and certainly

5135
03:21:04,389 --> 03:21:02,560
nasa is interested in

5136
03:21:06,229 --> 03:21:04,399
viruses in extreme environments just

5137
03:21:08,830 --> 03:21:06,239
thinking about extreme environments and

5138
03:21:12,309 --> 03:21:08,840

what this means for

5139
03:21:13,349 --> 03:21:12,319
understanding viruses in general and we

5140
03:21:16,070 --> 03:21:13,359
also

5141
03:21:18,469 --> 03:21:16,080
got a microbial observatory project

5142
03:21:20,229 --> 03:21:18,479
funded by the nsf for this place called

5143
03:21:22,229 --> 03:21:20,239
boiling springs lake in lassen volcanic

5144
03:21:24,870 --> 03:21:22,239
national park and this sort of defines

5145
03:21:27,910 --> 03:21:24,880
extreme environment um high temperatures

5146
03:21:29,990 --> 03:21:27,920
95 degrees celsius a ph of two

5147
03:21:32,389 --> 03:21:30,000
relatively easy to get to low

5148
03:21:34,950 --> 03:21:32,399
temperature is 50 degrees and so we

5149
03:21:37,190 --> 03:21:34,960
actually were able to detect some

5150
03:21:39,429 --> 03:21:37,200
eukaryotes in the system as well as

5151
03:21:41,830 --> 03:21:39,439
bacteria and archaea and we of course

5152
03:21:43,269 --> 03:21:41,840
were really interested in the viruses

5153
03:21:44,630 --> 03:21:43,279
and none of my students wanted to go out

5154
03:21:46,630 --> 03:21:44,640
on the lake and collect samples so we

5155
03:21:49,429 --> 03:21:46,640
had this little rov to go out on the

5156
03:21:51,429 --> 03:21:49,439
lake and collect some samples for us

5157
03:21:53,510 --> 03:21:51,439
we tried to find viruses here we had no

5158
03:21:55,030 --> 03:21:53,520
luck whatsoever so we actually did a

5159
03:21:56,790 --> 03:21:55,040
meta genome and we've heard lots about

5160
03:21:58,630 --> 03:21:56,800
meta genomes already

5161
03:22:01,110 --> 03:21:58,640
one of the things about meta genomes is

5162
03:22:02,630 --> 03:22:01,120
usually lots of material and so these

5163
03:22:05,670 --> 03:22:02,640

are a couple of my graduate students

5164
03:22:07,510 --> 03:22:05,680
again eric iverson and then the leader

5165
03:22:10,070 --> 03:22:07,520
of this set of work which was jeff

5166
03:22:12,150 --> 03:22:10,080
deemer collecting samples bringing them

5167
03:22:13,590 --> 03:22:12,160
out and doing some metagenomes and to

5168
03:22:16,630 --> 03:22:13,600
make a long story short in this

5169
03:22:18,950 --> 03:22:16,640
particular metagenome we found this very

5170
03:22:21,110 --> 03:22:18,960
very strange

5171
03:22:23,590 --> 03:22:21,120
viral genome that look like a

5172
03:22:26,150 --> 03:22:23,600
recombination between a single-stranded

5173
03:22:28,229 --> 03:22:26,160
dna virus and a single-stranded rna

5174
03:22:32,150 --> 03:22:28,239
virus and

5175
03:22:34,630 --> 03:22:32,160
this is a recombination that was never

5176
03:22:36,229 --> 03:22:34,640
supposed to happen um everything that

5177
03:22:37,429 --> 03:22:36,239
people had said it doesn't happen but as

5178
03:22:39,750 --> 03:22:37,439
we heard from one of the previous talks

5179
03:22:41,830 --> 03:22:39,760
viruses don't care um i think that was

5180
03:22:44,150 --> 03:22:41,840
evelyn as well viruses don't care what

5181
03:22:45,030 --> 03:22:44,160
we call them they just do it anyway and

5182
03:22:50,950 --> 03:22:45,040
so

5183
03:22:53,349 --> 03:22:50,960
a single-stranded rna virus got together

5184
03:22:55,510 --> 03:22:53,359
and generated this

5185
03:22:56,309 --> 03:22:55,520
what we're calling a now a crucivirus

5186
03:23:00,469 --> 03:22:56,319
and

5187
03:23:02,870 --> 03:23:00,479
publishing

5188
03:23:04,469 --> 03:23:02,880

a really interesting paper on this

5189
03:23:07,349 --> 03:23:04,479
that will kind of put some of his

5190
03:23:10,870 --> 03:23:07,359
recombination uh machineries to shame at

5191
03:23:13,190 --> 03:23:10,880
least we think that might be the case

5192
03:23:14,309 --> 03:23:13,200
so finishing up really quickly i wanted

5193
03:23:16,150 --> 03:23:14,319
to talk about virus bio signatures

5194
03:23:18,950 --> 03:23:16,160
biosignatures

5195
03:23:20,710 --> 03:23:18,960
this is a sampling site in yellowstone

5196
03:23:23,190 --> 03:23:20,720
national park you can tell the hair is a

5197
03:23:24,630 --> 03:23:23,200
little less gray so this is a few years

5198
03:23:27,590 --> 03:23:24,640
ago now

5199
03:23:29,910 --> 03:23:27,600
here i discovered a virus that we called

5200
03:23:33,030 --> 03:23:29,920
sulfolobus turreted a causahedral virus

5201
03:23:35,110 --> 03:23:33,040
when i was a postdoc in mark young's lab

5202
03:23:36,790 --> 03:23:35,120
liang tang and jack johnson's lab solved

5203
03:23:38,790 --> 03:23:36,800
the structure of this absolutely

5204
03:23:40,950 --> 03:23:38,800
beautiful particle we got really excited

5205
03:23:42,950 --> 03:23:40,960
about these projections at the five-fold

5206
03:23:45,349 --> 03:23:42,960
axes of symmetry but they noticed

5207
03:23:47,990 --> 03:23:45,359
something i think even more interesting

5208
03:23:49,990 --> 03:23:48,000
and that is is that the code protein of

5209
03:23:52,309 --> 03:23:50,000
this virus which is now infecting

5210
03:23:56,150 --> 03:23:52,319
archaea was practically identical

5211
03:23:57,190 --> 03:23:56,160
between archaea bacteria and eukaryotic

5212
03:23:59,830 --> 03:23:57,200
cells

5213
03:24:01,510 --> 03:23:59,840

is this horizontal gene transfer maybe

5214
03:24:03,110 --> 03:24:01,520
but that would be between archaea

5215
03:24:04,469 --> 03:24:03,120
bacteria and eukaryote we think this is

5216
03:24:06,630 --> 03:24:04,479
unlikely

5217
03:24:08,950 --> 03:24:06,640
could be convergent evolution

5218
03:24:11,030 --> 03:24:08,960
we think it's an example for a common

5219
03:24:12,469 --> 03:24:11,040
ancestor but a common ancestor that now

5220
03:24:14,790 --> 03:24:12,479
would be around

5221
03:24:17,030 --> 03:24:14,800
since you had the

5222
03:24:20,550 --> 03:24:17,040
distribution of

5223
03:24:22,790 --> 03:24:20,560
bacteria archaea and eukarya so overview

5224
03:24:25,349 --> 03:24:22,800
of this is the papers

5225
03:24:27,590 --> 03:24:25,359
basically if you look at one of these

5226
03:24:29,990 --> 03:24:27,600
cellular trees of life

5227
03:24:32,150 --> 03:24:30,000
eukaryotes here archaea here and

5228
03:24:33,990 --> 03:24:32,160
bacteria here

5229
03:24:36,070 --> 03:24:34,000
now we think that there are some of

5230
03:24:38,229 --> 03:24:36,080
these structures again that might

5231
03:24:39,190 --> 03:24:38,239
provide some biomarkers

5232
03:24:42,150 --> 03:24:39,200
that are

5233
03:24:44,469 --> 03:24:42,160
really potentially extremely ancient

5234
03:24:46,389 --> 03:24:44,479
whether that primordial virus back here

5235
03:24:48,309 --> 03:24:46,399
looked like this one um probably a

5236
03:24:51,030 --> 03:24:48,319
little bit open to question

5237
03:24:53,110 --> 03:24:51,040
but this is all clearly extremely

5238
03:24:54,710 --> 03:24:53,120

indirect and so

5239
03:24:56,550 --> 03:24:54,720
this is where i think one of the people

5240
03:24:59,590 --> 03:24:56,560
was chatting about you know virus

5241
03:25:01,910 --> 03:24:59,600
taxidermy as well as virus taxonomy

5242
03:25:04,229 --> 03:25:01,920
we got interested in viruses in the rock

5243
03:25:06,790 --> 03:25:04,239
record and this is where i have to thank

5244
03:25:09,750 --> 03:25:06,800
one of penny's um

5245
03:25:13,190 --> 03:25:09,760
i guess predecessors um for funding this

5246
03:25:15,429 --> 03:25:13,200
absolutely crazy project on virus

5247
03:25:17,110 --> 03:25:15,439
fossils so we got some directors

5248
03:25:19,670 --> 03:25:17,120
discretionary fund money for this from

5249
03:25:22,229 --> 03:25:19,680
carl pilcher and basically went looking

5250
03:25:24,389 --> 03:25:22,239
for virus fossils in some of these hot

5251
03:25:27,910 --> 03:25:24,399
spring environments and when i say we

5252
03:25:29,750 --> 03:25:27,920
most of that is dr jim ladler who was an

5253
03:25:32,790 --> 03:25:29,760
anesthesiologist before he came to work

5254
03:25:34,710 --> 03:25:32,800
in my lab and got a phd

5255
03:25:36,550 --> 03:25:34,720
nice to have all of his

5256
03:25:38,870 --> 03:25:36,560
anesthesia equipment when you're out in

5257
03:25:41,990 --> 03:25:38,880
the field but to make a long story short

5258
03:25:45,110 --> 03:25:42,000
what jim found was you could take

5259
03:25:48,309 --> 03:25:45,120
virions particularly those like t4 and

5260
03:25:50,070 --> 03:25:48,319
solidify them i.e coat them in silica

5261
03:25:51,510 --> 03:25:50,080
which is not fossilization as the

5262
03:25:53,830 --> 03:25:51,520
geologist will tell me it's

5263
03:25:55,990 --> 03:25:53,840

mineralization of various different

5264
03:25:58,389 --> 03:25:56,000
viruses we mineralized

5265
03:26:01,269 --> 03:25:58,399
tobacco mosaic virus we mineralized

5266
03:26:03,429 --> 03:26:01,279
bacteriophage t4 we also mineralized

5267
03:26:05,910 --> 03:26:03,439
some of our hot spring viruses and what

5268
03:26:09,110 --> 03:26:05,920
those look like for short periods of

5269
03:26:10,710 --> 03:26:09,120
time literally hours here's some silica

5270
03:26:13,510 --> 03:26:10,720
treated and here's some unsilic

5271
03:26:14,710 --> 03:26:13,520
non-treated bacteriophage t4

5272
03:26:17,750 --> 03:26:14,720
that

5273
03:26:19,269 --> 03:26:17,760
you've got some

5274
03:26:21,429 --> 03:26:19,279
solicification happening but

5275
03:26:23,190 --> 03:26:21,439
unfortunately over

5276
03:26:25,750 --> 03:26:23,200
time periods more than literally a

5277
03:26:28,070 --> 03:26:25,760
couple of days you lost any kind of

5278
03:26:30,469 --> 03:26:28,080
morphological signal at least at the

5279
03:26:31,590 --> 03:26:30,479
resolution that we were able to look at

5280
03:26:32,870 --> 03:26:31,600
here

5281
03:26:35,349 --> 03:26:32,880
but then

5282
03:26:37,510 --> 03:26:35,359
serendipity raised its i think beautiful

5283
03:26:39,429 --> 03:26:37,520
head in this particular case

5284
03:26:41,990 --> 03:26:39,439
and jim took a look at what actually

5285
03:26:43,670 --> 03:26:42,000
happened to these viruses when they were

5286
03:26:46,070 --> 03:26:43,680
silica treated and when they're silicon

5287
03:26:50,070 --> 03:26:46,080
coated for a short period of time and

5288
03:26:51,750 --> 03:26:50,080

what he found was is that some viruses

5289
03:26:54,070 --> 03:26:51,760
didn't lose their infectivity at all in

5290
03:26:57,269 --> 03:26:54,080
silica treatment some lost a lot of

5291
03:26:59,590 --> 03:26:57,279
infectivity which is here on the y-axis

5292
03:27:02,950 --> 03:26:59,600
and some were kind of intermediate hot

5293
03:27:05,910 --> 03:27:02,960
spring viruses lost less infectivity

5294
03:27:09,429 --> 03:27:05,920
than bacteriophage t4

5295
03:27:10,550 --> 03:27:09,439
and some human viruses this is vaccinia

5296
03:27:12,550 --> 03:27:10,560
virus

5297
03:27:14,630 --> 03:27:12,560
actually lost

5298
03:27:16,389 --> 03:27:14,640
a huge amount of activity and this is

5299
03:27:18,389 --> 03:27:16,399
kind of what you'd expect once you get a

5300
03:27:19,910 --> 03:27:18,399
silica coating on things our big

5301
03:27:21,110 --> 03:27:19,920
surprise came

5302
03:27:25,110 --> 03:27:21,120
when

5303
03:27:28,630 --> 03:27:25,120
conditions where there wasn't silica

5304
03:27:30,950 --> 03:27:28,640
regain their infectivity so we have a

5305
03:27:33,429 --> 03:27:30,960
what we're calling the zombie viruses

5306
03:27:35,590 --> 03:27:33,439
here um which have then

5307
03:27:37,309 --> 03:27:35,600
come back to life

5308
03:27:40,389 --> 03:27:37,319
based on this

5309
03:27:41,910 --> 03:27:40,399
solicification and desolicification and

5310
03:27:43,190 --> 03:27:41,920
a number of people really excited about

5311
03:27:44,790 --> 03:27:43,200
this we've got viruses we're gonna have

5312
03:27:46,150 --> 03:27:44,800
viruses on meteorites they're gonna be

5313
03:27:48,469 --> 03:27:46,160

transferring we're gonna get all the

5314
03:27:51,349 --> 03:27:48,479
martian nasty viruses being transferred

5315
03:27:53,510 --> 03:27:51,359
to this planet um turns out that after

5316
03:27:56,790 --> 03:27:53,520
about a month these lose their

5317
03:27:58,790 --> 03:27:56,800
infectivity completely so um unlikely

5318
03:28:00,950 --> 03:27:58,800
that this process at the very least will

5319
03:28:04,469 --> 03:28:00,960
be able to preserve something

5320
03:28:07,510 --> 03:28:04,479
via space flight but endolythic viruses

5321
03:28:09,830 --> 03:28:07,520
maybe it's certainly a possibility

5322
03:28:13,110 --> 03:28:09,840
so with that i just wanted to finish up

5323
03:28:16,389 --> 03:28:13,120
with some of our proposals that we had

5324
03:28:18,309 --> 03:28:16,399
in our astrobiology article

5325
03:28:20,070 --> 03:28:18,319
and basically i'm just going to leave

5326
03:28:22,469 --> 03:28:20,080
these up here

5327
03:28:24,550 --> 03:28:22,479
and answer questions on the rest of my

5328
03:28:27,030 --> 03:28:24,560
talk and then hopefully

5329
03:28:30,070 --> 03:28:27,040
we can come back and revisit some of

5330
03:28:32,389 --> 03:28:30,080
these questions um a little bit later on

5331
03:28:33,750 --> 03:28:32,399
so with that i'm only five minutes over

5332
03:28:34,550 --> 03:28:33,760
which is an organizer i think is pretty

5333
03:28:36,229 --> 03:28:34,560
good

5334
03:28:49,990 --> 03:28:36,239
and i'll take questions if anyone has

5335
03:28:58,229 --> 03:28:51,190
turn the lights off so maybe you can

5336
03:29:03,030 --> 03:29:00,630
nobody wants to ask questions

5337
03:29:04,309 --> 03:29:03,040
i have a question it's probably pretty

5338
03:29:06,790 --> 03:29:04,319

naive but

5339
03:29:08,950 --> 03:29:06,800
is there any way that you could imagine

5340
03:29:11,190 --> 03:29:08,960
kind of creating that zombie virus other

5341
03:29:13,110 --> 03:29:11,200
than mineralization or maybe a different

5342
03:29:14,950 --> 03:29:13,120
type of mineralization where they could

5343
03:29:18,150 --> 03:29:14,960
be preserved for

5344
03:29:20,150 --> 03:29:18,160
uh longer durations of time and then

5345
03:29:21,910 --> 03:29:20,160
their you know morphological features

5346
03:29:23,670 --> 03:29:21,920
might not be preserved but that they

5347
03:29:25,590 --> 03:29:23,680
could be brought back after an extended

5348
03:29:28,550 --> 03:29:25,600
period of time

5349
03:29:29,269 --> 03:29:28,560
so that's um that's a great question um

5350
03:29:33,110 --> 03:29:29,279
we

5351
03:29:36,389 --> 03:29:33,120
relatively straightforward um there has

5352
03:29:38,309 --> 03:29:36,399
been some work um of jennifer kyle

5353
03:29:40,870 --> 03:29:38,319
looking at iron

5354
03:29:43,269 --> 03:29:40,880
and seeing if um buyer's interactions

5355
03:29:45,269 --> 03:29:43,279
with iron could uh help to preserve them

5356
03:29:47,510 --> 03:29:45,279
but nothing really that seems to be able

5357
03:29:50,150 --> 03:29:47,520
to stabilize them

5358
03:29:53,030 --> 03:29:50,160
more than silica has been able to so i

5359
03:29:54,790 --> 03:29:53,040
think that that's um

5360
03:29:57,269 --> 03:29:54,800
we've probably done about as good as we

5361
03:29:58,710 --> 03:29:57,279
can with these particular viruses and

5362
03:30:01,110 --> 03:29:58,720
mostly we've been looking at you know

5363
03:30:03,110 --> 03:30:01,120

bacteriophage t4

5364
03:30:04,790 --> 03:30:03,120
i have much higher hopes for things that

5365
03:30:07,590 --> 03:30:04,800
you might find some viruses and endless

5366
03:30:08,870 --> 03:30:07,600
maybe some of gary's soil viruses

5367
03:30:10,630 --> 03:30:08,880
might have

5368
03:30:13,429 --> 03:30:10,640
some kind of

5369
03:30:15,429 --> 03:30:13,439
way that we can look at those i i really

5370
03:30:18,550 --> 03:30:15,439
don't know but i think it's a wide open

5371
03:30:20,469 --> 03:30:18,560
area for people to start looking at

5372
03:30:23,030 --> 03:30:20,479
yeah one of the things i wanted to

5373
03:30:24,550 --> 03:30:23,040
mention just as a follow-on to that ken

5374
03:30:27,070 --> 03:30:24,560
is that

5375
03:30:30,229 --> 03:30:27,080
microbial communities in general

5376
03:30:32,070 --> 03:30:30,239
self-fossilize in the highly mineralized

5377
03:30:33,269 --> 03:30:32,080
environments of gaze

5378
03:30:34,950 --> 03:30:33,279
and so

5379
03:30:36,070 --> 03:30:34,960
a lot of that is

5380
03:30:38,550 --> 03:30:36,080
excuse me because there are

5381
03:30:40,309 --> 03:30:38,560
chemolythotrophic organisms there that

5382
03:30:43,429 --> 03:30:40,319
are getting their energy out of

5383
03:30:44,469 --> 03:30:43,439
transforming uh one mineral into another

5384
03:30:47,190 --> 03:30:44,479
typically

5385
03:30:49,429 --> 03:30:47,200
from you know an oxidation state to a

5386
03:30:51,030 --> 03:30:49,439
higher oxidation state and because of

5387
03:30:52,790 --> 03:30:51,040
that there's essentially coating

5388
03:30:55,030 --> 03:30:52,800

themselves in

5389
03:30:56,950 --> 03:30:55,040
minerals and so and then they're also

5390
03:30:58,550 --> 03:30:56,960
undisturbed by weather

5391
03:31:01,349 --> 03:30:58,560
so it's a wonderful preservation

5392
03:31:03,269 --> 03:31:01,359
environment and i think uh not to keep

5393
03:31:04,870 --> 03:31:03,279
harping on caves as wonderful as they

5394
03:31:07,590 --> 03:31:04,880
are but i think you know it's one of

5395
03:31:10,830 --> 03:31:07,600
those places where one might be able to

5396
03:31:15,110 --> 03:31:10,840
actually find undisturbed in

5397
03:31:16,469 --> 03:31:15,120
situ uh viral fossil materials and caves

5398
03:31:19,030 --> 03:31:16,479
come in all different kinds of

5399
03:31:21,590 --> 03:31:19,040
geochemical flavors so sometimes there's

5400
03:31:23,670 --> 03:31:21,600
silica gel that you know is fossilizing

5401
03:31:26,309 --> 03:31:23,680
things very often it's carbonate

5402
03:31:29,110 --> 03:31:26,319
sometimes it's gypsum uh it can be

5403
03:31:31,269 --> 03:31:29,120
copper it can be iron it can be sulfur

5404
03:31:32,790 --> 03:31:31,279
so there are a lot of possibilities for

5405
03:31:34,150 --> 03:31:32,800
actually looking

5406
03:31:37,269 --> 03:31:34,160
you know if we knew what the heck we

5407
03:31:41,670 --> 03:31:39,750
we might be able to point

5408
03:31:43,030 --> 03:31:41,680
folks in the right direction or provide

5409
03:31:44,870 --> 03:31:43,040
samples

5410
03:31:47,510 --> 03:31:44,880
just saying

5411
03:31:49,590 --> 03:31:47,520
yeah no and i i completely agree i do

5412
03:31:51,110 --> 03:31:49,600
think that this is something that we

5413
03:31:52,710 --> 03:31:51,120

should be looking for

5414
03:31:54,710 --> 03:31:52,720
um and something that we should

5415
03:31:56,550 --> 03:31:54,720
definitely put into a

5416
03:31:57,750 --> 03:31:56,560
review article and or white paper is i

5417
03:31:59,349 --> 03:31:57,760
think we need to

5418
03:32:01,750 --> 03:31:59,359
look at some of these things and part of

5419
03:32:03,429 --> 03:32:01,760
the problem that we had with gym study

5420
03:32:07,030 --> 03:32:03,439
and the original

5421
03:32:08,950 --> 03:32:07,040
silica coding of t4

5422
03:32:11,269 --> 03:32:08,960
was actually a lack of resolution it

5423
03:32:12,070 --> 03:32:11,279
wasn't a case of

5424
03:32:14,550 --> 03:32:12,080
uh

5425
03:32:16,550 --> 03:32:14,560
really you know the the morphology was

5426
03:32:18,389 --> 03:32:16,560
gone but if we had some kind of higher

5427
03:32:20,630 --> 03:32:18,399
resolution way of looking at chemical

5428
03:32:23,990 --> 03:32:20,640
structures we probably would have a

5429
03:32:26,229 --> 03:32:24,000
really nice bio signature it's just that

5430
03:32:28,469 --> 03:32:26,239
i don't know enough about that kind of

5431
03:32:30,309 --> 03:32:28,479
technology and i'd you know love to talk

5432
03:32:32,790 --> 03:32:30,319
to more of the engineers who are trying

5433
03:32:35,830 --> 03:32:32,800
to think about some of these things

5434
03:32:37,510 --> 03:32:35,840
right great thanks has anyone subjected

5435
03:32:39,830 --> 03:32:37,520
them to like a space environment like

5436
03:32:43,670 --> 03:32:39,840
the tardigrades to see if they

5437
03:32:46,150 --> 03:32:43,680
would be okay after a while

5438
03:32:48,070 --> 03:32:46,160

so that's that's another great question

5439
03:32:49,590 --> 03:32:48,080
and i think penny sort of alluded to

5440
03:32:51,830 --> 03:32:49,600
this when she was mentioning you know

5441
03:32:53,510 --> 03:32:51,840
numbers of papers in english

5442
03:32:55,269 --> 03:32:53,520
there are a couple of russian papers

5443
03:32:57,110 --> 03:32:55,279
actually so some of the very early

5444
03:32:59,510 --> 03:32:57,120
russian space program

5445
03:33:01,750 --> 03:32:59,520
looked at some viruses including tobacco

5446
03:33:02,550 --> 03:33:01,760
mosaic virus in the space environment

5447
03:33:05,110 --> 03:33:02,560
and

5448
03:33:06,870 --> 03:33:05,120
they seem to do reasonably well

5449
03:33:09,030 --> 03:33:06,880
um and they didn't lose too much in the

5450
03:33:11,190 --> 03:33:09,040
way of infectivity but these were all

5451
03:33:13,830 --> 03:33:11,200
really pretty short-term exposure kinds

5452
03:33:16,070 --> 03:33:13,840
of experiments and so uh the longer-term

5453
03:33:17,429 --> 03:33:16,080
exposure experiments we were hoping that

5454
03:33:19,990 --> 03:33:17,439
with our silica treatment we would get

5455
03:33:22,389 --> 03:33:20,000
something that we could try but

5456
03:33:25,590 --> 03:33:22,399
if we lose if we lose infectivity in a

5457
03:33:27,830 --> 03:33:25,600
month with desiccation at

5458
03:33:28,710 --> 03:33:27,840
room temperature i'm highly unlikely

5459
03:33:30,150 --> 03:33:28,720
that's going to be something which is

5460
03:33:32,550 --> 03:33:30,160
going to survive in a space environment

5461
03:33:33,990 --> 03:33:32,560
we could try but um i think that we need

5462
03:33:35,990 --> 03:33:34,000
some more

5463
03:33:38,070 --> 03:33:36,000

maybe more simulation environment before

5464
03:33:39,510 --> 03:33:38,080
we actually go into a space environment

5465
03:33:42,710 --> 03:33:39,520
yeah

5466
03:33:46,870 --> 03:33:44,710
there are some um

5467
03:33:48,550 --> 03:33:46,880
other comments here

5468
03:33:51,429 --> 03:33:48,560
yeah some i've lost my

5469
03:33:56,389 --> 03:33:51,439
um comment window here

5470
03:34:03,030 --> 03:33:58,469
people are talking are asking about

5471
03:34:11,429 --> 03:34:07,510
ah i'm not sure that um you know

5472
03:34:14,950 --> 03:34:13,429
okay so let me see

5473
03:34:17,349 --> 03:34:14,960
i'll elaborate more on the ancient folds

5474
03:34:19,349 --> 03:34:17,359
of viral proteins so sh3 or ob folds

5475
03:34:21,510 --> 03:34:19,359
okay yeah yeah so

5476
03:34:22,710 --> 03:34:21,520
sorry about that again trying to deal

5477
03:34:25,349 --> 03:34:22,720
with too many screens here

5478
03:34:26,830 --> 03:34:25,359
simultaneously

5479
03:34:30,150 --> 03:34:26,840
but yeah so

5480
03:34:32,309 --> 03:34:30,160
the um the fold that i was i didn't have

5481
03:34:34,309 --> 03:34:32,319
a chance to get into um but these the

5482
03:34:37,510 --> 03:34:34,319
folds that i mentioned for our viruses

5483
03:34:39,349 --> 03:34:37,520
the ones that again the stiv that we

5484
03:34:40,950 --> 03:34:39,359
find in yellowstone hot springs and also

5485
03:34:44,229 --> 03:34:40,960
a similar structure to the bacterial

5486
03:34:46,070 --> 03:34:44,239
virus and the ones that you find in

5487
03:34:47,349 --> 03:34:46,080
some eukaryotic viruses as well not just

5488
03:34:49,110 --> 03:34:47,359

adenovirus but a number of different

5489
03:34:52,229 --> 03:34:49,120
ones these are what are called double

5490
03:34:53,990 --> 03:34:52,239
jelly roll structures um and a lot of

5491
03:34:56,150 --> 03:34:54,000
work has been done on this with mark

5492
03:34:57,590 --> 03:34:56,160
group of egg eugene coonan but also

5493
03:35:00,550 --> 03:34:57,600
dennis banford some of the structural

5494
03:35:02,469 --> 03:35:00,560
biologists and so this double jelly roll

5495
03:35:05,349 --> 03:35:02,479
and also single jelly roll kinds of

5496
03:35:07,030 --> 03:35:05,359
proteins are very well conserved in

5497
03:35:09,510 --> 03:35:07,040
terms of their structures

5498
03:35:11,990 --> 03:35:09,520
but not very well conserved in terms of

5499
03:35:14,070 --> 03:35:12,000
their sequences and so

5500
03:35:15,670 --> 03:35:14,080
what's uh i think would be really

5501
03:35:17,830 --> 03:35:15,680
interesting to look at again potentially

5502
03:35:20,309 --> 03:35:17,840
from a biosignature point of view is if

5503
03:35:21,830 --> 03:35:20,319
we can get to that kind of resolution

5504
03:35:24,630 --> 03:35:21,840
can we find

5505
03:35:25,910 --> 03:35:24,640
double jelly roll kinds of proteins

5506
03:35:27,590 --> 03:35:25,920
anywhere

5507
03:35:28,950 --> 03:35:27,600
and so i think that's a

5508
03:35:30,389 --> 03:35:28,960
an open

5509
03:35:33,269 --> 03:35:30,399
open question

5510
03:35:34,630 --> 03:35:33,279
but certainly there's a lot of these

5511
03:35:37,670 --> 03:35:34,640
double jelly roll products but there's

5512
03:35:39,830 --> 03:35:37,680
also something called the hk97 fold

5513
03:35:42,630 --> 03:35:39,840

which is a very different kind of

5514
03:35:45,190 --> 03:35:42,640
protein structure that's also found in

5515
03:35:47,830 --> 03:35:45,200
bacterial viruses and

5516
03:35:49,830 --> 03:35:47,840
eukaryotic viruses not in

5517
03:35:51,670 --> 03:35:49,840
our kale viruses yet but our kale

5518
03:35:54,229 --> 03:35:51,680
viruses are incredibly understudied so

5519
03:36:01,110 --> 03:35:54,239
it's distinctly possible that that kind

5520
03:36:04,870 --> 03:36:02,870
let's see

5521
03:36:06,630 --> 03:36:04,880
okay there's one about the

5522
03:36:08,070 --> 03:36:06,640
examples of the russian papers on

5523
03:36:10,469 --> 03:36:08,080
viruses in their space program that i

5524
03:36:12,710 --> 03:36:10,479
mentioned those are referenced in our

5525
03:36:13,830 --> 03:36:12,720
review article i could go and um dig

5526
03:36:15,990 --> 03:36:13,840
them up but if you go into our

5527
03:36:17,670 --> 03:36:16,000
astrobiology article and if you don't

5528
03:36:21,269 --> 03:36:17,680
have a copy of it just let me know i can

5529
03:36:23,750 --> 03:36:21,279
send it to you um so they those are the

5530
03:36:25,990 --> 03:36:23,760
two that i was able to find um there may

5531
03:36:27,590 --> 03:36:26,000
be some other ones there may be other

5532
03:36:29,590 --> 03:36:27,600
you know space literature that hasn't

5533
03:36:30,790 --> 03:36:29,600
been published but at least those the

5534
03:36:32,070 --> 03:36:30,800
ones i'm able to find but they were from

5535
03:36:44,229 --> 03:36:32,080
the

5536
03:36:46,389 --> 03:36:44,239
she's saying to to suggest to collect

5537
03:36:48,550 --> 03:36:46,399
viruses from places that normally spend

5538
03:36:51,269 --> 03:36:48,560

most of the time desiccated

5539
03:36:52,630 --> 03:36:51,279
so i guess arid environments um to have

5540
03:36:55,030 --> 03:36:52,640
a better chance of long-term

5541
03:36:57,429 --> 03:36:55,040
preservation hunt for the desiccation

5542
03:36:59,590 --> 03:36:57,439
virus yeah

5543
03:37:01,670 --> 03:36:59,600
sounds like an expedition to me

5544
03:37:05,990 --> 03:37:01,680
i think i think people looked in the

5545
03:37:08,469 --> 03:37:06,000
atacama um i'm not absolutely certain

5546
03:37:13,910 --> 03:37:10,630
i think arvin did you have a question

5547
03:37:23,429 --> 03:37:15,269
you're muted arvin so if you want to

5548
03:37:23,439 --> 03:37:26,070
maybe not

5549
03:37:31,349 --> 03:37:29,349
okay um no he said yes but

5550
03:37:37,110 --> 03:37:31,359
maybe you'll um marco can everybody

5551
03:37:41,990 --> 03:37:40,070
yes i can

5552
03:37:44,389 --> 03:37:42,000
ken i had a quick question and it was to

5553
03:37:46,309 --> 03:37:44,399
do with um

5554
03:37:47,670 --> 03:37:46,319
investigating the viruses and i'm

5555
03:37:50,630 --> 03:37:47,680
wondering because in those kind of

5556
03:37:53,349 --> 03:37:50,640
scenarios you're deriving the water away

5557
03:37:54,150 --> 03:37:53,359
and that destabilizes nucleic acid

5558
03:37:56,790 --> 03:37:54,160
so

5559
03:37:59,269 --> 03:37:56,800
i'm wondering if anyone has done any

5560
03:38:01,510 --> 03:37:59,279
kind of analysis at a sequence level to

5561
03:38:04,150 --> 03:38:01,520
see are these viruses more prone to

5562
03:38:06,150 --> 03:38:04,160
deamination any other kind of

5563
03:38:08,469 --> 03:38:06,160

base changes

5564
03:38:11,110 --> 03:38:08,479
coupled with um

5565
03:38:14,070 --> 03:38:11,120
what when the viruses do recover are you

5566
03:38:15,349 --> 03:38:14,080
seeing variants of mutations that are at

5567
03:38:17,349 --> 03:38:15,359
a higher rate

5568
03:38:19,670 --> 03:38:17,359
so what i'm trying to get to is are we

5569
03:38:21,510 --> 03:38:19,680
in these kind of cycles where we might

5570
03:38:23,190 --> 03:38:21,520
be getting into say the

5571
03:38:26,950 --> 03:38:23,200
environments where they go through these

5572
03:38:31,670 --> 03:38:28,870
environments from

5573
03:38:33,110 --> 03:38:31,680
wet climatic conditions to dry

5574
03:38:36,070 --> 03:38:33,120
to wet again

5575
03:38:38,070 --> 03:38:36,080
are we seeing an accelerated boom burst

5576
03:38:40,389 --> 03:38:38,080
period of virus evolution

5577
03:38:43,269 --> 03:38:40,399
at any given period of time

5578
03:38:44,870 --> 03:38:43,279
so a great question arvind um the short

5579
03:38:47,429 --> 03:38:44,880
answer is we don't know

5580
03:38:48,630 --> 03:38:47,439
um the longer answer is actually one

5581
03:38:50,630 --> 03:38:48,640
that ties into someone else that

5582
03:38:53,030 --> 03:38:50,640
digressed on on the chat room also about

5583
03:38:56,469 --> 03:38:53,040
um basically what happens to

5584
03:39:00,790 --> 03:38:56,479
the viruses when we desiccate them and

5585
03:39:03,189 --> 03:39:00,800
when we do the coatings and as you know

5586
03:39:06,550 --> 03:39:03,199
this is actually really critical to my

5587
03:39:08,950 --> 03:39:06,560
side hustle which is i'm trying to use

5588
03:39:11,670 --> 03:39:08,960

this technology to preserve vaccines to

5589
03:39:15,189 --> 03:39:11,680
get them out of the developing world so

5590
03:39:17,429 --> 03:39:15,199
if coding and drying does something to

5591
03:39:19,189 --> 03:39:17,439
the viral genome this could clearly be a

5592
03:39:21,030 --> 03:39:19,199
big problem so something that we

5593
03:39:23,510 --> 03:39:21,040
definitely need to look at

5594
03:39:26,630 --> 03:39:23,520
oh thank you

5595
03:39:28,309 --> 03:39:26,640
there's a question from ishmael

5596
03:39:30,469 --> 03:39:28,319
somewhere here

5597
03:39:32,150 --> 03:39:30,479
yeah alabama isotope fractionation by

5598
03:39:33,830 --> 03:39:32,160
and just

5599
03:39:37,110 --> 03:39:33,840
yeah yeah

5600
03:39:39,429 --> 03:39:37,120
so i am so a wonderful question i'm by

5601
03:39:41,670 --> 03:39:39,439
no stretch of the imagination an expert

5602
03:39:44,150 --> 03:39:41,680
in looking at isotope fractionation but

5603
03:39:46,630 --> 03:39:44,160
i certainly have looked at

5604
03:39:48,950 --> 03:39:46,640
um and heard papers from red papers of

5605
03:39:51,990 --> 03:39:48,960
people who've looked at fractionation

5606
03:39:55,990 --> 03:39:52,000
particularly of c13 um in

5607
03:39:58,150 --> 03:39:56,000
rocks and in microbes and what they see

5608
03:40:00,710 --> 03:39:58,160
is that depending on the particular kind

5609
03:40:02,790 --> 03:40:00,720
of metabolism that's being used they'll

5610
03:40:04,950 --> 03:40:02,800
be more or less enrichment in this case

5611
03:40:07,349 --> 03:40:04,960
particularly of c13

5612
03:40:11,030 --> 03:40:07,359
um if i remember correctly methanogens

5613
03:40:12,309 --> 03:40:11,040

enrich a lot more for c12 than c13 and

5614
03:40:13,590 --> 03:40:12,319
you know kathy may know more about this

5615
03:40:14,710 --> 03:40:13,600
penny probably knows more about this

5616
03:40:16,870 --> 03:40:14,720
than i do

5617
03:40:19,429 --> 03:40:16,880
but i don't think that anybody's looked

5618
03:40:21,590 --> 03:40:19,439
at any of these auxiliary metabolic

5619
03:40:23,830 --> 03:40:21,600
genes and whether there's any difference

5620
03:40:26,309 --> 03:40:23,840
in terms of fractionation and the reason

5621
03:40:27,910 --> 03:40:26,319
we put this into our paper and you may

5622
03:40:28,950 --> 03:40:27,920
remember from that last slide that i had

5623
03:40:30,309 --> 03:40:28,960
up before

5624
03:40:33,030 --> 03:40:30,319
is that

5625
03:40:35,910 --> 03:40:33,040
this might be some kind of virus

5626
03:40:39,189 --> 03:40:35,920
biosignature or a by biosignature that

5627
03:40:41,269 --> 03:40:39,199
there had been a virus infection at some

5628
03:40:43,429 --> 03:40:41,279
point in the past that could be

5629
03:40:48,229 --> 03:40:43,439
detectable in the rock record

5630
03:40:52,150 --> 03:40:49,910
i'd like to add maybe

5631
03:40:54,790 --> 03:40:52,160
just hypothesizing thinking about this

5632
03:40:57,189 --> 03:40:54,800
that when a virus infects its host

5633
03:40:59,750 --> 03:40:57,199
it wants its host to be efficient it

5634
03:41:00,710 --> 03:40:59,760
wants to do its duty so you could say if

5635
03:41:02,870 --> 03:41:00,720
it's going to

5636
03:41:04,630 --> 03:41:02,880
be increasing respiration or any type of

5637
03:41:06,389 --> 03:41:04,640
metabolic outputs it's going to make it

5638
03:41:09,269 --> 03:41:06,399

more efficient and by more efficient

5639
03:41:11,990 --> 03:41:09,279
does that mean that's more likely to

5640
03:41:14,309 --> 03:41:12,000
select if we're going to use carbon 12c

5641
03:41:16,309 --> 03:41:14,319
versus 13c or because it wants to push

5642
03:41:18,309 --> 03:41:16,319
things out so quickly it's going to

5643
03:41:22,630 --> 03:41:18,319
discriminate less and there's actually

5644
03:41:24,070 --> 03:41:22,640
an enrichment of 13c and that um isotope

5645
03:41:25,910 --> 03:41:24,080
of that isotope

5646
03:41:27,269 --> 03:41:25,920
sounds like we need to do experiments

5647
03:41:29,110 --> 03:41:27,279
yeah

5648
03:41:30,710 --> 03:41:29,120
or write proposals to get funding to do

5649
03:41:32,630 --> 03:41:30,720
experiments

5650
03:41:34,950 --> 03:41:32,640
yeah i mean we can if we look at

5651
03:41:37,990 --> 03:41:34,960
anything like i i studied nitrous oxide

5652
03:41:41,349 --> 03:41:38,000
before yeah for biotic

5653
03:41:43,429 --> 03:41:41,359
the fractionation factor can change and

5654
03:41:45,510 --> 03:41:43,439
it can change by who's doing it even

5655
03:41:46,710 --> 03:41:45,520
within the organism that's doing it you

5656
03:41:48,550 --> 03:41:46,720
can look at various ones and there's

5657
03:41:50,150 --> 03:41:48,560
this fluctuation and

5658
03:41:53,429 --> 03:41:50,160
i never really understood this but i

5659
03:41:55,830 --> 03:41:53,439
definitely think it has viral influence

5660
03:41:57,830 --> 03:41:55,840
yeah and i think just the fact that

5661
03:41:59,670 --> 03:41:57,840
those auxiliary metabolic genes is a

5662
03:42:01,110 --> 03:41:59,680
pretty new thing

5663
03:42:02,469 --> 03:42:01,120

um that people are thinking about and

5664
03:42:04,870 --> 03:42:02,479
it's great that nigel talked about it

5665
03:42:07,510 --> 03:42:04,880
earlier as well and certainly simon has

5666
03:42:09,670 --> 03:42:07,520
been um very very involved in looking at

5667
03:42:11,189 --> 03:42:09,680
some of these things as well so i think

5668
03:42:12,389 --> 03:42:11,199
this is a

5669
03:42:14,229 --> 03:42:12,399
very much uh

5670
03:42:15,910 --> 03:42:14,239
expanding field

5671
03:42:17,110 --> 03:42:15,920
and hopefully something that you know

5672
03:42:19,990 --> 03:42:17,120
people are going to start to look at in

5673
03:42:23,670 --> 03:42:21,830
i've got something else to add which i

5674
03:42:25,910 --> 03:42:23,680
think might be very important and this

5675
03:42:27,269 --> 03:42:25,920
is kind of looking at environments

5676
03:42:28,790 --> 03:42:27,279
and uh

5677
03:42:30,309 --> 03:42:28,800
one of the things is that people have

5678
03:42:32,389 --> 03:42:30,319
been looking at obviously climate change

5679
03:42:35,830 --> 03:42:32,399
experiments and impacts on viruses as a

5680
03:42:36,950 --> 03:42:35,840
consequence of elevated co2 levels or

5681
03:42:39,590 --> 03:42:36,960
temperature

5682
03:42:41,429 --> 03:42:39,600
right and there's very little been done

5683
03:42:42,870 --> 03:42:41,439
on it but there are one or two papers

5684
03:42:44,870 --> 03:42:42,880
that have been done in the context

5685
03:42:46,790 --> 03:42:44,880
within the plant industry

5686
03:42:48,870 --> 03:42:46,800
and they've identified that elevating

5687
03:42:51,670 --> 03:42:48,880
the co2 level from say where we are

5688
03:42:54,150 --> 03:42:51,680

right now roughly 400 100 plus to about

5689
03:42:56,469 --> 03:42:54,160
700 uh

5690
03:42:59,030 --> 03:42:56,479
what we do get is threefold increase in

5691
03:43:01,349 --> 03:42:59,040
the rate of replication of viruses so

5692
03:43:02,950 --> 03:43:01,359
the viral load within a host increases

5693
03:43:05,110 --> 03:43:02,960
by threefold

5694
03:43:06,710 --> 03:43:05,120
and this is in the case of

5695
03:43:08,150 --> 03:43:06,720
bali yellow leaf

5696
03:43:09,189 --> 03:43:08,160
coal power so at least this is something

5697
03:43:12,070 --> 03:43:09,199
that we know

5698
03:43:13,670 --> 03:43:12,080
with uh some viruses but

5699
03:43:16,550 --> 03:43:13,680
this might be something else to think

5700
03:43:18,469 --> 03:43:16,560
about in terms of environmental factors

5701
03:43:21,110 --> 03:43:18,479
that could govern replication rates

5702
03:43:23,110 --> 03:43:21,120
turnover times and

5703
03:43:24,950 --> 03:43:23,120
exploration of sequence space in terms

5704
03:43:26,070 --> 03:43:24,960
of adaptation

5705
03:43:38,389 --> 03:43:26,080
yeah it's a great idea i think it's

5706
03:43:43,510 --> 03:43:40,389
okay i think this is also the time in

5707
03:43:46,950 --> 03:43:43,520
the schedule where i'm supposed to be um

5708
03:43:48,790 --> 03:43:46,960
talking about what's next um

5709
03:43:50,150 --> 03:43:48,800
gotta don't want to i know people you

5710
03:43:51,510 --> 03:43:50,160
need to need to leave here and it's

5711
03:43:53,670 --> 03:43:51,520
getting late for people in other time

5712
03:43:55,269 --> 03:43:53,680
zones so um

5713
03:43:56,630 --> 03:43:55,279

penny did you want to say another couple

5714
03:43:59,830 --> 03:43:56,640
of words

5715
03:44:01,590 --> 03:43:59,840
um you know just thanks to everybody for

5716
03:44:04,229 --> 03:44:01,600
attending

5717
03:44:05,990 --> 03:44:04,239
we've had a huge participation and so

5718
03:44:08,469 --> 03:44:06,000
we're delighted about that

5719
03:44:10,870 --> 03:44:08,479
and i can't wait to carry this forward

5720
03:44:14,150 --> 03:44:10,880
into watching everybody shape this into

5721
03:44:19,750 --> 03:44:16,790
many revelations to me as a

5722
03:44:22,389 --> 03:44:19,760
microbiologist focusing on cellular

5723
03:44:24,790 --> 03:44:22,399
organisms and lots of jargon i didn't

5724
03:44:27,269 --> 03:44:24,800
understand but i'm getting um

5725
03:44:29,110 --> 03:44:27,279
i'm getting up to speed and i think

5726
03:44:32,309 --> 03:44:29,120
there are a lot of subtle

5727
03:44:35,429 --> 03:44:32,319
evolutionary concepts that are emerging

5728
03:44:36,790 --> 03:44:35,439
from the viral field as i understand

5729
03:44:39,510 --> 03:44:36,800
what you guys have been talking about

5730
03:44:42,150 --> 03:44:39,520
for the last two days that are different

5731
03:44:43,429 --> 03:44:42,160
in a lot of ways from the way we think

5732
03:44:45,830 --> 03:44:43,439
about things

5733
03:44:48,630 --> 03:44:45,840
with respect to bacteria and archaea and

5734
03:44:50,550 --> 03:44:48,640
even protists and whatnot so i think

5735
03:44:53,830 --> 03:44:50,560
you know coming together of the minds

5736
03:44:56,710 --> 03:44:53,840
across this uh this divide is is really

5737
03:44:59,189 --> 03:44:56,720
very useful and getting these ideas into

5738
03:45:00,870 --> 03:44:59,199

the astrobiological context

5739
03:45:04,150 --> 03:45:00,880
out into the literature

5740
03:45:05,189 --> 03:45:04,160
will not only help the astrovirology

5741
03:45:07,990 --> 03:45:05,199
case

5742
03:45:09,990 --> 03:45:08,000
that this is a valid way to

5743
03:45:12,550 --> 03:45:10,000
spend some of nasa's money in the future

5744
03:45:13,910 --> 03:45:12,560
for funding some of the stuff but also i

5745
03:45:15,750 --> 03:45:13,920
think it's

5746
03:45:18,550 --> 03:45:15,760
a way to

5747
03:45:20,550 --> 03:45:18,560
get some of these ideas into the other

5748
03:45:22,229 --> 03:45:20,560
part of microbiology that is thinking

5749
03:45:24,229 --> 03:45:22,239
about things in a different way and

5750
03:45:25,750 --> 03:45:24,239
perhaps vice versa so

5751
03:45:27,590 --> 03:45:25,760
i'm seeing the potential for cross

5752
03:45:29,590 --> 03:45:27,600
fertilization so i want to thank

5753
03:45:32,070 --> 03:45:29,600
everybody i want to thank all of you

5754
03:45:35,590 --> 03:45:32,080
guys who uh scrambled around to get this

5755
03:45:36,389 --> 03:45:35,600
organized gary and ucan and kathy

5756
03:45:48,950 --> 03:45:36,399
and

5757
03:45:51,830 --> 03:45:48,960
there's something gary mentioned as well

5758
03:45:56,389 --> 03:45:51,840
is um whoever is interested in

5759
03:45:59,189 --> 03:45:56,399
contributing to review articles um white

5760
03:46:02,229 --> 03:45:59,199
papers etc please get in touch with us

5761
03:46:03,750 --> 03:46:02,239
at least with me email is the best way

5762
03:46:05,830 --> 03:46:03,760
i've seen some people you know ping me

5763
03:46:07,990 --> 03:46:05,840

on twitter the problem with twitter is

5764
03:46:09,429 --> 03:46:08,000
it's the fire hose approach and so

5765
03:46:12,070 --> 03:46:09,439
please

5766
03:46:14,870 --> 03:46:12,080
send me something either direct message

5767
03:46:16,870 --> 03:46:14,880
is great um but email is by far and away

5768
03:46:19,670 --> 03:46:16,880
the best way to reach me and probably

5769
03:46:20,710 --> 03:46:19,680
best to send to gary and or kathy as

5770
03:46:23,670 --> 03:46:20,720
well

5771
03:46:25,269 --> 03:46:23,680
and we will be in touch again probably

5772
03:46:27,349 --> 03:46:25,279
early next week

5773
03:46:29,590 --> 03:46:27,359
um sometime to at least sort of put

5774
03:46:32,790 --> 03:46:29,600
together an email list think about

5775
03:46:35,750 --> 03:46:32,800
um strategizing for these for these next

5776
03:46:42,389 --> 03:46:36,790
i also

5777
03:46:44,870 --> 03:46:42,399
feedback um please give it to us either

5778
03:46:46,550 --> 03:46:44,880
email or hear in the chat um it would

5779
03:46:49,510 --> 03:46:46,560
just be great to know

5780
03:46:51,269 --> 03:46:49,520
if this was a success in your opinion

5781
03:46:53,189 --> 03:46:51,279
any ways you think we can improve we are

5782
03:46:54,630 --> 03:46:53,199
looking to definitely improve anything

5783
03:46:57,910 --> 03:46:54,640
you thought that was missing that we

5784
03:47:00,229 --> 03:46:57,920
could help with an understanding um

5785
03:47:01,830 --> 03:47:00,239
we will keep the website live and it has

5786
03:47:03,269 --> 03:47:01,840
resources available there if you have

5787
03:47:07,750 --> 03:47:03,279
another question we're happy to add

5788
03:47:11,269 --> 03:47:09,830

yes and also

5789
03:47:12,630 --> 03:47:11,279
also um

5790
03:47:16,150 --> 03:47:12,640
you know we have kathy here as a

5791
03:47:18,630 --> 03:47:16,160
resource in terms of life detection and

5792
03:47:21,269 --> 03:47:18,640
you know we're very interested in

5793
03:47:24,229 --> 03:47:21,279
the translation of these science ideas

5794
03:47:27,189 --> 03:47:24,239
into what can fly on future missions and

5795
03:47:29,750 --> 03:47:27,199
uh it looks like we have you know robust

5796
03:47:32,150 --> 03:47:29,760
future in life detection missions to a

5797
03:47:33,189 --> 03:47:32,160
number of different bodies

5798
03:47:36,309 --> 03:47:33,199
certainly

5799
03:47:38,309 --> 03:47:36,319
some way to actually look for viruses

5800
03:47:40,469 --> 03:47:38,319
and identify them in some fashion or

5801
03:47:42,309 --> 03:47:40,479
virus-like things or what are they

5802
03:47:44,309 --> 03:47:42,319
called mobile elements or i don't care

5803
03:47:47,189 --> 03:47:44,319
what you call them doodads is what i

5804
03:47:48,790 --> 03:47:47,199
said on the chat um you know we're very

5805
03:47:51,030 --> 03:47:48,800
interested in figuring out how to do

5806
03:47:53,110 --> 03:47:51,040
that and all of these things take a very

5807
03:47:55,590 --> 03:47:53,120
long time to develop and figure out

5808
03:47:59,030 --> 03:47:55,600
right we still are really very much in

5809
03:48:00,630 --> 03:47:59,040
the infancy of uh actual life detection

5810
03:48:02,150 --> 03:48:00,640
rather than just looking for bags of

5811
03:48:03,990 --> 03:48:02,160
chemicals

5812
03:48:04,710 --> 03:48:04,000
which is you know part of the approach

5813
03:48:06,870 --> 03:48:04,720

but

5814
03:48:09,189 --> 03:48:06,880
we need more and so

5815
03:48:11,349 --> 03:48:09,199
folding in these other life forms as i

5816
03:48:12,950 --> 03:48:11,359
think of them their true life forms in

5817
03:48:14,950 --> 03:48:12,960
my view

5818
03:48:16,710 --> 03:48:14,960
i think is really very important and now

5819
03:48:19,349 --> 03:48:16,720
is the time to really get on board with

5820
03:48:20,550 --> 03:48:19,359
that so ideas for instruments ideas for

5821
03:48:22,550 --> 03:48:20,560
approaches

5822
03:48:24,309 --> 03:48:22,560
techniques techniques translate into

5823
03:48:26,950 --> 03:48:24,319
instruments techniques that exist that

5824
03:48:28,550 --> 03:48:26,960
haven't been used with viruses before or

5825
03:48:31,510 --> 03:48:28,560
that are used

5826
03:48:34,389 --> 03:48:31,520
on earth but are not part of our

5827
03:48:36,389 --> 03:48:34,399
pantheon of instrumentation that we're

5828
03:48:39,030 --> 03:48:36,399
already using on space

5829
03:48:41,990 --> 03:48:39,040
missions so all of that is very fluid

5830
03:48:43,429 --> 03:48:42,000
and we're dying to hear about that so uh

5831
03:48:46,710 --> 03:48:43,439
contact kathy

5832
03:48:48,070 --> 03:48:46,720
me both of us whatever any combination

5833
03:48:51,189 --> 03:48:48,080
of of us

5834
03:48:52,790 --> 03:48:51,199
and we will you know get back to you

5835
03:48:54,710 --> 03:48:52,800
yes maybe not immediately but we will

5836
03:48:58,389 --> 03:48:54,720
get back to you

5837
03:49:02,070 --> 03:49:00,070
were there any last minute questions

5838
03:49:10,469 --> 03:49:02,080

that anyone wanted to ask

5839
03:49:15,429 --> 03:49:12,950
okay my final part in comment virus is

5840
03:49:21,910 --> 03:49:17,590
viruses can be in rock that's what i

5841
03:49:26,710 --> 03:49:24,550
thanks everybody so much everybody

5842
03:49:29,030 --> 03:49:26,720
wonderful the rest of your evening or

5843
03:49:30,550 --> 03:49:29,040
day or morning or whatever time wherever

5844
03:49:32,790 --> 03:49:30,560
you happen to be

5845
03:49:34,389 --> 03:49:32,800
thank you marco thank you greg

5846
03:49:37,110 --> 03:49:34,399
thank you thank you for the nasa

5847
03:49:39,110 --> 03:49:37,120
technical staff that helped yes yes

5848
03:49:42,309 --> 03:49:39,120
flawless execution and several of the

5849
03:49:44,950 --> 03:49:42,319
comments are yeah fantastic super loved