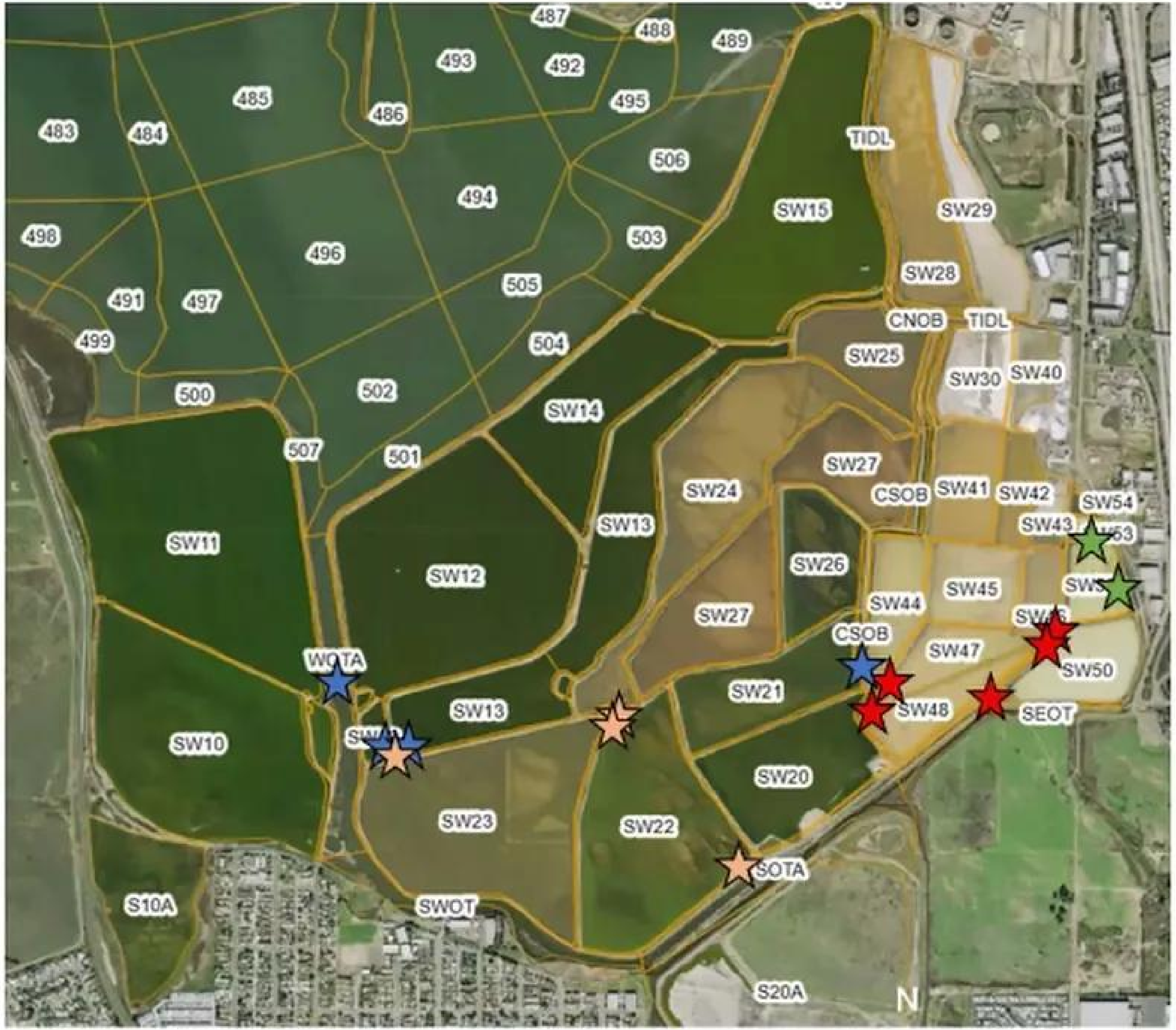
South Bay Salt Works!



1
00:00:04,230 --> 00:00:02,389
good morning or good afternoon or good

2
00:00:06,150 --> 00:00:04,240
night depending on where you are it's my

3
00:00:08,870 --> 00:00:06,160
pleasure to talk to you today about the

4
00:00:10,950 --> 00:00:08,880
flexible genomes of halo archaea that

5
00:00:12,950 --> 00:00:10,960
live in hypersaline brines

6
00:00:15,430 --> 00:00:12,960
so i want to start off by introducing

7
00:00:16,870 --> 00:00:15,440
you to our field site which i find quite

8
00:00:18,710 --> 00:00:16,880
beautiful this is the south base salt

9
00:00:21,269 --> 00:00:18,720
works and it's a salt harvesting

10
00:00:22,870 --> 00:00:21,279
facility here in southern california in

11
00:00:24,390 --> 00:00:22,880
the san diego bay

12
00:00:26,390 --> 00:00:24,400
and so as you can see here in the top

13
00:00:28,950 --> 00:00:26,400

right we have the san diego bay

14
00:00:31,109 --> 00:00:28,960
um and that's the source of the the

15
00:00:33,830 --> 00:00:31,119
water which is channeled through a

16
00:00:36,549 --> 00:00:33,840
series of evaporative ponds

17
00:00:37,510 --> 00:00:36,559
within the the salt works

18
00:00:39,510 --> 00:00:37,520
and

19
00:00:40,549 --> 00:00:39,520
in those ponds the

20
00:00:43,750 --> 00:00:40,559
salts

21
00:00:47,029 --> 00:00:43,760
begin to accumulate via vapor

22
00:00:49,029 --> 00:00:47,039
and eventually they they reach

23
00:00:50,630 --> 00:00:49,039
saturation and they begin to precipitate

24
00:00:53,029 --> 00:00:50,640
at that point

25
00:00:54,869 --> 00:00:53,039
and so they can be harvested for

26
00:00:56,549 --> 00:00:54,879
commercial purposes

27
00:00:58,549 --> 00:00:56,559
but in the meantime

28
00:01:02,069 --> 00:00:58,559
these hyper saline brines serve as a

29
00:01:04,710 --> 00:01:02,079
habitat for a surprisingly diverse array

30
00:01:08,390 --> 00:01:04,720
of microbial life

31
00:01:10,789 --> 00:01:08,400
including uh the the halo archaea which

32
00:01:13,109 --> 00:01:10,799
are the um

33
00:01:15,590 --> 00:01:13,119
the organisms that produce these

34
00:01:18,870 --> 00:01:15,600
characteristic carotenoid pigments that

35
00:01:21,749 --> 00:01:18,880
that produce this beautiful red color um

36
00:01:23,749 --> 00:01:21,759
that's common in hyper saline systems

37
00:01:26,070 --> 00:01:23,759
um and so this field site is actually a

38
00:01:28,469 --> 00:01:26,080

really good analog for ancient mars as

39
00:01:30,550 --> 00:01:28,479
the planet transitioned from wet to dry

40
00:01:31,910 --> 00:01:30,560
because as you can see here on the left

41
00:01:33,190 --> 00:01:31,920
side of the map

42
00:01:35,190 --> 00:01:33,200
we have

43
00:01:38,550 --> 00:01:35,200
approximately seawater

44
00:01:42,230 --> 00:01:38,560
to hypersaline brines here on the right

45
00:01:43,830 --> 00:01:42,240
and we did manage to sample over over

46
00:01:45,510 --> 00:01:43,840
two field efforts we did manage to

47
00:01:48,310 --> 00:01:45,520
sample um

48
00:01:51,350 --> 00:01:48,320
an extensive gradient from from seawater

49
00:01:55,510 --> 00:01:51,360
to hyper saline brines uh and so we

50
00:01:58,630 --> 00:01:55,520
covered a broad gradient in 2019

51
00:01:59,830 --> 00:01:58,640
and in 2020 we decided to focus actually

52
00:02:02,069 --> 00:01:59,840
on

53
00:02:03,670 --> 00:02:02,079
this part of the gradient to fill in

54
00:02:06,789 --> 00:02:03,680
some of these gaps

55
00:02:08,469 --> 00:02:06,799
and so we sampled extensively

56
00:02:11,350 --> 00:02:08,479
between

57
00:02:13,910 --> 00:02:11,360
seawater here on the left to

58
00:02:16,070 --> 00:02:13,920
brines where the the sodium chloride has

59
00:02:19,110 --> 00:02:16,080
just reached saturation and is just

60
00:02:21,990 --> 00:02:19,120
beginning to precipitate out as halite

61
00:02:24,309 --> 00:02:22,000
and so here on the on the the bottom on

62
00:02:27,510 --> 00:02:24,319
the x-axis if you don't know water

63
00:02:29,910 --> 00:02:27,520

activity is a measure of um

64
00:02:33,190 --> 00:02:29,920
the free water molecules in a solution

65
00:02:34,949 --> 00:02:33,200
and so a water activity of one is as

66
00:02:37,830 --> 00:02:34,959
pure water

67
00:02:39,190 --> 00:02:37,840
and down here around point seven is

68
00:02:41,430 --> 00:02:39,200
beginning to be

69
00:02:44,309 --> 00:02:41,440
quite hyper saline brines

70
00:02:46,309 --> 00:02:44,319
and in our first field effort we did

71
00:02:48,710 --> 00:02:46,319
sample all the way down to

72
00:02:50,550 --> 00:02:48,720
water activity of 0.4

73
00:02:52,790 --> 00:02:50,560
which is well below the known water

74
00:02:55,430 --> 00:02:52,800
activity limit of life

75
00:02:59,509 --> 00:02:55,440
um and so i will talk a little bit about

76
00:03:01,509 --> 00:02:59,519
our um results from our uh a micro

77
00:03:03,990 --> 00:03:01,519
uh environmental investigation of the

78
00:03:07,190 --> 00:03:04,000
diversity within these brines

79
00:03:09,030 --> 00:03:07,200
from our 2019

80
00:03:10,630 --> 00:03:09,040
and this study was actually just

81
00:03:12,390 --> 00:03:10,640
published recently in environmental

82
00:03:14,550 --> 00:03:12,400
microbiology

83
00:03:17,110 --> 00:03:14,560
our key findings here

84
00:03:18,949 --> 00:03:17,120
was that this microbial succession from

85
00:03:20,869 --> 00:03:18,959
from seawater to

86
00:03:23,589 --> 00:03:20,879
presumably sterile brines is actually

87
00:03:26,470 --> 00:03:23,599
characterized by three phases

88
00:03:28,710 --> 00:03:26,480

stage one which is um between water

89
00:03:31,190 --> 00:03:28,720
activity of one and zero point seven

90
00:03:34,070 --> 00:03:31,200
which is already starting to be a

91
00:03:36,789 --> 00:03:34,080
hyper saline brine

92
00:03:39,750 --> 00:03:36,799
this stage is is characterized by a

93
00:03:41,110 --> 00:03:39,760
surge of of archaeal

94
00:03:43,190 --> 00:03:41,120
microdiversity and we called it

95
00:03:46,630 --> 00:03:43,200
microdiversity because

96
00:03:48,949 --> 00:03:46,640
here on the left we have a an estimate

97
00:03:51,110 --> 00:03:48,959
of the number of unique organisms

98
00:03:53,270 --> 00:03:51,120
within these brines which you can see is

99
00:03:55,990 --> 00:03:53,280
is increasing rapidly

100
00:03:58,390 --> 00:03:56,000
but on the right we can see that the

101
00:04:01,190 --> 00:03:58,400
total phylogen phylogenetic diversity is

102
00:04:03,750 --> 00:04:01,200
actually quite low and so

103
00:04:05,110 --> 00:04:03,760
this this surge in the number of unique

104
00:04:08,550 --> 00:04:05,120
taxa

105
00:04:09,990 --> 00:04:08,560
is is actually due to very very fine

106
00:04:12,070 --> 00:04:10,000
scale

107
00:04:13,190 --> 00:04:12,080
genetic diversity

108
00:04:16,949 --> 00:04:13,200
and

109
00:04:20,229 --> 00:04:16,959
um

110
00:04:21,749 --> 00:04:20,239
but uh uh this this first stage which is

111
00:04:26,469 --> 00:04:21,759
characterized by the proliferation of

112
00:04:29,510 --> 00:04:26,479
halo archaea um is uh uh followed by

113
00:04:31,990 --> 00:04:29,520

a a stage where the the increasing

114
00:04:34,710 --> 00:04:32,000
selection pressure due to the

115
00:04:37,590 --> 00:04:34,720
increasingly harsh um hyper saline

116
00:04:39,670 --> 00:04:37,600
conditions within the brines

117
00:04:43,189 --> 00:04:39,680
actually begins to drive down the

118
00:04:45,189 --> 00:04:43,199
diversity of halo archaea and

119
00:04:47,030 --> 00:04:45,199
at this stage the brines are dominated

120
00:04:49,110 --> 00:04:47,040
by only a few highly specialized

121
00:04:51,590 --> 00:04:49,120
halophilic taxa

122
00:04:55,110 --> 00:04:51,600
at some point the the the limit of life

123
00:04:56,390 --> 00:04:55,120
in these brines is reached and uh the

124
00:04:58,870 --> 00:04:56,400
the

125
00:05:00,550 --> 00:04:58,880
remaining brines we we assume are

126
00:05:01,830 --> 00:05:00,560
sterile or at the very least devoid of

127
00:05:04,790 --> 00:05:01,840
active life

128
00:05:07,590 --> 00:05:04,800
um and yet we do see that um nonetheless

129
00:05:10,310 --> 00:05:07,600
we we have an increase in bacterial

130
00:05:12,070 --> 00:05:10,320
diversity here um which we initially

131
00:05:12,870 --> 00:05:12,080
thought was somewhat unusual but we did

132
00:05:15,110 --> 00:05:12,880
show

133
00:05:19,909 --> 00:05:15,120
in this

134
00:05:22,790 --> 00:05:19,919
bacterial diversity is actually probably

135
00:05:24,550 --> 00:05:22,800
uh in fall or or exogenous genetic

136
00:05:25,990 --> 00:05:24,560
material on this coming from the

137
00:05:27,749 --> 00:05:26,000
surrounding environment it's falling

138
00:05:29,990 --> 00:05:27,759

into the brines and is getting preserved

139
00:05:31,510 --> 00:05:30,000
there so rather than a living active

140
00:05:33,510 --> 00:05:31,520
community what we're seeing here is a

141
00:05:35,909 --> 00:05:33,520
fossil community

142
00:05:37,830 --> 00:05:35,919
um but i do want to focus for the

143
00:05:39,510 --> 00:05:37,840
remainder of my presentation on this

144
00:05:40,629 --> 00:05:39,520
first stage here

145
00:05:41,909 --> 00:05:40,639
this is

146
00:05:44,230 --> 00:05:41,919
the stage in which we see a

147
00:05:46,390 --> 00:05:44,240
proliferation of halo archaea this surge

148
00:05:48,230 --> 00:05:46,400
of microdiversity and so

149
00:05:50,390 --> 00:05:48,240
in our 2020

150
00:05:52,870 --> 00:05:50,400
field effort and in our subsequent

151
00:05:55,110 --> 00:05:52,880
analyses we're trying to really zoom in

152
00:05:56,790 --> 00:05:55,120
on on that stage one and try to

153
00:05:58,070 --> 00:05:56,800
understand what are the sources of this

154
00:06:00,790 --> 00:05:58,080
diversity

155
00:06:03,510 --> 00:06:00,800
and what are the patterns of evolution

156
00:06:05,830 --> 00:06:03,520
that are are enabling

157
00:06:08,190 --> 00:06:05,840
this adaptation and specialization to

158
00:06:09,270 --> 00:06:08,200
hyper saline brine

159
00:06:10,150 --> 00:06:09,280
[Music]

160
00:06:12,070 --> 00:06:10,160
so

161
00:06:14,469 --> 00:06:12,080
we did take a metagenomics approach in

162
00:06:18,390 --> 00:06:14,479
this 2020 field study

163
00:06:20,390 --> 00:06:18,400

and we we began with using a

164
00:06:22,870 --> 00:06:20,400
bioinformatics pipeline

165
00:06:25,189 --> 00:06:22,880
uh that was developed by

166
00:06:26,790 --> 00:06:25,199
a member of my lab abhishek duta who

167
00:06:29,110 --> 00:06:26,800
i'll be sure to

168
00:06:30,550 --> 00:06:29,120
give an extra shout out to at the end

169
00:06:33,029 --> 00:06:30,560
um

170
00:06:35,029 --> 00:06:33,039
so one of the reasons that i'm

171
00:06:36,790 --> 00:06:35,039
particularly interested in taking this

172
00:06:39,110 --> 00:06:36,800
metagenomics approach

173
00:06:41,029 --> 00:06:39,120
is in order to look for evidence of

174
00:06:44,230 --> 00:06:41,039
horizontal gene transfer

175
00:06:46,550 --> 00:06:44,240
within this this succession from

176
00:06:47,749 --> 00:06:46,560
seawater to sodium chloride saturated

177
00:06:48,790 --> 00:06:47,759
brines

178
00:06:49,909 --> 00:06:48,800
and so

179
00:06:52,150 --> 00:06:49,919
uh

180
00:06:53,749 --> 00:06:52,160
horizontal gene transfer is fairly

181
00:06:56,230 --> 00:06:53,759
well studied in halo archaea it's been

182
00:06:58,790 --> 00:06:56,240
demonstrated that many families of halo

183
00:07:01,990 --> 00:06:58,800
archaea are quite promiscuous and have

184
00:07:02,790 --> 00:07:02,000
higher rates of horizontal gene transfer

185
00:07:04,550 --> 00:07:02,800
and

186
00:07:07,110 --> 00:07:04,560
recombination events

187
00:07:09,909 --> 00:07:07,120
and so i want to understand how

188
00:07:11,749 --> 00:07:09,919

does horizontal gene transfer drive

189
00:07:15,029 --> 00:07:11,759
these patterns and diversity the surge

190
00:07:17,830 --> 00:07:15,039
of diversity if at all

191
00:07:20,469 --> 00:07:17,840
so we begin by looking for some of the

192
00:07:22,550 --> 00:07:20,479
usual suspects um

193
00:07:25,589 --> 00:07:22,560
here i have the uh

194
00:07:27,990 --> 00:07:25,599
the the frequency of gene transfer agent

195
00:07:29,350 --> 00:07:28,000
proteins identified by rast within our

196
00:07:31,830 --> 00:07:29,360
metagenomes

197
00:07:34,309 --> 00:07:31,840
and it didn't take exactly the shape

198
00:07:36,309 --> 00:07:34,319
that i was expecting here we see that

199
00:07:38,790 --> 00:07:36,319
the the peak in

200
00:07:42,710 --> 00:07:38,800
gene transfer agent frequency actually

201
00:07:44,469 --> 00:07:42,720
uh anticipates the the surge of uh of

202
00:07:45,990 --> 00:07:44,479
microdiversity of of archaeal

203
00:07:47,430 --> 00:07:46,000
microdiversity

204
00:07:49,990 --> 00:07:47,440
and so that came as a bit of a surprise

205
00:07:52,550 --> 00:07:50,000
to us but we saw the exact same pattern

206
00:07:55,029 --> 00:07:52,560
emerge uh when we're looking at the the

207
00:07:56,710 --> 00:07:55,039
frequency of mobile element proteins um

208
00:08:00,869 --> 00:07:56,720
identified by rest

209
00:08:03,670 --> 00:08:00,879
and so um uh this this does lead us to

210
00:08:04,790 --> 00:08:03,680
to believe that what we have is actually

211
00:08:06,950 --> 00:08:04,800
a uh

212
00:08:09,589 --> 00:08:06,960
increase in horizontal gene transfer

213
00:08:12,469 --> 00:08:09,599

early on as the conditions begin to to

214
00:08:15,749 --> 00:08:12,479
to become a little bit more extreme but

215
00:08:17,350 --> 00:08:15,759
what we have then is a latent a latent

216
00:08:19,350 --> 00:08:17,360
diverse community

217
00:08:21,589 --> 00:08:19,360
which we weren't actually

218
00:08:22,469 --> 00:08:21,599
sequencing since the the

219
00:08:25,510 --> 00:08:22,479
the

220
00:08:26,629 --> 00:08:25,520
latent diversity might be quite rare

221
00:08:30,309 --> 00:08:26,639
and so

222
00:08:32,070 --> 00:08:30,319
that's that's our hypothesis for now

223
00:08:36,230 --> 00:08:32,080
but what we wanted to

224
00:08:38,949 --> 00:08:36,240
investigate now is okay so the the

225
00:08:41,110 --> 00:08:38,959
meta genomes are showing that there is

226
00:08:43,110 --> 00:08:41,120
quite a bit of change between

227
00:08:45,670 --> 00:08:43,120
seawater environment to the hyper saline

228
00:08:47,829 --> 00:08:45,680
brines but what is the function of this

229
00:08:49,829 --> 00:08:47,839
change what are the functional changes

230
00:08:52,470 --> 00:08:49,839
and so we're looking at the

231
00:08:53,910 --> 00:08:52,480
cog functions of the genes that were

232
00:08:55,990 --> 00:08:53,920
annotated

233
00:08:56,870 --> 00:08:56,000
by by rast

234
00:08:57,990 --> 00:08:56,880
and

235
00:09:01,030 --> 00:08:58,000
eggnog

236
00:09:02,790 --> 00:09:01,040
and a couple of patterns emerge the the

237
00:09:05,269 --> 00:09:02,800
first one is

238
00:09:07,990 --> 00:09:05,279

that quite clearly

239
00:09:09,269 --> 00:09:08,000
the the most common function is unknown

240
00:09:11,269 --> 00:09:09,279
and so

241
00:09:12,949 --> 00:09:11,279
that does tell us that there may be some

242
00:09:14,550 --> 00:09:12,959
more work that needs to be done in order

243
00:09:16,949 --> 00:09:14,560
to understand

244
00:09:18,150 --> 00:09:16,959
what those genes are that that we

245
00:09:22,470 --> 00:09:18,160
haven't

246
00:09:24,710 --> 00:09:22,480
but then

247
00:09:27,110 --> 00:09:24,720
here we did see a pattern emerge that

248
00:09:28,790 --> 00:09:27,120
actually quite closely corresponds to

249
00:09:32,150 --> 00:09:28,800
the uh

250
00:09:35,350 --> 00:09:32,160
pattern of um mobile gene mobile element

251
00:09:36,870 --> 00:09:35,360
proteins and gene transfer agents

252
00:09:39,110 --> 00:09:36,880
and so

253
00:09:41,910 --> 00:09:39,120
this this functional class

254
00:09:44,470 --> 00:09:41,920
is is cell wall membrane

255
00:09:45,910 --> 00:09:44,480
and envelope biogenesis proteins

256
00:09:50,790 --> 00:09:45,920
and uh

257
00:09:53,750 --> 00:09:50,800
so our uh our hypothesis here is that

258
00:09:55,269 --> 00:09:53,760
these these proteins might be the target

259
00:09:58,550 --> 00:09:55,279
of horizontal gene transfer somehow

260
00:10:02,710 --> 00:09:58,560
involved in horizontal gene transfer

261
00:10:05,509 --> 00:10:02,720
but then we do have a third major trend

262
00:10:07,590 --> 00:10:05,519
that's showing up here and that's um the

263
00:10:08,310 --> 00:10:07,600

the the cog

264
00:10:16,310 --> 00:10:08,320
I

265
00:10:18,389 --> 00:10:16,320
and so

266
00:10:20,630 --> 00:10:18,399
that might be a little bit closer to our

267
00:10:22,389 --> 00:10:20,640
original hypothesis we may be seeing

268
00:10:24,550 --> 00:10:22,399
that that that

269
00:10:25,670 --> 00:10:24,560
this is evidence of recombination

270
00:10:27,910 --> 00:10:25,680
driving

271
00:10:30,150 --> 00:10:27,920
some of this surge in microdiversity we

272
00:10:32,310 --> 00:10:30,160
can see that this maps fairly well onto

273
00:10:34,389 --> 00:10:32,320
microdiversity obviously these analyses

274
00:10:38,550 --> 00:10:34,399
are still quite exploratory

275
00:10:45,030 --> 00:10:41,990
i want to thank everybody for listening

276
00:10:47,670 --> 00:10:45,040
and in particular i want to thank

277
00:10:50,470 --> 00:10:47,680
my my team with oceans across space and

278
00:10:52,389 --> 00:10:50,480
time which has enabled this research um

279
00:10:54,630 --> 00:10:52,399
ocean's crossbasin time is funded by the

280
00:10:57,110 --> 00:10:54,640
nasa astrobiology program

281
00:10:59,590 --> 00:10:57,120
i also want to thank all of my uh

282
00:11:01,750 --> 00:10:59,600
collaborators at uh

283
00:11:03,509 --> 00:11:01,760
my my institution but also my

284
00:11:04,630 --> 00:11:03,519
collaborators at georgia tech and at

285
00:11:06,949 --> 00:11:04,640
stanford

286
00:11:09,509 --> 00:11:06,959
that made this research possible

287
00:11:11,670 --> 00:11:09,519
and finally i want to thank the jeff

288
00:11:14,150 --> 00:11:11,680

bowman lab at scripps

289
00:11:16,470 --> 00:11:14,160
that's my lab they've been really really

290
00:11:18,310 --> 00:11:16,480
great listening to me

291
00:11:19,829 --> 00:11:18,320
talk about my research

292
00:11:21,030 --> 00:11:19,839
and supporting me and helping me when

293
00:11:22,710 --> 00:11:21,040
possible

294
00:11:24,550 --> 00:11:22,720
and in particular i do want to thank

295
00:11:26,069 --> 00:11:24,560
abhishek duta who developed the

296
00:11:28,470 --> 00:11:26,079
metagenomics pipeline that i'm working

297
00:11:32,470 --> 00:11:28,480
with and has been an incredible help

298
00:11:34,550 --> 00:11:32,480
with all of my my uh struggles with um

299
00:11:35,670 --> 00:11:34,560
with the the bioinformatics side of

300
00:11:37,509 --> 00:11:35,680
things and

301
00:11:39,590 --> 00:11:37,519
um