Thank you!

Harrison B. Smith

Research Scientist
Earth-Life Science Institute,
Tokyo Institute of Technology
hbs@elsi.jp

Seeding Biochemistry on Other Worlds: Enceladus as a Case Study— H.B. Smith et. al. — Astrobiology, 2021

科研費
KAKENHI

wpi

ELSI
EARTH-LIFE SCIENCE INSTITUTE
TOKYO INSTITUTE OF TECHNOLOGY

ASU School of Earth and Space Exploration
Arizona State University

1
00:00:04,789 --> 00:00:02,790
hi my name is harrison smith and i want

2
00:00:06,630 --> 00:00:04,799
to talk to you today about my research

3
00:00:09,350 --> 00:00:06,640
the title is called seeding biochemistry

4
00:00:10,709 --> 00:00:09,360
in other words enceladus is a case study

5
00:00:12,470 --> 00:00:10,719
and the first thing i'd like to mention

6
00:00:14,789 --> 00:00:12,480
is that this work was done

7
00:00:16,870 --> 00:00:14,799
um not by myself but also with

8
00:00:18,630 --> 00:00:16,880
co-authors alexa drew jamboy and sarah

9
00:00:20,150 --> 00:00:18,640
walker

10
00:00:22,790 --> 00:00:20,160
so the first thing i want to talk about

11
00:00:25,830 --> 00:00:22,800
today is the concept of habitability

12
00:00:27,589 --> 00:00:25,840
which is pervasive in astrobiology

13
00:00:29,269 --> 00:00:27,599

and i want to talk about it because it's

14
00:00:32,549 --> 00:00:29,279
something which

15
00:00:35,190 --> 00:00:32,559
i'm a little bit passionate about and

16
00:00:37,750 --> 00:00:35,200
the idea that only environments like the

17
00:00:40,630 --> 00:00:37,760
one shown above where you have

18
00:00:41,910 --> 00:00:40,640
water on the surface of an environment

19
00:00:43,430 --> 00:00:41,920
these are currently the only types of

20
00:00:45,190 --> 00:00:43,440
environments which are currently labeled

21
00:00:47,029 --> 00:00:45,200
as being habitable

22
00:00:49,590 --> 00:00:47,039
but of course we know that that's not

23
00:00:51,510 --> 00:00:49,600
necessarily the case

24
00:00:53,430 --> 00:00:51,520
there are places in our solar system for

25
00:00:56,229 --> 00:00:53,440
example on enceladus where

26
00:00:57,510 --> 00:00:56,239
a subsurface ocean seems like a

27
00:00:58,549 --> 00:00:57,520
completely reasonable habitable

28
00:01:00,229 --> 00:00:58,559
environment

29
00:01:03,510 --> 00:01:00,239
but this language that we use to define

30
00:01:05,509 --> 00:01:03,520
habitability kind of pigeonholes us into

31
00:01:07,030 --> 00:01:05,519
thinking about the concept in a kind of

32
00:01:10,390 --> 00:01:07,040
biased way

33
00:01:11,670 --> 00:01:10,400
even in places like earth and antarctica

34
00:01:14,310 --> 00:01:11,680
for example

35
00:01:15,990 --> 00:01:14,320
the south pole lake bostock is a

36
00:01:19,990 --> 00:01:16,000
subglacial lake

37
00:01:23,190 --> 00:01:20,000
where life has been found and

38
00:01:26,230 --> 00:01:23,200

this is an interesting fact because this

39
00:01:29,670 --> 00:01:26,240
environment on earth would be deemed

40
00:01:32,870 --> 00:01:29,680
non-inhabitable by the definition and

41
00:01:35,190 --> 00:01:32,880
that's pervasive in astrobiology

42
00:01:36,230 --> 00:01:35,200
and one of the reasons i'm talking about

43
00:01:39,830 --> 00:01:36,240
this for so long

44
00:01:40,870 --> 00:01:39,840
is because um i think habitability is a

45
00:01:44,310 --> 00:01:40,880
concept which

46
00:01:45,270 --> 00:01:44,320
is hard to quantify um it's kind of

47
00:01:48,389 --> 00:01:45,280
specific

48
00:01:49,590 --> 00:01:48,399
to the environmental context and you

49
00:01:53,590 --> 00:01:49,600
remove the concept

50
00:01:55,510 --> 00:01:53,600
of individual organisms or types of life

51
00:01:58,389 --> 00:01:55,520
when you talk about habitability

52
00:01:59,990 --> 00:01:58,399
it's implied but it's not explicit

53
00:02:02,709 --> 00:02:00,000
instead i think it makes more sense to

54
00:02:06,149 --> 00:02:02,719
talk about the concept of viability

55
00:02:09,430 --> 00:02:06,159
and viability is specific to the type

56
00:02:11,510 --> 00:02:09,440
of life that you're interested in

57
00:02:13,510 --> 00:02:11,520
so you can actually quantify whether or

58
00:02:15,350 --> 00:02:13,520
not a particular type of organism or

59
00:02:17,589 --> 00:02:15,360
particular type of life

60
00:02:19,030 --> 00:02:17,599
is viable in a given environment and

61
00:02:22,869 --> 00:02:19,040
it's a little bit more

62
00:02:25,750 --> 00:02:22,879
rigidly explorable scientifically than

63
00:02:29,110 --> 00:02:25,760

the concept of habitability which is at

64
00:02:30,869 --> 00:02:29,120
some levels intentionally vague

65
00:02:32,630 --> 00:02:30,879
so the question that i have is kind of

66
00:02:34,550 --> 00:02:32,640
how can we predict if earth life can be

67
00:02:37,350 --> 00:02:34,560
viable on other planets

68
00:02:38,150 --> 00:02:37,360
um and i'm going to list a kind of set

69
00:02:41,110 --> 00:02:38,160
of features

70
00:02:43,270 --> 00:02:41,120
that it'd be nice um to have when we

71
00:02:44,869 --> 00:02:43,280
want to test this question

72
00:02:46,309 --> 00:02:44,879
and they're going to be biased by the

73
00:02:47,589 --> 00:02:46,319
fact that i'm a computational

74
00:02:50,550 --> 00:02:47,599
estrobiologist

75
00:02:52,309 --> 00:02:50,560
so ideally if you want a predictive

76
00:02:53,750 --> 00:02:52,319
earth life as well on their planets

77
00:02:55,430 --> 00:02:53,760
we don't have to require growing

78
00:02:58,229 --> 00:02:55,440
anything in the lab it's more easily

79
00:02:59,750 --> 00:02:58,239
scalable to ask this question this way

80
00:03:01,430 --> 00:02:59,760
it'd be nice to vary kind of the big

81
00:03:03,190 --> 00:03:01,440
unknowns so things like catalytic

82
00:03:06,630 --> 00:03:03,200
properties of organisms

83
00:03:08,550 --> 00:03:06,640
basically the reactions which are

84
00:03:10,869 --> 00:03:08,560
catalyzed by the enzymes encoded in the

85
00:03:12,470 --> 00:03:10,879
organism's genomes

86
00:03:14,390 --> 00:03:12,480
it'd be nice people vary the growth or

87
00:03:16,710 --> 00:03:14,400
sustainability requirements of organisms

88
00:03:18,630 --> 00:03:16,720

as we learn more about those things

89
00:03:20,070 --> 00:03:18,640
and having a first order model is

90
00:03:21,589 --> 00:03:20,080
probably good enough we don't want to

91
00:03:23,190 --> 00:03:21,599
really over constrain ourselves with

92
00:03:24,630 --> 00:03:23,200
thinking about things like reactions

93
00:03:26,710 --> 00:03:24,640
stoichiometries

94
00:03:29,750 --> 00:03:26,720
kinetic constants more environmental

95
00:03:32,869 --> 00:03:29,760
concentrations where we don't have to

96
00:03:34,470 --> 00:03:32,879
so the way that i approach this

97
00:03:36,789 --> 00:03:34,480
is using this technique called network

98
00:03:41,030 --> 00:03:36,799
expansion network expansion

99
00:03:42,630 --> 00:03:41,040
ideally gets around all of the issues or

100
00:03:44,149 --> 00:03:42,640
has the advantages that were listed on

101
00:03:45,910 --> 00:03:44,159
the last slide so

102
00:03:47,830 --> 00:03:45,920
the concept behind network expansion

103
00:03:49,110 --> 00:03:47,840
which is a computational approach

104
00:03:51,990 --> 00:03:49,120
is that you initialize possible

105
00:03:55,030 --> 00:03:52,000
reactions based on organism's genome or

106
00:03:58,550 --> 00:03:55,040
possibly a community's metagenome

107
00:03:59,270 --> 00:03:58,560
you look at the compounds which are

108
00:04:01,589 --> 00:03:59,280
available

109
00:04:03,270 --> 00:04:01,599
in the environment so in the case of

110
00:04:05,110 --> 00:04:03,280
enceladus for example these are

111
00:04:06,630 --> 00:04:05,120
observed based on spacecraft

112
00:04:08,390 --> 00:04:06,640
measurements

113
00:04:09,990 --> 00:04:08,400

you come up with a target you can come

114
00:04:11,429 --> 00:04:10,000
up with a target compound set so

115
00:04:14,630 --> 00:04:11,439
basically you define

116
00:04:16,229 --> 00:04:14,640
predefined the compounds which you think

117
00:04:18,629 --> 00:04:16,239
are required

118
00:04:21,030 --> 00:04:18,639
to define an organism an organism's

119
00:04:22,629 --> 00:04:21,040
viability

120
00:04:25,350 --> 00:04:22,639
in order to actually execute the network

121
00:04:26,790 --> 00:04:25,360
expansion what you do is you check

122
00:04:28,629 --> 00:04:26,800
if reactions are possible based on

123
00:04:30,550 --> 00:04:28,639
compounds in the environment so

124
00:04:32,150 --> 00:04:30,560
you say this organism can catalyze this

125
00:04:35,430 --> 00:04:32,160
set of reactions

126
00:04:36,950 --> 00:04:35,440
particular compounds

127
00:04:40,390 --> 00:04:36,960
and if these compounds are available we

128
00:04:43,270 --> 00:04:40,400
assume that the reaction is possible

129
00:04:43,909 --> 00:04:43,280
and what you do then is you add the

130
00:04:45,670 --> 00:04:43,919
products

131
00:04:47,830 --> 00:04:45,680
of whatever reactions are possible back

132
00:04:48,390 --> 00:04:47,840
into the seed set that you used to start

133
00:04:50,310 --> 00:04:48,400
out with

134
00:04:52,950 --> 00:04:50,320
and what you do what ends up happening

135
00:04:54,830 --> 00:04:52,960
is you start collecting compounds over

136
00:04:57,590 --> 00:04:54,840
time over the iteration of this

137
00:04:59,189 --> 00:04:57,600
simulation you repeat this process

138
00:05:00,710 --> 00:04:59,199

until no new compounds can be produced

139
00:05:01,430 --> 00:05:00,720
and at the end of the simulation you

140
00:05:03,110 --> 00:05:01,440
check

141
00:05:04,870 --> 00:05:03,120
how many target compounds are produced

142
00:05:09,670 --> 00:05:04,880
based on this predefined list

143
00:05:11,909 --> 00:05:09,680
that we're using to measure viability

144
00:05:14,070 --> 00:05:11,919
so this this meets all the requirements

145
00:05:16,790 --> 00:05:14,080
that i was listing before

146
00:05:19,189 --> 00:05:16,800
conveniently enough um and i want to

147
00:05:22,710 --> 00:05:19,199
point out that this is a technique which

148
00:05:24,310 --> 00:05:22,720
has some history and being used

149
00:05:25,749 --> 00:05:24,320
in particular studies in the origin of

150
00:05:26,629 --> 00:05:25,759
life but this is a really prominent

151
00:05:28,790 --> 00:05:26,639
paper that it was

152
00:05:29,909 --> 00:05:28,800
recently used in from a few years back

153
00:05:31,430 --> 00:05:29,919
looking at how

154
00:05:33,909 --> 00:05:31,440
we could use network expansion kind of

155
00:05:35,430 --> 00:05:33,919
at a global level to understand

156
00:05:37,430 --> 00:05:35,440
uh what reactions might have been

157
00:05:40,790 --> 00:05:37,440
possible in a world before

158
00:05:41,670 --> 00:05:40,800
phosphate okay so to kind of jump to the

159
00:05:44,070 --> 00:05:41,680
punch line

160
00:05:46,550 --> 00:05:44,080
when i asked the question could earth

161
00:05:49,670 --> 00:05:46,560
biochemistry be viable in enceladus

162
00:05:51,350 --> 00:05:49,680
uh if the answer is yes then a

163
00:05:53,590 --> 00:05:51,360

natural follow-up question might be kind

164
00:05:56,469 --> 00:05:53,600
of what organisms or combinational organ

165
00:05:57,909 --> 00:05:56,479
organisms are viable and if no what

166
00:06:00,550 --> 00:05:57,919
things would you have to add in order to

167
00:06:02,070 --> 00:06:00,560
make organisms viable

168
00:06:05,270 --> 00:06:02,080
and again the viability is something

169
00:06:07,189 --> 00:06:05,280
that we pre-depo predefined based on a

170
00:06:09,110 --> 00:06:07,199
list of target compounds and this comes

171
00:06:11,029 --> 00:06:09,120
from the literature and comes from

172
00:06:12,469 --> 00:06:11,039
studies on organisms and what types of

173
00:06:13,830 --> 00:06:12,479
environments they can grow in and

174
00:06:15,110 --> 00:06:13,840
what things they produce which are

175
00:06:17,270 --> 00:06:15,120
necessary for them to continue to

176
00:06:20,710 --> 00:06:17,280
survive

177
00:06:23,430 --> 00:06:20,720
okay so again here's just a reminder

178
00:06:24,550 --> 00:06:23,440
of the initialization process and going

179
00:06:26,550 --> 00:06:24,560
over a little more detail

180
00:06:28,710 --> 00:06:26,560
in kind of where the genomic data comes

181
00:06:30,550 --> 00:06:28,720
from we look at these genomic databases

182
00:06:32,230 --> 00:06:30,560
for example jgi which i show a

183
00:06:33,110 --> 00:06:32,240
screenshot of here we also get some data

184
00:06:35,830 --> 00:06:33,120
from

185
00:06:37,990 --> 00:06:35,840
phimat2 we're picking a kind of

186
00:06:39,990 --> 00:06:38,000
reasonable starting point for

187
00:06:41,749 --> 00:06:40,000
organisms for enceladus so we look at

188
00:06:44,230 --> 00:06:41,759

high ph

189
00:06:45,029 --> 00:06:44,240
we found about 307 organisms mostly

190
00:06:47,110 --> 00:06:45,039
bacteria

191
00:06:49,749 --> 00:06:47,120
that meet the condition of living these

192
00:06:52,710 --> 00:06:49,759
high ph environments

193
00:06:53,990 --> 00:06:52,720
um yes what you can do is you can go in

194
00:06:56,230 --> 00:06:54,000
and you can select an

195
00:06:57,589 --> 00:06:56,240
organism where you can select a

196
00:07:00,309 --> 00:06:57,599
metagenome

197
00:07:02,629 --> 00:07:00,319
and you can get the list of ecs which

198
00:07:06,309 --> 00:07:02,639
are enzymes associated with an organism

199
00:07:07,830 --> 00:07:06,319
or a community and you can tie these

200
00:07:10,230 --> 00:07:07,840
seeds to particular reactions through a

201
00:07:12,070 --> 00:07:10,240
database called keg

202
00:07:14,870 --> 00:07:12,080
and you can turn these reactions into a

203
00:07:16,710 --> 00:07:14,880
network that looks something like this

204
00:07:18,469 --> 00:07:16,720
um okay so that's how we initialize the

205
00:07:20,230 --> 00:07:18,479
possible reactions

206
00:07:22,230 --> 00:07:20,240
like i said to initialize the compounds

207
00:07:24,390 --> 00:07:22,240
what you do is you take um

208
00:07:25,990 --> 00:07:24,400
observations from enceladus taken from

209
00:07:26,469 --> 00:07:26,000
spacecraft measurements and this gives

210
00:07:29,830 --> 00:07:26,479
you the

211
00:07:33,029 --> 00:07:29,840
compound set the target set comes again

212
00:07:34,710 --> 00:07:33,039
from previous literature okay so here's

213
00:07:37,830 --> 00:07:34,720

jumping to the punch line

214
00:07:38,309 --> 00:07:37,840
um we find that none of the prokaryotes

215
00:07:41,110 --> 00:07:38,319
that

216
00:07:42,710 --> 00:07:41,120
we studied were viable even when you add

217
00:07:44,469 --> 00:07:42,720
phosphate phosphate is something which

218
00:07:46,469 --> 00:07:44,479
is missing in detections that enceladus

219
00:07:47,909 --> 00:07:46,479
people think mostly just because

220
00:07:49,670 --> 00:07:47,919
of instrument sensitivity and low

221
00:07:51,749 --> 00:07:49,680
concentrations but not because it's

222
00:07:54,309 --> 00:07:51,759
completely absent

223
00:07:55,909 --> 00:07:54,319
what we find here this is just going

224
00:07:57,510 --> 00:07:55,919
into a little bit more detail on kind of

225
00:07:58,629 --> 00:07:57,520
how we're defining the viability of an

226
00:08:06,469 --> 00:07:58,639
organism

227
00:08:07,830 --> 00:08:06,479
all these 65 global target compounds

228
00:08:09,670 --> 00:08:07,840
that they could possibly produce in

229
00:08:12,390 --> 00:08:09,680
their their

230
00:08:14,150 --> 00:08:12,400
metabolic network and so instead what we

231
00:08:14,950 --> 00:08:14,160
do is we don't penalize the organism for

232
00:08:16,710 --> 00:08:14,960
not having

233
00:08:18,469 --> 00:08:16,720
the possibility to produce those

234
00:08:19,990 --> 00:08:18,479
compounds based on the reactions that

235
00:08:22,710 --> 00:08:20,000
are coded for

236
00:08:24,309 --> 00:08:22,720
um through the proteins that are coded

237
00:08:26,070 --> 00:08:24,319
for in the genome instead what we do is

238
00:08:27,830 --> 00:08:26,080

we take the overlapping list

239
00:08:29,189 --> 00:08:27,840
of all the compounds which an organism

240
00:08:29,749 --> 00:08:29,199
could possibly produce with that target

241
00:08:32,709 --> 00:08:29,759
list

242
00:08:33,509 --> 00:08:32,719
and so for example this bacteria here

243
00:08:36,070 --> 00:08:33,519
its goal

244
00:08:37,829 --> 00:08:36,080
is producing 60 target compounds not the

245
00:08:41,430 --> 00:08:37,839
65 because 65

246
00:08:43,509 --> 00:08:41,440
and 4 aren't possible

247
00:08:45,190 --> 00:08:43,519
same example here is the other extreme

248
00:08:48,790 --> 00:08:45,200
so there's ikea that

249
00:08:51,030 --> 00:08:48,800
only overlap with 32 of these 65 target

250
00:08:52,550 --> 00:08:51,040
compounds and if it can produce all 32

251
00:08:55,670 --> 00:08:52,560
of these target compounds we say that

252
00:08:58,070 --> 00:08:55,680
it's 100 viable

253
00:08:58,870 --> 00:08:58,080
um so then the next follow-up question

254
00:09:00,310 --> 00:08:58,880
becomes

255
00:09:02,230 --> 00:09:00,320
what compounds would be necessary to

256
00:09:04,230 --> 00:09:02,240
enhance organismal viability if you

257
00:09:05,750 --> 00:09:04,240
could add particular compounds to the

258
00:09:07,190 --> 00:09:05,760
environment and solids could you

259
00:09:09,430 --> 00:09:07,200
generate

260
00:09:10,790 --> 00:09:09,440
environments which allow these organisms

261
00:09:13,190 --> 00:09:10,800
to be viable

262
00:09:14,389 --> 00:09:13,200
so what we do here is we generate what

263
00:09:16,630 --> 00:09:14,399

are called irreducible

264
00:09:18,389 --> 00:09:16,640
compound seed sets for each organism so

265
00:09:19,350 --> 00:09:18,399
basically working backwards to figure

266
00:09:22,790 --> 00:09:19,360
out

267
00:09:24,389 --> 00:09:22,800
um what are the seeds that need to be an

268
00:09:27,590 --> 00:09:24,399
environment in order

269
00:09:30,630 --> 00:09:27,600
for an organism to produce the full

270
00:09:32,230 --> 00:09:30,640
um intersecting set of target compounds

271
00:09:33,269 --> 00:09:32,240
necessary for viability

272
00:09:35,190 --> 00:09:33,279
and here's some details in that

273
00:09:35,990 --> 00:09:35,200
algorithm you can pause the video and

274
00:09:39,350 --> 00:09:36,000
look back at it later

275
00:09:40,630 --> 00:09:39,360
if you'd like so asking the questions

276
00:09:43,030 --> 00:09:40,640
which compounds would be necessary to

277
00:09:44,949 --> 00:09:43,040
enhance organizational viability

278
00:09:46,710 --> 00:09:44,959
if we just look at a particular organism

279
00:09:48,710 --> 00:09:46,720
there's a set of organisms and ask

280
00:09:50,230 --> 00:09:48,720
kind of how many additional seeds are

281
00:09:51,829 --> 00:09:50,240
required for viability

282
00:09:53,829 --> 00:09:51,839
and we ask this question independent of

283
00:09:55,430 --> 00:09:53,839
environment we get a result like this

284
00:09:57,030 --> 00:09:55,440
on the x-axis is the number of seed

285
00:09:59,829 --> 00:09:57,040
compounds and the y-axis is the

286
00:10:02,630 --> 00:09:59,839
molecular weight so kind of a proxy for

287
00:10:03,190 --> 00:10:02,640
the complexity of the compounds and you

288
00:10:05,190 --> 00:10:03,200

see that

289
00:10:06,710 --> 00:10:05,200
in general there are archaea for example

290
00:10:10,069 --> 00:10:06,720
on the left hand that plot

291
00:10:13,350 --> 00:10:10,079
that only need 50 uh kind of

292
00:10:14,630 --> 00:10:13,360
compounds for viability um

293
00:10:16,550 --> 00:10:14,640
if they're the right compounds and these

294
00:10:18,470 --> 00:10:16,560
tend to be very complex things

295
00:10:20,550 --> 00:10:18,480
on average based on what you can see the

296
00:10:22,790 --> 00:10:20,560
y-axis there

297
00:10:24,069 --> 00:10:22,800
and the dash the vertical dashed line is

298
00:10:25,910 --> 00:10:24,079
the um

299
00:10:27,670 --> 00:10:25,920
just number of compounds which are

300
00:10:29,590 --> 00:10:27,680
observed in the environment on enceladus

301
00:10:32,870 --> 00:10:29,600
that were incorporated into our seed set

302
00:10:34,790 --> 00:10:32,880
and so you can also break this plot by

303
00:10:36,389 --> 00:10:34,800
not looking at the seed sets with the

304
00:10:37,829 --> 00:10:36,399
fewest number of compounds but instead

305
00:10:39,269 --> 00:10:37,839
looking at seed sets with lowest mean

306
00:10:40,870 --> 00:10:39,279
molecular weight so if we want to kind

307
00:10:43,030 --> 00:10:40,880
of bias our results for

308
00:10:44,630 --> 00:10:43,040
simpler things you notice that now the

309
00:10:48,150 --> 00:10:44,640
left-hand side of the plot

310
00:10:50,710 --> 00:10:48,160
the archaea which requires the fewest

311
00:10:52,389 --> 00:10:50,720
number of seeds is actually 18 instead

312
00:10:54,710 --> 00:10:52,399
of 15 because here we're looking

313
00:10:56,310 --> 00:10:54,720

specifically at things that are less

314
00:10:57,990 --> 00:10:56,320
complex

315
00:10:59,990 --> 00:10:58,000
maybe a more relevant question though is

316
00:11:01,430 --> 00:11:00,000
to ask based on what we already know in

317
00:11:03,190 --> 00:11:01,440
enceladus

318
00:11:04,790 --> 00:11:03,200
what additional compounds would you need

319
00:11:05,990 --> 00:11:04,800
to add to the environment so not just

320
00:11:07,590 --> 00:11:06,000
not thinking environmentally

321
00:11:09,190 --> 00:11:07,600
agnostically but thinking particularly

322
00:11:10,310 --> 00:11:09,200
to the environment that we're interested

323
00:11:12,790 --> 00:11:10,320
in here

324
00:11:14,230 --> 00:11:12,800
so what are the seed sets with the fused

325
00:11:16,470 --> 00:11:14,240
compounds how many

326
00:11:18,310 --> 00:11:16,480
seeds do we need to add to those

327
00:11:20,389 --> 00:11:18,320
environments and we see that

328
00:11:21,750 --> 00:11:20,399
you have to add a minimum of seven seeds

329
00:11:23,509 --> 00:11:21,760
for some bacteria

330
00:11:25,110 --> 00:11:23,519
but it the average seems to be something

331
00:11:27,750 --> 00:11:25,120
more like about

332
00:11:29,670 --> 00:11:27,760
adding about 17 seeds and we can make

333
00:11:32,470 --> 00:11:29,680
the same plot where instead of

334
00:11:34,150 --> 00:11:32,480
looking at the seed sets which produce

335
00:11:35,829 --> 00:11:34,160
viability for the

336
00:11:37,350 --> 00:11:35,839
fewest number of compounds instead

337
00:11:41,750 --> 00:11:37,360
looking at kind of the

338
00:11:43,990 --> 00:11:41,760

least complex environments that produce

339
00:11:45,509 --> 00:11:44,000
viability in different organisms and you

340
00:11:48,150 --> 00:11:45,519
see these require slightly

341
00:11:49,590 --> 00:11:48,160
more compounds so kind of the takeaway

342
00:11:51,590 --> 00:11:49,600
is for viability you can get away with

343
00:11:53,110 --> 00:11:51,600
fewer compounds which are more complex

344
00:11:54,790 --> 00:11:53,120
or much larger number of simpler

345
00:11:55,990 --> 00:11:54,800
compounds and depends kind of what your

346
00:11:59,509 --> 00:11:56,000
constraints are going into

347
00:12:01,269 --> 00:11:59,519
things um we can do some other just kind

348
00:12:03,110 --> 00:12:01,279
of looking at interesting statistics of

349
00:12:05,110 --> 00:12:03,120
the data for

350
00:12:06,629 --> 00:12:05,120
viability for organisms so for example

351
00:12:08,310 --> 00:12:06,639
we can look at the jaccard index which

352
00:12:09,350 --> 00:12:08,320
is kind of the self-similarity of the

353
00:12:13,030 --> 00:12:09,360
seed sets

354
00:12:13,430 --> 00:12:13,040
for each organism and so what we find is

355
00:12:15,269 --> 00:12:13,440
even

356
00:12:16,790 --> 00:12:15,279
within organisms that irreducible seed

357
00:12:18,230 --> 00:12:16,800
sets can vary considerably they only

358
00:12:19,910 --> 00:12:18,240
have about 10 to 25

359
00:12:23,430 --> 00:12:19,920
overlap which is what that y-axis is

360
00:12:26,150 --> 00:12:23,440
there so there's many paths to viability

361
00:12:27,509 --> 00:12:26,160
and um the similarity between seed sets

362
00:12:29,750 --> 00:12:27,519
tends to be a lot higher

363
00:12:31,430 --> 00:12:29,760

when you compare archaea to archaea than

364
00:12:32,629 --> 00:12:31,440
than when you compare keto bacteria for

365
00:12:35,829 --> 00:12:32,639
instance

366
00:12:37,350 --> 00:12:35,839
um so that's not too surprising so

367
00:12:38,629 --> 00:12:37,360
kind of revisiting the question which

368
00:12:40,230 --> 00:12:38,639
compounds would be necessary to enhance

369
00:12:42,550 --> 00:12:40,240
organized liability

370
00:12:43,750 --> 00:12:42,560
um kind of from a very high level we

371
00:12:46,949 --> 00:12:43,760
find that

372
00:12:49,030 --> 00:12:46,959
uh prokaryotes can be viable using

373
00:12:50,310 --> 00:12:49,040
fewer compounds than what's on enceladus

374
00:12:52,230 --> 00:12:50,320
but if we're really

375
00:12:53,750 --> 00:12:52,240
curious in in terms of the particular

376
00:12:54,710 --> 00:12:53,760
environment of enceladus you still have

377
00:12:57,030 --> 00:12:54,720
to add

378
00:12:59,110 --> 00:12:57,040
at least seven compounds usually more

379
00:13:00,870 --> 00:12:59,120
especially if you limit yourself to

380
00:13:02,389 --> 00:13:00,880
simpler compounds and then to kind of

381
00:13:04,389 --> 00:13:02,399
answer the ultimate question of

382
00:13:06,629 --> 00:13:04,399
course earth biochemistry would be up on

383
00:13:08,790 --> 00:13:06,639
another planet

384
00:13:10,629 --> 00:13:08,800
i would say yes but with a big caveat

385
00:13:12,629 --> 00:13:10,639
which is you know

386
00:13:13,670 --> 00:13:12,639
if complex compounds are present in the

387
00:13:15,269 --> 00:13:13,680
environment or

388
00:13:16,870 --> 00:13:15,279

the organisms have sufficiently

389
00:13:18,629 --> 00:13:16,880
extensive catalytic capabilities we

390
00:13:20,949 --> 00:13:18,639
didn't limit ourselves to

391
00:13:22,870 --> 00:13:20,959
organisms which are known to be found in

392
00:13:24,629 --> 00:13:22,880
these high ph environments

393
00:13:27,030 --> 00:13:24,639
we might be able to find organisms which

394
00:13:28,230 --> 00:13:27,040
are viable or if we assume that there's

395
00:13:29,430 --> 00:13:28,240
more complex compounds in the

396
00:13:30,230 --> 00:13:29,440
environment that haven't been observed

397
00:13:32,310 --> 00:13:30,240
yet

398
00:13:34,389 --> 00:13:32,320
there may be organisms that are viable

399
00:13:36,150 --> 00:13:34,399
but out of the organisms that we tested

400
00:13:38,629 --> 00:13:36,160
none were viable with with what we know

401
00:13:40,389 --> 00:13:38,639
to be present and that's all it is

402
00:13:42,870 --> 00:13:40,399
so kind of the takeaway here is that

403
00:13:44,949 --> 00:13:42,880
this network expansion technique can be

404
00:13:46,550 --> 00:13:44,959
um useful for conservatively

405
00:13:50,069 --> 00:13:46,560
conservatively bounding

406
00:13:51,670 --> 00:13:50,079
environments for viability of organisms

407
00:13:53,590 --> 00:13:51,680
and that is implications for things like

408
00:13:55,030 --> 00:13:53,600
accidental panspermia and it could even

409
00:13:57,670 --> 00:13:55,040
help guide things like

410
00:13:59,910 --> 00:13:57,680
directed panspermia thinking about the

411
00:14:01,590 --> 00:13:59,920
big picture questions in astrobiology

412
00:14:02,550 --> 00:14:01,600
and here's some follow-up questions of

413
00:14:04,470 --> 00:14:02,560

of things that we're interested in

414
00:14:06,150 --> 00:14:04,480
exploring in future work related to this

415
00:14:06,470 --> 00:14:06,160
so thinking about ecosystems thinking

416
00:14:08,310 --> 00:14:06,480
about

417
00:14:10,310 --> 00:14:08,320
more details in the environment more

418
00:14:12,389 --> 00:14:10,320
constraints related to thermodynamics

419
00:14:13,670 --> 00:14:12,399
and concentrations for instance

420
00:14:15,430 --> 00:14:13,680
i'd just like to thank everyone today

421
00:14:17,509 --> 00:14:15,440
for listening um the

422
00:14:19,670 --> 00:14:17,519
reference for the paper that this is the

423
00:14:21,910 --> 00:14:19,680
talk is based on is on the right there

424
00:14:23,350 --> 00:14:21,920
and i'd like to thank um kenny for

425
00:14:25,990 --> 00:14:23,360
sponsoring this work and

426
00:14:27,910 --> 00:14:26,000
the wpialc and the work that i did to