



EXPLORING LIFE'S EARLY EVOLUTION  
WITH MOLECULAR PALEOBIOLOGY  
**DR. BETÜL KAÇAR**



1  
00:00:00,240 --> 00:00:03,429  
good morning folks

2  
00:00:07,430 --> 00:00:05,110  
i hope everybody's having a wonderful

3  
00:00:08,870 --> 00:00:07,440  
time at absycon it's good to see a nice

4  
00:00:11,669 --> 00:00:08,880  
crowd here

5  
00:00:13,110 --> 00:00:11,679  
at 8 30 for our plenary session before

6  
00:00:15,829 --> 00:00:13,120  
we get going

7  
00:00:17,990 --> 00:00:15,839  
i would like to remind everyone of

8  
00:00:19,590 --> 00:00:18,000  
what promises to be a really fun event

9  
00:00:23,109 --> 00:00:19,600  
tonight

10  
00:00:25,109 --> 00:00:23,119  
entitled when worlds collide

11  
00:00:26,310 --> 00:00:25,119  
many of us are

12  
00:00:28,390 --> 00:00:26,320  
aware

13  
00:00:31,189 --> 00:00:28,400

from personal experience how science and

14

00:00:34,310 --> 00:00:31,199

the arts grow from the same roots

15

00:00:36,310 --> 00:00:34,320

and reinforce one another and mutually

16

00:00:39,510 --> 00:00:36,320

inform one another

17

00:00:41,350 --> 00:00:39,520

and so this issue will be explored this

18

00:00:45,190 --> 00:00:41,360

evening

19

00:00:47,430 --> 00:00:45,200

and there will be a uh

20

00:00:51,350 --> 00:00:47,440

but a self-described

21

00:00:53,670 --> 00:00:51,360

geek musician uh mr griffin who will be

22

00:00:55,910 --> 00:00:53,680

uh performing and i think he'll all get

23

00:01:00,310 --> 00:00:55,920

a big kick out of that so please join us

24

00:01:06,070 --> 00:01:03,349

so it's my pleasure to start this

25

00:01:08,310 --> 00:01:06,080

morning's uh plenary session

26

00:01:09,590 --> 00:01:08,320

and give an introduction to detect a

27

00:01:10,870 --> 00:01:09,600

speaker

28

00:01:12,870 --> 00:01:10,880

professor

29

00:01:17,270 --> 00:01:12,880

batul

30

00:01:19,350 --> 00:01:17,280

so a long time ago in a galaxy far away

31

00:01:22,950 --> 00:01:19,360

called montana

32

00:01:24,070 --> 00:01:22,960

i received a a a message from michael

33

00:01:27,510 --> 00:01:24,080

new

34

00:01:30,710 --> 00:01:27,520

and specifically michael summoned me

35

00:01:33,109 --> 00:01:30,720

to ab grad khan which was taking place

36

00:01:34,789 --> 00:01:33,119

in the fly fishing capital of ennis

37

00:01:38,069 --> 00:01:34,799

montana

38

00:01:41,350 --> 00:01:38,079

specifically michael summoned me

39

00:01:43,990 --> 00:01:41,360

and so having just received uh an

40

00:01:46,469 --> 00:01:44,000

exobiology grant i immediately hopped

41

00:01:48,550 --> 00:01:46,479

into my old land cruiser and drove the

42

00:01:52,870 --> 00:01:48,560

three hours from missoula

43

00:01:55,109 --> 00:01:52,880

lucky for me the drive was as always

44

00:01:57,350 --> 00:01:55,119

spectacular

45

00:01:58,870 --> 00:01:57,360

but also the interaction with that young

46

00:02:01,670 --> 00:01:58,880

scientist was

47

00:02:03,749 --> 00:02:01,680

delightful and it was auspicious as well

48

00:02:04,950 --> 00:02:03,759

and so over the course of more than a

49

00:02:07,429 --> 00:02:04,960

decade

50

00:02:08,389 --> 00:02:07,439

it has been my pleasure as well as my

51

00:02:11,270 --> 00:02:08,399

privilege

52

00:02:14,309 --> 00:02:11,280

to witness the growth and maturation of

53

00:02:15,990 --> 00:02:14,319

today's plenary speaker

54

00:02:17,510 --> 00:02:16,000

dr khachar

55

00:02:21,350 --> 00:02:17,520

received her

56

00:02:23,270 --> 00:02:21,360

bs from marmara university in istanbul

57

00:02:25,830 --> 00:02:23,280

and then went on to earn her doctorate

58

00:02:28,470 --> 00:02:25,840

in biological chemistry

59

00:02:31,670 --> 00:02:28,480

at emory university

60

00:02:33,350 --> 00:02:31,680

she pursued postdoctoral studies at my

61

00:02:35,589 --> 00:02:33,360

home institution

62

00:02:37,030 --> 00:02:35,599

georgia tech where she began what i

63

00:02:40,150 --> 00:02:37,040

would call

64

00:02:42,470 --> 00:02:40,160

her scientific vision quest

65

00:02:44,710 --> 00:02:42,480

that she'll share with us today but what

66

00:02:47,509 --> 00:02:44,720

do i mean by that well i would direct

67

00:02:51,910 --> 00:02:47,519

your attention to a 2000

68

00:02:53,350 --> 00:02:51,920

uh 2011 paper in the journal artificial

69

00:02:56,150 --> 00:02:53,360

life

70

00:02:59,830 --> 00:02:56,160

entitled towards the recapitulation of

71

00:03:03,670 --> 00:02:59,840

ancient history in the laboratory

72

00:03:06,309 --> 00:03:03,680

this was what was on dr khachar's mind

73

00:03:09,270 --> 00:03:06,319

when we conversed in ennis

74

00:03:11,990 --> 00:03:09,280

now at that time ancestral state

75

00:03:14,710 --> 00:03:12,000

reconstruction was a well-established

76

00:03:17,589 --> 00:03:14,720

practice in phylogenetics

77

00:03:21,030 --> 00:03:17,599

what was daring about

78

00:03:23,030 --> 00:03:21,040

batul was that she aspired to carry on

79

00:03:25,589 --> 00:03:23,040

ancestral state

80

00:03:27,350 --> 00:03:25,599

reconstruction in the laboratory

81

00:03:29,910 --> 00:03:27,360

using

82

00:03:32,550 --> 00:03:29,920

integrated synthetic biology and

83

00:03:35,910 --> 00:03:32,560

experimental evolution

84

00:03:38,390 --> 00:03:35,920

so did she realize these aspirations

85

00:03:41,990 --> 00:03:38,400

i would direct your attention to

86

00:03:45,190 --> 00:03:42,000

last month's article in cell

87

00:03:47,190 --> 00:03:45,200

entitled resurrected rubisco suggests

88

00:03:50,390 --> 00:03:47,200

uniform carbon

89

00:03:51,270 --> 00:03:50,400

isotope signatures over geologic time a

90

00:03:52,710 --> 00:03:51,280

real

91

00:03:54,470 --> 00:03:52,720

masterpiece

92

00:03:56,949 --> 00:03:54,480

uh integrating

93

00:03:59,190 --> 00:03:56,959

all of these different fields uh with

94

00:04:02,309 --> 00:03:59,200

geology

95

00:04:05,190 --> 00:04:02,319

now i would like to emphasize to the

96

00:04:08,149 --> 00:04:05,200

young investigators gathered with us

97

00:04:11,110 --> 00:04:08,159

this morning that the path that led from

98

00:04:13,110 --> 00:04:11,120

istanbul to ennis montana

99

00:04:15,830 --> 00:04:13,120

to madison wisconsin where she's

100

00:04:17,509 --> 00:04:15,840

currently a professor was not an easy

101  
00:04:20,710 --> 00:04:17,519  
one

102  
00:04:22,150 --> 00:04:20,720  
but along that path baitou held fast to

103  
00:04:24,790 --> 00:04:22,160  
her vision

104  
00:04:27,830 --> 00:04:24,800  
and showed true grit

105  
00:04:30,950 --> 00:04:27,840  
taking on challenges and overcoming

106  
00:04:34,550 --> 00:04:30,960  
obstacles and adversity that would have

107  
00:04:36,790 --> 00:04:34,560  
defeated a lesser human being

108  
00:04:39,830 --> 00:04:36,800  
that to my mind more than anything else

109  
00:04:42,710 --> 00:04:39,840  
is what makes professor kachar's

110  
00:04:43,670 --> 00:04:42,720  
journey her story as a scientist and as

111  
00:04:45,749 --> 00:04:43,680  
a woman

112  
00:04:47,749 --> 00:04:45,759  
so compelling

113  
00:04:50,310 --> 00:04:47,759

after georgia tech

114

00:04:53,030 --> 00:04:50,320

baitul continued postdoctoral work with

115

00:04:55,670 --> 00:04:53,040

dan anderson uppsala

116

00:04:56,790 --> 00:04:55,680

university and rich linsky at michigan

117

00:04:59,590 --> 00:04:56,800

state

118

00:05:01,909 --> 00:04:59,600

and then she spent three years at

119

00:05:03,990 --> 00:05:01,919

harvard

120

00:05:07,270 --> 00:05:04,000

working with scott edwards before

121

00:05:09,270 --> 00:05:07,280

accepting her first faculty appointment

122

00:05:11,430 --> 00:05:09,280

just a few years ago at the university

123

00:05:14,629 --> 00:05:11,440

of arizona

124

00:05:17,510 --> 00:05:14,639

this past year her burgeoning laboratory

125

00:05:20,270 --> 00:05:17,520

group moved from tucson to madison where

126

00:05:22,150 --> 00:05:20,280

she is a professor in the department of

127

00:05:24,390 --> 00:05:22,160

bacteriology

128

00:05:27,350 --> 00:05:24,400

dr gachar has been the recipient of

129

00:05:29,830 --> 00:05:27,360

numerous accolades most recently the

130

00:05:31,830 --> 00:05:29,840

stanley I miller award from the

131

00:05:33,749 --> 00:05:31,840

international society for the study of

132

00:05:36,070 --> 00:05:33,759

the origin of life

133

00:05:40,469 --> 00:05:36,080

a cylon fellowship from the simons

134

00:05:43,270 --> 00:05:40,479

foundation a nasa early career award

135

00:05:46,629 --> 00:05:43,280

her laboratory in madison is currently

136

00:05:49,909 --> 00:05:46,639

supported by a nasa templeton foundation

137

00:05:50,870 --> 00:05:49,919

grant the human frontier science program

138

00:05:53,909 --> 00:05:50,880

and

139

00:05:56,150 --> 00:05:53,919

the nasa icar grant that serves as a

140

00:05:57,990 --> 00:05:56,160

cornerstone for the new research

141

00:06:00,469 --> 00:05:58,000

coordination network

142

00:06:02,950 --> 00:06:00,479

life from early cells to

143

00:06:05,430 --> 00:06:02,960

multicellularity

144

00:06:08,309 --> 00:06:05,440

please join me in giving me a warm abs

145

00:06:09,420 --> 00:06:08,319

icon welcome to one of our own

146

00:06:18,830 --> 00:06:09,430

batul

147

00:06:23,670 --> 00:06:21,430

kachor wow

148

00:06:26,550 --> 00:06:23,680

thank you so much um frank

149

00:06:27,990 --> 00:06:26,560

i um it really feels like homecoming

150

00:06:30,390 --> 00:06:28,000

being here

151  
00:06:31,590 --> 00:06:30,400  
i'm i always define my job as one of the

152  
00:06:34,469 --> 00:06:31,600  
most

153  
00:06:37,590 --> 00:06:34,479  
special jobs one can ever have to be a

154  
00:06:39,189 --> 00:06:37,600  
biologist on the only planet that you

155  
00:06:42,150 --> 00:06:39,199  
can be one

156  
00:06:43,749 --> 00:06:42,160  
right you can be a physicist or chem or

157  
00:06:45,749 --> 00:06:43,759  
you can study physics or chemistry and

158  
00:06:47,990 --> 00:06:45,759  
geology anywhere else in the universe

159  
00:06:51,510 --> 00:06:48,000  
but this is the only place where we can

160  
00:06:53,589 --> 00:06:51,520  
be biologists that's to me amazing and

161  
00:06:55,510 --> 00:06:53,599  
of course this is the crowd that that's

162  
00:06:57,990 --> 00:06:55,520  
trying to change that so we can be

163  
00:06:59,189 --> 00:06:58,000

biologists on different planets

164

00:07:02,710 --> 00:06:59,199

and we know

165

00:07:05,830 --> 00:07:02,720

that the challenge is big

166

00:07:07,510 --> 00:07:05,840

what if i told you that these two images

167

00:07:09,189 --> 00:07:07,520

well it kind of feels weird to not have

168

00:07:11,189 --> 00:07:09,199

slides here there

169

00:07:16,870 --> 00:07:11,199

all right one of these images is

170

00:07:19,510 --> 00:07:16,880

lifeless whereas the other one is living

171

00:07:20,710 --> 00:07:19,520

at the very superficial level life can

172

00:07:22,309 --> 00:07:20,720

trick us

173

00:07:23,990 --> 00:07:22,319

for his existence

174

00:07:25,270 --> 00:07:24,000

the image on the

175

00:07:28,150 --> 00:07:25,280

um

176  
00:07:29,430 --> 00:07:28,160  
left in fact is lifeless

177  
00:07:32,150 --> 00:07:29,440  
the fern

178  
00:07:34,230 --> 00:07:32,160  
that may be tricking you to be a fern is

179  
00:07:35,589 --> 00:07:34,240  
a silver dendrite

180  
00:07:37,990 --> 00:07:35,599  
this is a

181  
00:07:39,430 --> 00:07:38,000  
outcome of a malfunctioning engineering

182  
00:07:42,070 --> 00:07:39,440  
experiment

183  
00:07:43,430 --> 00:07:42,080  
and on the right you're looking at a

184  
00:07:44,710 --> 00:07:43,440  
forest

185  
00:07:46,550 --> 00:07:44,720  
a fern

186  
00:07:48,629 --> 00:07:46,560  
a lot of life

187  
00:07:50,230 --> 00:07:48,639  
so while at the superficial level life

188  
00:07:52,550 --> 00:07:50,240

can trick us and i think this very

189

00:07:54,869 --> 00:07:52,560

aspect makes our jobs as a cerebiologist

190

00:07:57,589 --> 00:07:54,879

very difficult things start to fall

191

00:07:59,909 --> 00:07:57,599

apart the moment we go up a scale or

192

00:08:02,230 --> 00:07:59,919

down a scale and ask deeper questions

193

00:08:04,629 --> 00:08:02,240

about the image that we are

194

00:08:06,550 --> 00:08:04,639

we are examining

195

00:08:09,990 --> 00:08:06,560

what i mean by that is that if you

196

00:08:11,670 --> 00:08:10,000

travel up a level or down

197

00:08:14,230 --> 00:08:11,680

you have to

198

00:08:16,550 --> 00:08:14,240

experience or what you are looking needs

199

00:08:18,790 --> 00:08:16,560

to express a level of complexity or in

200

00:08:21,430 --> 00:08:18,800

increasing complexity or a decrease in

201  
00:08:24,790 --> 00:08:21,440  
complexity which is what we generally

202  
00:08:26,390 --> 00:08:24,800  
think a living organism or living state

203  
00:08:28,390 --> 00:08:26,400  
attributes

204  
00:08:31,670 --> 00:08:28,400  
so let's just start with

205  
00:08:33,509 --> 00:08:31,680  
an organism we have uh if you go up a

206  
00:08:35,670 --> 00:08:33,519  
level we have populations that form

207  
00:08:37,430 --> 00:08:35,680  
communities that form ecosystems they

208  
00:08:39,190 --> 00:08:37,440  
impact the ecosystem and ultimately

209  
00:08:40,870 --> 00:08:39,200  
shape the biosphere

210  
00:08:43,670 --> 00:08:40,880  
if you go down a level we have cell

211  
00:08:46,310 --> 00:08:43,680  
groups and cell at the very bottom we

212  
00:08:47,910 --> 00:08:46,320  
have molecules and you can also uh strip

213  
00:08:49,590 --> 00:08:47,920

them apart

214

00:08:51,350 --> 00:08:49,600

we know that there are self similar

215

00:08:53,590 --> 00:08:51,360

patterns of interactions between each

216

00:08:55,110 --> 00:08:53,600

layers and i am i'm showing you b

217

00:08:57,509 --> 00:08:55,120

because i think they're fascinating is a

218

00:08:59,829 --> 00:08:57,519

very top-level example of what life

219

00:09:01,910 --> 00:08:59,839

organ or how it organizes itself there

220

00:09:03,590 --> 00:09:01,920

are rules between levels and there is

221

00:09:05,430 --> 00:09:03,600

different level of complexity in each

222

00:09:07,110 --> 00:09:05,440

layer

223

00:09:08,470 --> 00:09:07,120

so let's just drill down this image a

224

00:09:11,670 --> 00:09:08,480

little bit

225

00:09:13,910 --> 00:09:11,680

at the top we have biogeochemical cycles

226

00:09:16,150 --> 00:09:13,920

catalyzed or driven or moderated by

227

00:09:18,630 --> 00:09:16,160

populations and again if you go down a

228

00:09:20,470 --> 00:09:18,640

level if you have organisms and cells

229

00:09:22,710 --> 00:09:20,480

and subcellular micromolecules that

230

00:09:24,310 --> 00:09:22,720

compose cells and finally at the very

231

00:09:27,670 --> 00:09:24,320

bottom we have energy input in

232

00:09:30,710 --> 00:09:27,680

substrates and i know that this group is

233

00:09:32,310 --> 00:09:30,720

studying this aspect of life it made at

234

00:09:35,350 --> 00:09:32,320

various levels

235

00:09:37,430 --> 00:09:35,360

so let's consider a modern ecosystem and

236

00:09:39,030 --> 00:09:37,440

i'm showing an image of the dinosaurs

237

00:09:40,630 --> 00:09:39,040

and i know this is tricky because when i

238

00:09:43,910 --> 00:09:40,640

talk about early life at least when i

239

00:09:45,829 --> 00:09:43,920

teach at university most students think

240

00:09:47,430 --> 00:09:45,839

that i'm talking about dinosaurs because

241

00:09:49,269 --> 00:09:47,440

they're ancient right

242

00:09:51,990 --> 00:09:49,279

but of course this group knows that

243

00:09:53,670 --> 00:09:52,000

dinosaurs existed in a relatively modern

244

00:09:55,870 --> 00:09:53,680

ecosystem

245

00:09:58,070 --> 00:09:55,880

and what we care about when when an

246

00:09:59,190 --> 00:09:58,080

astrobiologist looks at an image like

247

00:10:00,710 --> 00:09:59,200

this

248

00:10:02,710 --> 00:10:00,720

what will really

249

00:10:04,790 --> 00:10:02,720

get their interest is that it's not

250

00:10:06,790 --> 00:10:04,800

going to be the individuals per se but

251  
00:10:09,269 --> 00:10:06,800  
the effect of the individual on the

252  
00:10:10,790 --> 00:10:09,279  
planet scale reservations

253  
00:10:13,269 --> 00:10:10,800  
export service

254  
00:10:15,750 --> 00:10:13,279  
atmosphere maybe temperature we are

255  
00:10:17,590 --> 00:10:15,760  
interested in understanding the impact

256  
00:10:20,550 --> 00:10:17,600  
of whatever the living state is in that

257  
00:10:22,230 --> 00:10:20,560  
very body with it is environment

258  
00:10:25,509 --> 00:10:22,240  
and of course the challenge might then

259  
00:10:26,949 --> 00:10:25,519  
be to detect these impacts

260  
00:10:28,790 --> 00:10:26,959  
and i'm not even talking about the

261  
00:10:31,990 --> 00:10:28,800  
detectability problem just yet but we

262  
00:10:34,150 --> 00:10:32,000  
know that direct observation is not easy

263  
00:10:36,069 --> 00:10:34,160

of course if we find dinosaurs on a

264

00:10:39,670 --> 00:10:36,079

different planet in our solar system at

265

00:10:45,269 --> 00:10:42,949

so biologists are aware of these issues

266

00:10:47,829 --> 00:10:45,279

and and have amazing tools to study

267

00:10:50,230 --> 00:10:47,839

these interactions but what biologists

268

00:10:51,990 --> 00:10:50,240

are lacking especially molecular studies

269

00:10:55,190 --> 00:10:52,000

i think is how

270

00:10:57,829 --> 00:10:55,200

are the the concept of how environments

271

00:10:59,190 --> 00:10:57,839

and evolutionary innovations impacted

272

00:11:01,430 --> 00:10:59,200

our planets

273

00:11:03,910 --> 00:11:01,440

we lack an understanding of the

274

00:11:06,389 --> 00:11:03,920

underlying variation and the impact of

275

00:11:09,030 --> 00:11:06,399

this variation at the planetary scale

276

00:11:10,470 --> 00:11:09,040

over long time scales

277

00:11:12,949 --> 00:11:10,480

we also don't understand the

278

00:11:15,750 --> 00:11:12,959

circumstances and timing of events that

279

00:11:16,790 --> 00:11:15,760

made the early earth a much different

280

00:11:19,190 --> 00:11:16,800

place

281

00:11:21,110 --> 00:11:19,200

and we know these things through various

282

00:11:23,590 --> 00:11:21,120

studies that come from paleontology

283

00:11:26,389 --> 00:11:23,600

paleobiology geology geobiology

284

00:11:29,990 --> 00:11:26,399

geochemistry so on so forth

285

00:11:32,150 --> 00:11:30,000

so let's review what i'm talking about

286

00:11:34,069 --> 00:11:32,160

uh of course you know we're interested

287

00:11:36,470 --> 00:11:34,079

in the birth of life how this planet

288

00:11:38,710 --> 00:11:36,480

gave birth to this magnificent thing

289

00:11:40,790 --> 00:11:38,720

that we get to study right now and but

290

00:11:42,230 --> 00:11:40,800

in some cases if you think about our our

291

00:11:44,230 --> 00:11:42,240

kind earth or the hideon earth it's

292

00:11:47,670 --> 00:11:44,240

really the best of times but also the

293

00:11:49,269 --> 00:11:47,680

worst of times to study

294

00:11:51,269 --> 00:11:49,279

it's the best of times because life

295

00:11:53,269 --> 00:11:51,279

emerged it's the worst of times because

296

00:11:55,350 --> 00:11:53,279

we have very little to work with when it

297

00:11:57,590 --> 00:11:55,360

comes to studying the past

298

00:11:59,990 --> 00:11:57,600

there is very limited record it comes

299

00:12:01,829 --> 00:12:00,000

from very limited data

300

00:12:03,829 --> 00:12:01,839

that represents a very limited

301  
00:12:06,710 --> 00:12:03,839  
subfraction of whatever the living

302  
00:12:08,310 --> 00:12:06,720  
organism or the population was reciting

303  
00:12:09,350 --> 00:12:08,320  
in that very locality that we are

304  
00:12:11,829 --> 00:12:09,360  
studying

305  
00:12:13,030 --> 00:12:11,839  
and the bigger the mystery the fewer the

306  
00:12:15,750 --> 00:12:13,040  
data

307  
00:12:18,310 --> 00:12:15,760  
so let's review what these major

308  
00:12:20,230 --> 00:12:18,320  
innovations are over time

309  
00:12:24,310 --> 00:12:20,240  
we have cellular life

310  
00:12:26,150 --> 00:12:24,320  
emerging around 3.8 say 4 billion i put

311  
00:12:28,069 --> 00:12:26,160  
these arrows as absolute numbers but of

312  
00:12:29,509 --> 00:12:28,079  
course they do not represent an absolute

313  
00:12:32,470 --> 00:12:29,519

day

314

00:12:36,150 --> 00:12:32,480

emergence of cyanobacteria

315

00:12:38,150 --> 00:12:36,160

eukaryotes followed by green algae and

316

00:12:39,670 --> 00:12:38,160

of course these are underlined by in

317

00:12:42,230 --> 00:12:39,680

this case and the second in the

318

00:12:43,910 --> 00:12:42,240

symbiosis event that gave rise to these

319

00:12:46,389 --> 00:12:43,920

what i'm sharing here and then we have

320

00:12:48,150 --> 00:12:46,399

animals and land plants and surface

321

00:12:50,870 --> 00:12:48,160

ecosystems

322

00:12:53,590 --> 00:12:50,880

so when we look at the history of life

323

00:12:55,990 --> 00:12:53,600

on earth these are the top innovations i

324

00:12:57,750 --> 00:12:56,000

think that really impacted what we see

325

00:13:00,310 --> 00:12:57,760

around us today

326

00:13:02,230 --> 00:13:00,320

this is magnificent to me and if you

327

00:13:05,110 --> 00:13:02,240

think about it all of these events are

328

00:13:07,030 --> 00:13:05,120

interconnected it all comes back to the

329

00:13:08,790 --> 00:13:07,040

cellular life and some innovation that

330

00:13:09,990 --> 00:13:08,800

happened at that point that got

331

00:13:12,069 --> 00:13:10,000

transferred

332

00:13:12,790 --> 00:13:12,079

i'm using the word transfer with a lot

333

00:13:14,389 --> 00:13:12,800

of

334

00:13:16,629 --> 00:13:14,399

conservation here

335

00:13:18,870 --> 00:13:16,639

um maybe in some cases adapted or it was

336

00:13:21,509 --> 00:13:18,880

a fluke accident we don't know but they

337

00:13:23,829 --> 00:13:21,519

all rely on each other so if i am here

338

00:13:26,069 --> 00:13:23,839

and talking to you it is because some

339

00:13:27,030 --> 00:13:26,079

innovation at the molecular level

340

00:13:28,629 --> 00:13:27,040

happened

341

00:13:30,550 --> 00:13:28,639

billions of years ago the fact that we

342

00:13:33,670 --> 00:13:30,560

are having this very conversation relies

343

00:13:34,470 --> 00:13:33,680

on these interactions of the past

344

00:13:35,829 --> 00:13:34,480

and

345

00:13:38,949 --> 00:13:35,839

some of these

346

00:13:41,350 --> 00:13:38,959

in fact majority of them are seeing what

347

00:13:42,870 --> 00:13:41,360

we like to think of as singular events

348

00:13:44,790 --> 00:13:42,880

they happened once

349

00:13:47,189 --> 00:13:44,800

and nothing like them happened ever

350

00:13:49,269 --> 00:13:47,199

again and we cannot compare them to

351

00:13:51,030 --> 00:13:49,279

independent data sets we don't have any

352

00:13:53,189 --> 00:13:51,040

other planet with life so this is what

353

00:13:55,189 --> 00:13:53,199

we have even in this place we can we

354

00:13:58,550 --> 00:13:55,199

don't have an independent record yet at

355

00:14:01,430 --> 00:13:58,560

least to compare these events to

356

00:14:03,269 --> 00:14:01,440

and they happened billions of years ago

357

00:14:05,590 --> 00:14:03,279

you cannot have an ecosystem or

358

00:14:07,670 --> 00:14:05,600

biosphere without cells

359

00:14:09,750 --> 00:14:07,680

i just recently read a paper that

360

00:14:10,949 --> 00:14:09,760

animals and and the rise of animals may

361

00:14:13,110 --> 00:14:10,959

be due to

362

00:14:14,870 --> 00:14:13,120

evolution of nitrogen fixation and the

363

00:14:16,949 --> 00:14:14,880

impact of that with the rise of oxygen

364

00:14:19,509 --> 00:14:16,959

on this planet so these are all

365

00:14:22,470 --> 00:14:19,519

interconnected and it's quite remarkable

366

00:14:23,990 --> 00:14:22,480

that these events happened at singular

367

00:14:25,189 --> 00:14:24,000

themselves

368

00:14:26,710 --> 00:14:25,199

in early life

369

00:14:29,110 --> 00:14:26,720

it's complicated

370

00:14:31,430 --> 00:14:29,120

if you were a bacteria looking for

371

00:14:33,509 --> 00:14:31,440

a date with a really really old organism

372

00:14:35,030 --> 00:14:33,519

and go to your tinder profile or just

373

00:14:36,949 --> 00:14:35,040

check it out this is what you would see

374

00:14:39,430 --> 00:14:36,959

it's some sort of

375

00:14:41,430 --> 00:14:39,440

i mean as a biologist who's studying you

376

00:14:43,269 --> 00:14:41,440

know biochemical and chemical and

377

00:14:45,430 --> 00:14:43,279

molecular systems when i first saw this

378

00:14:47,350 --> 00:14:45,440

i thought this is it

379

00:14:49,670 --> 00:14:47,360

so when we look at past this this is

380

00:14:52,150 --> 00:14:49,680

what we see so the favorite food is

381

00:14:54,949 --> 00:14:52,160

complicated relationship status is very

382

00:14:57,509 --> 00:14:54,959

complicated age it's very complicated

383

00:14:59,030 --> 00:14:57,519

family lame let's not even go there

384

00:15:01,269 --> 00:14:59,040

so this is what we have is an early

385

00:15:04,069 --> 00:15:01,279

pre-cambrian microbe

386

00:15:06,470 --> 00:15:04,079

and the foundations of these jumps like

387

00:15:09,430 --> 00:15:06,480

call them jumps that i just showed you

388

00:15:10,949 --> 00:15:09,440

that define our planet are completely

389

00:15:12,949 --> 00:15:10,959

unknown

390

00:15:15,350 --> 00:15:12,959

and i will add another layer to this

391

00:15:18,949 --> 00:15:15,360

challenge it is not even clear whether

392

00:15:21,030 --> 00:15:18,959

these events themselves were inevitable

393

00:15:23,269 --> 00:15:21,040

right so that's another that's another

394

00:15:24,949 --> 00:15:23,279

challenge we got there so let's revisit

395

00:15:27,350 --> 00:15:24,959

the image i just showed you this time

396

00:15:29,990 --> 00:15:27,360

with some pretty pictures but what we

397

00:15:32,389 --> 00:15:30,000

know comes from maybe some chemical

398

00:15:34,310 --> 00:15:32,399

isotopic probing and and those if we are

399

00:15:36,790 --> 00:15:34,320

lucky we have some fossil data again as

400

00:15:39,509 --> 00:15:36,800

we go back in time the the we have

401  
00:15:41,910 --> 00:15:39,519  
scattered information and and the record

402  
00:15:43,910 --> 00:15:41,920  
gets spottier and spottier

403  
00:15:45,910 --> 00:15:43,920  
but i want you to look at the gaps

404  
00:15:48,230 --> 00:15:45,920  
between these major biological

405  
00:15:51,910 --> 00:15:48,240  
innovations in the history of life and

406  
00:15:53,430 --> 00:15:51,920  
by gaps i mean the timing right hundreds

407  
00:15:57,189 --> 00:15:53,440  
millions

408  
00:16:01,910 --> 00:15:59,350  
and i will tell you now why as a as a

409  
00:16:04,069 --> 00:16:01,920  
biologist i find that amazing

410  
00:16:06,790 --> 00:16:04,079  
because we know that at the molecular

411  
00:16:08,310 --> 00:16:06,800  
scale variations variation happens very

412  
00:16:10,949 --> 00:16:08,320  
rapidly

413  
00:16:13,110 --> 00:16:10,959

okay we do not have

414

00:16:15,269 --> 00:16:13,120  
an obvious way deterministic or

415

00:16:17,829 --> 00:16:15,279  
mechanistic

416

00:16:19,430 --> 00:16:17,839  
to link variability on the scale of

417

00:16:21,829 --> 00:16:19,440  
minutes

418

00:16:24,069 --> 00:16:21,839  
with micro evolutionary innovation that

419

00:16:26,870 --> 00:16:24,079  
expresses themselves on the scale of

420

00:16:27,990 --> 00:16:26,880  
billions of years

421

00:16:29,990 --> 00:16:28,000  
think about

422

00:16:31,990 --> 00:16:30,000  
translation machinery it's it's the

423

00:16:35,189 --> 00:16:32,000  
computer of life right it processes

424

00:16:37,110 --> 00:16:35,199  
everything that in every organism that

425

00:16:39,590 --> 00:16:37,120  
ever existed on this planet has this

426

00:16:41,670 --> 00:16:39,600

computer in itself

427

00:16:44,069 --> 00:16:41,680

it is able to generate a peptide chain

428

00:16:46,710 --> 00:16:44,079

in some cases less than a second

429

00:16:48,470 --> 00:16:46,720

that is very fast we know when we study

430

00:16:50,470 --> 00:16:48,480

molecule and molecular life that the

431

00:16:53,030 --> 00:16:50,480

variation exists in that very rapid

432

00:16:54,870 --> 00:16:53,040

speed so how is it possible that a

433

00:16:56,470 --> 00:16:54,880

variation that happens at that really

434

00:16:58,470 --> 00:16:56,480

fast pace

435

00:17:03,350 --> 00:16:58,480

leads to these major innovations that

436

00:17:06,150 --> 00:17:04,630

i think this is

437

00:17:07,990 --> 00:17:06,160

a miracle

438

00:17:09,990 --> 00:17:08,000

but also a curse at the same time that's

439

00:17:12,230 --> 00:17:10,000

why i think studying early life is very

440

00:17:14,150 --> 00:17:12,240

fascinating but also very challenging

441

00:17:16,789 --> 00:17:14,160

because we don't have ways to link this

442

00:17:18,390 --> 00:17:16,799

thing these two events and and

443

00:17:19,909 --> 00:17:18,400

i think the fact that we are struggling

444

00:17:21,510 --> 00:17:19,919

to even comprehend these things or at

445

00:17:24,069 --> 00:17:21,520

least i do

446

00:17:26,150 --> 00:17:24,079

also perhaps indicates that we do not

447

00:17:28,150 --> 00:17:26,160

teach biology and biological concepts

448

00:17:30,870 --> 00:17:28,160

the way we should we don't learn them in

449

00:17:33,990 --> 00:17:30,880

the way we should we need to integrate

450

00:17:36,630 --> 00:17:34,000

life's history when we teach life

451  
00:17:37,590 --> 00:17:36,640  
and i don't think we do that

452  
00:17:39,110 --> 00:17:37,600  
at least

453  
00:17:40,710 --> 00:17:39,120  
comprehensively at the moment in in

454  
00:17:41,909 --> 00:17:40,720  
biology

455  
00:17:43,990 --> 00:17:41,919  
so

456  
00:17:46,549 --> 00:17:44,000  
as i mentioned just adding another layer

457  
00:17:48,950 --> 00:17:46,559  
to the challenge we have this issue of

458  
00:17:51,430 --> 00:17:48,960  
inevitability and i think

459  
00:17:53,270 --> 00:17:51,440  
in my mind contingency is sort of

460  
00:17:55,270 --> 00:17:53,280  
it's kind of a reasonable

461  
00:17:57,029 --> 00:17:55,280  
assumption to make here

462  
00:18:00,310 --> 00:17:57,039  
just because it happened that way

463  
00:18:02,310 --> 00:18:00,320

doesn't mean it had to happen that way

464

00:18:04,230 --> 00:18:02,320

right so the case for this in especially

465

00:18:06,630 --> 00:18:04,240

in deep time evolution of these earliest

466

00:18:08,549 --> 00:18:06,640

organisms and biomolecules is a pretty

467

00:18:10,390 --> 00:18:08,559

reasonable one

468

00:18:12,830 --> 00:18:10,400

and we are aware of this because of the

469

00:18:16,549 --> 00:18:12,840

studies that come from chemistry and

470

00:18:18,789 --> 00:18:16,559

physics that they provide us this all of

471

00:18:21,190 --> 00:18:18,799

the chemical space that is possible in

472

00:18:23,510 --> 00:18:21,200

the universe and then when you think of

473

00:18:24,789 --> 00:18:23,520

it about the innovations on this planet

474

00:18:26,789 --> 00:18:24,799

they each

475

00:18:29,669 --> 00:18:26,799

are our sort of

476

00:18:32,470 --> 00:18:29,679

universe of possibilities themselves

477

00:18:35,190 --> 00:18:32,480

so we know this vast chemical space and

478

00:18:38,870 --> 00:18:35,200

we know how narrow life reduced it down

479

00:18:40,549 --> 00:18:38,880

to to some selective amino acids we know

480

00:18:42,789 --> 00:18:40,559

there's a handful of amino acids like

481

00:18:44,870 --> 00:18:42,799

life selected or certain backbones that

482

00:18:46,470 --> 00:18:44,880

life likes to use if you think about it

483

00:18:48,470 --> 00:18:46,480

considering all the possibilities it's a

484

00:18:52,549 --> 00:18:48,480

pretty selected

485

00:18:58,310 --> 00:18:54,549

so while physics and chemistry can give

486

00:19:01,270 --> 00:18:58,320

us this oven wonder of the general what

487

00:19:06,549 --> 00:19:01,280

we need is to understand the depths of

488

00:19:09,990 --> 00:19:07,909

in this regard

489

00:19:12,390 --> 00:19:10,000

there is no substitute for fundamental

490

00:19:14,710 --> 00:19:12,400

biology in the assessment of planetary

491

00:19:16,549 --> 00:19:14,720

scale life in the universe

492

00:19:17,510 --> 00:19:16,559

we need to understand these rules of

493

00:19:21,029 --> 00:19:17,520

life

494

00:19:22,789 --> 00:19:21,039

that enabled the selection the subset of

495

00:19:25,990 --> 00:19:22,799

the chemicals that it studies and

496

00:19:27,830 --> 00:19:26,000

utilizes frequently

497

00:19:29,190 --> 00:19:27,840

we need to understand these events on

498

00:19:30,870 --> 00:19:29,200

their own terms

499

00:19:32,789 --> 00:19:30,880

and with what little data we have

500

00:19:34,710 --> 00:19:32,799

available

501  
00:19:36,710 --> 00:19:34,720  
and this is also coming

502  
00:19:38,710 --> 00:19:36,720  
without the benefit and of course maybe

503  
00:19:41,909 --> 00:19:38,720  
in some cases the limitation of

504  
00:19:43,830 --> 00:19:41,919  
comparing them to any other evolutionary

505  
00:19:46,390 --> 00:19:43,840  
history

506  
00:19:49,990 --> 00:19:46,400  
this was really where

507  
00:19:52,470 --> 00:19:50,000  
my focus uh has has been drawn into

508  
00:19:54,470 --> 00:19:52,480  
and and and this is really the crowd

509  
00:19:56,310 --> 00:19:54,480  
that made me think about these things

510  
00:19:58,710 --> 00:19:56,320  
and and showed me these gaps in our

511  
00:20:01,110 --> 00:19:58,720  
understanding of life and and i've been

512  
00:20:03,830 --> 00:20:01,120  
thinking how do we do this how do how do

513  
00:20:07,270 --> 00:20:03,840

we draw the benefit of all the tools and

514

00:20:10,470 --> 00:20:07,280

systems that we now develop

515

00:20:12,149 --> 00:20:10,480

to questions such as this

516

00:20:14,549 --> 00:20:12,159

and it was inevitable that we need to

517

00:20:17,029 --> 00:20:14,559

develop systems however limited that

518

00:20:19,350 --> 00:20:17,039

might be and we i'm i would love to talk

519

00:20:23,029 --> 00:20:19,360

to you about that more and we will be

520

00:20:25,830 --> 00:20:23,039

to travel backwards from present

521

00:20:28,549 --> 00:20:25,840

that let's develop systems where we can

522

00:20:29,830 --> 00:20:28,559

at least make an attempt to understand

523

00:20:31,430 --> 00:20:29,840

the

524

00:20:33,590 --> 00:20:31,440

house

525

00:20:35,669 --> 00:20:33,600

how can a molecular system give birth to

526  
00:20:39,510 --> 00:20:35,679  
such major innovation

527  
00:20:42,230 --> 00:20:40,310  
i

528  
00:20:45,990 --> 00:20:42,240  
narrowed my focus to

529  
00:20:49,350 --> 00:20:46,000  
not only organisms and cells but the

530  
00:20:51,909 --> 00:20:49,360  
proteins that enable majority of the bio

531  
00:20:53,990 --> 00:20:51,919  
geochemical cycles and metabolisms that

532  
00:20:55,669 --> 00:20:54,000  
whose outputs we get to study in the

533  
00:20:57,669 --> 00:20:55,679  
rock record

534  
00:21:00,470 --> 00:20:57,679  
and i spent quite a bit of time trying

535  
00:21:03,029 --> 00:21:00,480  
to link what sort of protein system can

536  
00:21:06,390 --> 00:21:03,039  
completely realize not only the history

537  
00:21:08,630 --> 00:21:06,400  
of life but but enable us to study the

538  
00:21:10,630 --> 00:21:08,640

impact at the planetary scale

539

00:21:13,909 --> 00:21:10,640

it seems like a big stretch

540

00:21:16,710 --> 00:21:13,919

even though i showed you how a cell at

541

00:21:18,470 --> 00:21:16,720

the at the if we go up a layer is

542

00:21:21,909 --> 00:21:18,480

connected to what we study at the

543

00:21:27,350 --> 00:21:24,470

overall the approach in my mind at least

544

00:21:30,789 --> 00:21:27,360

was quite straightforward

545

00:21:32,630 --> 00:21:30,799

i nasa loves to use the term

546

00:21:34,950 --> 00:21:32,640

not a fruit salad but we're making a

547

00:21:36,549 --> 00:21:34,960

smoothie here and i always thought well

548

00:21:38,470 --> 00:21:36,559

that's great but you have to be very

549

00:21:39,990 --> 00:21:38,480

careful with what you put in the blender

550

00:21:41,990 --> 00:21:40,000

when you make a smoothie because if you

551  
00:21:43,510 --> 00:21:42,000  
use it wrong ingredient you can ruin the

552  
00:21:44,789 --> 00:21:43,520  
entire smoothie and there's no going

553  
00:21:46,549 --> 00:21:44,799  
back

554  
00:21:49,350 --> 00:21:46,559  
so uh

555  
00:21:51,110 --> 00:21:49,360  
with this dramatic pursue i i thought

556  
00:21:53,110 --> 00:21:51,120  
about all right so we have the genes and

557  
00:21:55,510 --> 00:21:53,120  
proteins and and yes we have

558  
00:21:58,390 --> 00:21:55,520  
phylogenetic approaches that allow us to

559  
00:22:01,110 --> 00:21:58,400  
study the revolution today we

560  
00:22:03,029 --> 00:22:01,120  
know so much by studying viral evolution

561  
00:22:04,789 --> 00:22:03,039  
predicting where mutations might be to

562  
00:22:06,710 --> 00:22:04,799  
some degree yes they have their

563  
00:22:09,350 --> 00:22:06,720

restrictions but let's try to stretch

564

00:22:12,230 --> 00:22:09,360

this at the deep time scale

565

00:22:14,470 --> 00:22:12,240

and can we engineer systems with these

566

00:22:17,350 --> 00:22:14,480

phylogenetically reconstructed genes to

567

00:22:20,549 --> 00:22:17,360

reprogram the behavior of the cell which

568

00:22:25,669 --> 00:22:20,559

we can then study geochemically in the

569

00:22:29,270 --> 00:22:26,870

with that

570

00:22:31,510 --> 00:22:29,280

started this pursuit of evolutionary

571

00:22:33,510 --> 00:22:31,520

construction of ancient enzymes yes the

572

00:22:35,350 --> 00:22:33,520

ancestral state reconstruction was

573

00:22:37,270 --> 00:22:35,360

frequently applied to understand the

574

00:22:40,470 --> 00:22:37,280

evolutionary mechanisms and particularly

575

00:22:42,230 --> 00:22:40,480

how mutations drive adaptation but

576

00:22:44,390 --> 00:22:42,240

wasn't being stretched i think deep

577

00:22:45,590 --> 00:22:44,400

enough to understand the planetary scale

578

00:22:47,350 --> 00:22:45,600

phenomena

579

00:22:49,590 --> 00:22:47,360

and of course that was

580

00:22:52,149 --> 00:22:49,600

those applications i was exposed to

581

00:22:55,350 --> 00:22:52,159

those possibilities here

582

00:22:57,669 --> 00:22:55,360

and here is a paper by margaret day hall

583

00:22:59,669 --> 00:22:57,679

published in 1981

584

00:23:02,549 --> 00:22:59,679

this might be one of the papers that she

585

00:23:04,070 --> 00:23:02,559

has that have the fewest citations maybe

586

00:23:06,549 --> 00:23:04,080

20 30

587

00:23:08,630 --> 00:23:06,559

we might be half of those citations

588

00:23:10,549 --> 00:23:08,640

this is a remarkable paper evolution of

589

00:23:12,630 --> 00:23:10,559

major metabolic innovations in the

590

00:23:14,789 --> 00:23:12,640

pre-cambrian margaret dayhoff was a

591

00:23:16,789 --> 00:23:14,799

biophysicist

592

00:23:18,950 --> 00:23:16,799

she revolutionized how we study

593

00:23:22,630 --> 00:23:18,960

compute how we utilize computational

594

00:23:26,870 --> 00:23:22,640

power to study protein structures

595

00:23:28,470 --> 00:23:26,880

in this paper they examine pathways

596

00:23:30,710 --> 00:23:28,480

where there is information on the

597

00:23:32,870 --> 00:23:30,720

chemical structure of enzyme say so they

598

00:23:35,190 --> 00:23:32,880

focus on extracting some information

599

00:23:36,870 --> 00:23:35,200

from the structure of the proteins

600

00:23:38,789 --> 00:23:36,880

and they assume

601  
00:23:40,789 --> 00:23:38,799  
that the gene products

602  
00:23:42,390 --> 00:23:40,799  
whose structures have been conserved of

603  
00:23:45,110 --> 00:23:42,400  
these proteins

604  
00:23:47,269 --> 00:23:45,120  
have remains relatively unchanged in

605  
00:23:48,310 --> 00:23:47,279  
their basic functions

606  
00:23:49,510 --> 00:23:48,320  
over

607  
00:23:51,430 --> 00:23:49,520  
um

608  
00:23:53,029 --> 00:23:51,440  
throughout pre-cambrian

609  
00:23:54,549 --> 00:23:53,039  
so there is this assumption in their

610  
00:23:58,470 --> 00:23:54,559  
paper it's a very brief paper i highly

611  
00:24:01,909 --> 00:24:00,870  
we make this assumption very frequently

612  
00:24:04,710 --> 00:24:01,919  
especially when it comes to

613  
00:24:07,590 --> 00:24:04,720

understanding your life we in some cases

614

00:24:09,110 --> 00:24:07,600

have to but we know that it's not that

615

00:24:11,190 --> 00:24:09,120

straightforward when it comes to

616

00:24:13,590 --> 00:24:11,200

understanding how mutations and

617

00:24:16,070 --> 00:24:13,600

variation impacts protein structure and

618

00:24:18,870 --> 00:24:16,080

how that consequently impacts protein

619

00:24:20,870 --> 00:24:18,880

function

620

00:24:22,789 --> 00:24:20,880

single insertions or deletions can

621

00:24:24,149 --> 00:24:22,799

change the behavior of a protein

622

00:24:25,830 --> 00:24:24,159

altogether

623

00:24:28,789 --> 00:24:25,840

we now know that it's not all about the

624

00:24:31,510 --> 00:24:28,799

active sites meaning the the heart of

625

00:24:33,830 --> 00:24:31,520

the protein where the reaction

626  
00:24:36,149 --> 00:24:33,840  
whatever the specific function happens

627  
00:24:38,310 --> 00:24:36,159  
we know they operate in a complex

628  
00:24:39,190 --> 00:24:38,320  
network inside the system inside the

629  
00:24:41,909 --> 00:24:39,200  
cell

630  
00:24:44,070 --> 00:24:41,919  
and there a single mutation that changes

631  
00:24:46,310 --> 00:24:44,080  
the structure of a protein can impact

632  
00:24:49,190 --> 00:24:46,320  
the entire cellular function because of

633  
00:24:51,350 --> 00:24:49,200  
this disturbance in the network

634  
00:24:55,190 --> 00:24:51,360  
a single insertion can be as dramatic as

635  
00:24:55,200 --> 00:24:59,430  
it's true

636  
00:25:04,070 --> 00:25:01,190  
with that in hindsight

637  
00:25:06,149 --> 00:25:04,080  
we work towards establishing a method

638  
00:25:07,750 --> 00:25:06,159

the idea was to start with protein

639

00:25:09,750 --> 00:25:07,760

sequence alignments

640

00:25:11,269 --> 00:25:09,760

form phylogenies

641

00:25:12,390 --> 00:25:11,279

we spend a lot of time on this and i

642

00:25:15,269 --> 00:25:12,400

will show

643

00:25:17,269 --> 00:25:15,279

an example uh how we utilize our

644

00:25:19,350 --> 00:25:17,279

assessment of protein to uh tree

645

00:25:21,350 --> 00:25:19,360

topology and our alignment output so i

646

00:25:22,710 --> 00:25:21,360

will be drilling down a little bit there

647

00:25:25,269 --> 00:25:22,720

towards the end

648

00:25:27,029 --> 00:25:25,279

well we infer the ancestral sequence

649

00:25:29,029 --> 00:25:27,039

this is where the synthetic biology part

650

00:25:31,350 --> 00:25:29,039

kicks in we can synthesize these genes

651  
00:25:33,830 --> 00:25:31,360  
it's getting cheaper every year to

652  
00:25:35,110 --> 00:25:33,840  
create these things it's not cheap but

653  
00:25:38,070 --> 00:25:35,120  
it's cheaper

654  
00:25:40,789 --> 00:25:38,080  
and it and in one level we purify these

655  
00:25:42,870 --> 00:25:40,799  
proteins to look at their uh behavior in

656  
00:25:44,470 --> 00:25:42,880  
vitro outside of the cell environment

657  
00:25:46,549 --> 00:25:44,480  
but what i really was interested in is

658  
00:25:47,750 --> 00:25:46,559  
again change the behavior of the

659  
00:25:50,310 --> 00:25:47,760  
organism

660  
00:25:51,110 --> 00:25:50,320  
to the degree that we can then

661  
00:25:54,149 --> 00:25:51,120  
um

662  
00:25:56,470 --> 00:25:54,159  
implement biogeochemical applications to

663  
00:25:58,310 --> 00:25:56,480

interpret or at least provide

664

00:26:00,390 --> 00:25:58,320

who knows an entirely independent data

665

00:26:02,950 --> 00:26:00,400

set using artificial

666

00:26:07,750 --> 00:26:02,960

biology in the lab to see what life can

667

00:26:11,669 --> 00:26:08,630

we

668

00:26:13,269 --> 00:26:11,679

started uh this is the translation was

669

00:26:16,390 --> 00:26:13,279

my first focus

670

00:26:19,510 --> 00:26:16,400

uh i started studying how this machinery

671

00:26:22,549 --> 00:26:19,520

that i just referred to you elongates

672

00:26:25,029 --> 00:26:22,559

peptides these are very uh not only they

673

00:26:27,510 --> 00:26:25,039

are singular in their own rights right

674

00:26:29,190 --> 00:26:27,520

there's only one translation machinery

675

00:26:31,110 --> 00:26:29,200

it may be different if you look at

676  
00:26:33,190 --> 00:26:31,120  
bacteria or archaea and eukaryotes there

677  
00:26:34,149 --> 00:26:33,200  
might be differences but there is only

678  
00:26:37,110 --> 00:26:34,159  
one

679  
00:26:39,350 --> 00:26:37,120  
computation center inside us inside all

680  
00:26:41,750 --> 00:26:39,360  
of us

681  
00:26:44,310 --> 00:26:41,760  
so i started by talking with the

682  
00:26:47,190 --> 00:26:44,320  
elongation step of this translation

683  
00:26:49,350 --> 00:26:47,200  
machinery inside bacteria i was curious

684  
00:26:51,350 --> 00:26:49,360  
to see what will it do

685  
00:26:53,430 --> 00:26:51,360  
to the entire cell

686  
00:26:55,909 --> 00:26:53,440  
well when you poke with something that

687  
00:26:58,950 --> 00:26:55,919  
is really essential turns out it breaks

688  
00:27:00,549 --> 00:26:58,960

so this was a very childlike i think

689

00:27:02,149 --> 00:27:00,559

venture you know how kids like to break

690

00:27:03,190 --> 00:27:02,159

things so we were just breaking cells

691

00:27:05,029 --> 00:27:03,200

all the time

692

00:27:07,669 --> 00:27:05,039

but i was interested in understanding

693

00:27:09,669 --> 00:27:07,679

how far can the cell tolerate the

694

00:27:11,830 --> 00:27:09,679

perturbance that we are introducing by

695

00:27:14,230 --> 00:27:11,840

poking

696

00:27:16,149 --> 00:27:14,240

and by introducing ancient genes that

697

00:27:18,389 --> 00:27:16,159

are older and older and translation is

698

00:27:20,710 --> 00:27:18,399

one of the first innovations of life

699

00:27:22,630 --> 00:27:20,720

anyway that dates at about four billion

700

00:27:24,230 --> 00:27:22,640

year right so we're

701  
00:27:27,110 --> 00:27:24,240  
messing with a four billion year old

702  
00:27:29,110 --> 00:27:27,120  
machine by asking it to interact and

703  
00:27:31,669 --> 00:27:29,120  
recognize a three billion year old

704  
00:27:37,750 --> 00:27:34,870  
after spending some time working on this

705  
00:27:39,510 --> 00:27:37,760  
system and then subjecting these perdarb

706  
00:27:41,430 --> 00:27:39,520  
systems to evolution to see how they

707  
00:27:43,110 --> 00:27:41,440  
will repeat themselves and i was

708  
00:27:45,430 --> 00:27:43,120  
interested in this replaying tape of

709  
00:27:47,830 --> 00:27:45,440  
life phenomena introduced by gold at

710  
00:27:51,590 --> 00:27:47,840  
that time

711  
00:27:53,909 --> 00:27:51,600  
recognize that we cannot directly link

712  
00:27:55,750 --> 00:27:53,919  
such engineering to the environment

713  
00:27:57,269 --> 00:27:55,760

you can break translation in very

714

00:27:59,350 --> 00:27:57,279

different ways that may have nothing to

715

00:28:01,350 --> 00:27:59,360

do with deep past

716

00:28:02,470 --> 00:28:01,360

so we need to focus on proteins that

717

00:28:04,389 --> 00:28:02,480

actually

718

00:28:07,190 --> 00:28:04,399

interact with the environment in a bit

719

00:28:09,029 --> 00:28:07,200

more direct way than the translation

720

00:28:12,710 --> 00:28:09,039

so while we are continuing that work for

721

00:28:14,470 --> 00:28:12,720

different purposes we moved on

722

00:28:16,149 --> 00:28:14,480

because i developed an experimental

723

00:28:17,669 --> 00:28:16,159

framework for reconstructing these

724

00:28:20,230 --> 00:28:17,679

metabolisms

725

00:28:21,909 --> 00:28:20,240

i thought okay here is what i learned

726  
00:28:24,310 --> 00:28:21,919  
we need to have

727  
00:28:25,830 --> 00:28:24,320  
a geologic history that we are tying the

728  
00:28:27,909 --> 00:28:25,840  
behavior of these

729  
00:28:30,149 --> 00:28:27,919  
engineered organisms to

730  
00:28:33,430 --> 00:28:30,159  
is there a fossil or a geochemical

731  
00:28:35,430 --> 00:28:33,440  
record related to this protein function

732  
00:28:38,310 --> 00:28:35,440  
there's vast amount of literature coming

733  
00:28:40,310 --> 00:28:38,320  
from biogeochemistry on these topics

734  
00:28:43,350 --> 00:28:40,320  
do we have a phenotype

735  
00:28:45,350 --> 00:28:43,360  
a characterizable and geologically

736  
00:28:47,269 --> 00:28:45,360  
relevant phenotype that we can generate

737  
00:28:49,909 --> 00:28:47,279  
in the lab

738  
00:28:51,669 --> 00:28:49,919

what are we linking this to

739

00:28:54,470 --> 00:28:51,679

is there an ancestry

740

00:28:57,350 --> 00:28:54,480

are we dealing with a tractable system

741

00:28:59,029 --> 00:28:57,360

how far can we extend it to in deep time

742

00:29:01,669 --> 00:28:59,039

is there an appropriate modern host

743

00:29:03,190 --> 00:29:01,679

biological system yes we can engineer a

744

00:29:04,389 --> 00:29:03,200

spinach that's not a problem but that

745

00:29:07,510 --> 00:29:04,399

doesn't tell you much if you're

746

00:29:09,750 --> 00:29:07,520

interested in arcane innovations

747

00:29:13,669 --> 00:29:09,760

do we have those model systems that will

748

00:29:15,909 --> 00:29:13,679

allow us to study the past

749

00:29:16,870 --> 00:29:15,919

with that we turned into many different

750

00:29:18,630 --> 00:29:16,880

venues

751  
00:29:19,669 --> 00:29:18,640  
one of this that i'm going to focus on

752  
00:29:22,950 --> 00:29:19,679  
is

753  
00:29:24,789 --> 00:29:22,960  
our way connecting protein evolution to

754  
00:29:27,430 --> 00:29:24,799  
mental availabilities of earth over

755  
00:29:28,389 --> 00:29:27,440  
geologic time

756  
00:29:30,149 --> 00:29:28,399  
there is a

757  
00:29:31,990 --> 00:29:30,159  
there are some authors that contributed

758  
00:29:34,789 --> 00:29:32,000  
to these papers that are already in this

759  
00:29:36,070 --> 00:29:34,799  
room and and this is really remarkable i

760  
00:29:39,510 --> 00:29:36,080  
think generation of what astro

761  
00:29:41,430 --> 00:29:39,520  
biologists also created over decades

762  
00:29:43,350 --> 00:29:41,440  
that we understand

763  
00:29:45,029 --> 00:29:43,360

to some degree that the metal

764

00:29:47,190 --> 00:29:45,039  
concentration

765

00:29:48,549 --> 00:29:47,200  
over geologic time did not remain

766

00:29:51,190 --> 00:29:48,559  
constant

767

00:29:54,230 --> 00:29:51,200  
for majority of these metals

768

00:29:56,549 --> 00:29:54,240  
we know this for iron that has a higher

769

00:29:58,230 --> 00:29:56,559  
concentration in the archaean then

770

00:30:01,029 --> 00:29:58,240  
there's a drop

771

00:30:03,029 --> 00:30:01,039  
and there's a lot of work into why such

772

00:30:05,750 --> 00:30:03,039  
changes happened the links of these

773

00:30:07,190 --> 00:30:05,760  
changes to great oxidation event or

774

00:30:10,710 --> 00:30:07,200  
different

775

00:30:12,630 --> 00:30:10,720  
earth surface related phenomena

776

00:30:13,510 --> 00:30:12,640

so this was very interesting

777

00:30:16,630 --> 00:30:13,520

uh

778

00:30:18,870 --> 00:30:16,640

to look at and i can tell you having

779

00:30:21,190 --> 00:30:18,880

interacting with a lot of metalloenzyme

780

00:30:22,549 --> 00:30:21,200

biochemists when i shove down this image

781

00:30:24,470 --> 00:30:22,559

they're mind blown

782

00:30:26,389 --> 00:30:24,480

they tell me they never thought about

783

00:30:28,549 --> 00:30:26,399

history of where these metals came from

784

00:30:30,630 --> 00:30:28,559

and i get it but that's what we do as

785

00:30:31,990 --> 00:30:30,640

astrobiologists right and that's amazing

786

00:30:33,110 --> 00:30:32,000

that we are able to connect these

787

00:30:36,870 --> 00:30:33,120

seemingly

788

00:30:38,870 --> 00:30:36,880

distant fields together

789

00:30:41,669 --> 00:30:38,880

nitrogenase is one of these

790

00:30:44,230 --> 00:30:41,679

enzymes a molybdenum dependent nitrogen

791

00:30:47,029 --> 00:30:44,240

fixing enzyme that also represents a

792

00:30:52,630 --> 00:30:50,149

it has a female code iron molybdenum

793

00:30:53,909 --> 00:30:52,640

co-factor there's different versions of

794

00:30:56,470 --> 00:30:53,919

this enzyme but this is the most

795

00:30:59,350 --> 00:30:56,480

prevalent version that exists today that

796

00:31:02,389 --> 00:30:59,360

fixes the advising inert nitrogen in the

797

00:31:07,669 --> 00:31:05,669

i think it is very amazing that life

798

00:31:08,950 --> 00:31:07,679

came up with this innovation once and

799

00:31:12,549 --> 00:31:08,960

has one

800

00:31:14,310 --> 00:31:12,559

protein to rely on to fix the nitrogen

801  
00:31:16,710 --> 00:31:14,320  
in the atmosphere at least so far that

802  
00:31:18,389 --> 00:31:16,720  
seems to be the case

803  
00:31:19,669 --> 00:31:18,399  
it's not very uh

804  
00:31:21,509 --> 00:31:19,679  
creative it is really putting all your

805  
00:31:26,149 --> 00:31:21,519  
eggs in one basket so we rely on this

806  
00:31:29,509 --> 00:31:26,950  
we

807  
00:31:32,950 --> 00:31:29,519  
looked at the literature to

808  
00:31:35,110 --> 00:31:32,960  
study the previous studies that tried to

809  
00:31:36,470 --> 00:31:35,120  
probe the evolution of nitrogenase over

810  
00:31:38,549 --> 00:31:36,480  
deep time

811  
00:31:41,110 --> 00:31:38,559  
and not surprisingly this came from

812  
00:31:44,070 --> 00:31:41,120  
astrobiologists here

813  
00:31:46,470 --> 00:31:44,080

yet we don't quite know the exact timing

814

00:31:47,990 --> 00:31:46,480

of nitrogen fixation

815

00:31:50,389 --> 00:31:48,000

there's there's a big question mark

816

00:31:51,430 --> 00:31:50,399

there but in the origin of nitrogenases

817

00:31:53,669 --> 00:31:51,440

we think

818

00:31:55,750 --> 00:31:53,679

must have coincided with the emergence

819

00:31:57,590 --> 00:31:55,760

of nitrogen fixation because that's what

820

00:31:59,830 --> 00:31:57,600

we see now

821

00:32:03,350 --> 00:31:59,840

and that makes sense at the top level as

822

00:32:07,509 --> 00:32:05,830

our questions were as follows what was

823

00:32:08,870 --> 00:32:07,519

the metal dependence of ancestral

824

00:32:11,509 --> 00:32:08,880

enzymes

825

00:32:15,029 --> 00:32:11,519

where did this enzyme come from

826

00:32:17,750 --> 00:32:15,039

how did it evolve

827

00:32:20,630 --> 00:32:17,760

and big shout out to dr amanda garcia

828

00:32:22,549 --> 00:32:20,640

who joined my lab i was a first-year

829

00:32:25,029 --> 00:32:22,559

faculty member at the university of

830

00:32:27,029 --> 00:32:25,039

arizona 2018

831

00:32:28,950 --> 00:32:27,039

and that's when she joined my lab as an

832

00:32:31,590 --> 00:32:28,960

a postdoctoral fellow

833

00:32:33,669 --> 00:32:31,600

to take this journey having been trained

834

00:32:37,110 --> 00:32:33,679

in in paleobiology

835

00:32:40,830 --> 00:32:37,120

uh this was amazing for me to have

836

00:32:44,549 --> 00:32:40,840

someone coming directly from the fields

837

00:32:46,470 --> 00:32:44,559

in trained in paleontology at ucla

838

00:32:48,630 --> 00:32:46,480

to join our group to study molecular

839

00:32:50,710 --> 00:32:48,640

revolution system it was exactly the

840

00:32:53,669 --> 00:32:50,720

kind of scholars that i was hoping to

841

00:32:59,110 --> 00:32:57,269

amanda asked the question can we create

842

00:33:01,190 --> 00:32:59,120

phylogenies and now

843

00:33:02,950 --> 00:33:01,200

on top of building on the previous

844

00:33:04,549 --> 00:33:02,960

studies

845

00:33:05,430 --> 00:33:04,559

a lot of these work came from montana

846

00:33:06,389 --> 00:33:05,440

state

847

00:33:07,590 --> 00:33:06,399

can we

848

00:33:10,549 --> 00:33:07,600

use the

849

00:33:13,669 --> 00:33:10,559

upgraded updated sequence information in

850

00:33:16,789 --> 00:33:13,679

the database rebuilt the tree and it not

851  
00:33:19,669 --> 00:33:16,799  
only read the tree that we created but

852  
00:33:21,669 --> 00:33:19,679  
also reconstruct the ancestors using

853  
00:33:22,870 --> 00:33:21,679  
this tree to understand

854  
00:33:28,070 --> 00:33:22,880  
the

855  
00:33:30,149 --> 00:33:28,080  
preference of this enzyme

856  
00:33:32,389 --> 00:33:30,159  
she built machine learning models this

857  
00:33:35,269 --> 00:33:32,399  
was a computational study that were

858  
00:33:39,830 --> 00:33:35,279  
trained on the active site amino acids

859  
00:33:41,430 --> 00:33:39,840  
of old nitrogenases that exist today

860  
00:33:42,630 --> 00:33:41,440  
because not all of them prefer

861  
00:33:44,870 --> 00:33:42,640  
molybdenum

862  
00:33:48,310 --> 00:33:44,880  
we have others such as vanadium

863  
00:33:51,029 --> 00:33:48,320

preferring nitrogenases to come up with

864

00:33:53,509 --> 00:33:51,039

a classification to determine whether

865

00:33:57,269 --> 00:33:53,519

the ancestral metal dependence focuses

866

00:33:59,909 --> 00:33:57,279

on the molybdenum or not

867

00:34:02,870 --> 00:33:59,919

we published this paper

868

00:34:04,870 --> 00:34:02,880

as our indicated what indicate that what

869

00:34:07,590 --> 00:34:04,880

the output of these analysis indicate

870

00:34:09,750 --> 00:34:07,600

that ancestral nitrogenases were in fact

871

00:34:11,109 --> 00:34:09,760

molybdenum periphery

872

00:34:12,470 --> 00:34:11,119

which wasn't

873

00:34:14,790 --> 00:34:12,480

in contradiction with what was

874

00:34:16,869 --> 00:34:14,800

previously thought

875

00:34:19,430 --> 00:34:16,879

we moved on to our second question where

876

00:34:22,310 --> 00:34:19,440

did this enzyme come from

877

00:34:24,310 --> 00:34:22,320

again we just because nitrogen fixation

878

00:34:26,389 --> 00:34:24,320

or something like it's

879

00:34:27,829 --> 00:34:26,399

existed and and that we read these

880

00:34:29,589 --> 00:34:27,839

signatures in the rug record doesn't

881

00:34:31,589 --> 00:34:29,599

mean and

882

00:34:33,589 --> 00:34:31,599

nitrogenase as we know today existed

883

00:34:35,190 --> 00:34:33,599

back then don't we love to talk about

884

00:34:37,510 --> 00:34:35,200

life is we don't know it

885

00:34:39,430 --> 00:34:37,520

this is an opportunity for us to explore

886

00:34:41,190 --> 00:34:39,440

a potential life as we don't know we

887

00:34:43,030 --> 00:34:41,200

should not make these assumptions so we

888

00:34:45,430 --> 00:34:43,040

asked the question where did this enzyme

889

00:34:48,790 --> 00:34:45,440

come from did early nitrogenase was in

890

00:34:50,710 --> 00:34:48,800

fact a nitrogenase

891

00:34:55,270 --> 00:34:50,720

for this we looked at the partnering

892

00:35:00,950 --> 00:34:58,470

the protein network that enables

893

00:35:03,270 --> 00:35:00,960

this enzyme to have a metal factor

894

00:35:05,109 --> 00:35:03,280

requires match raises as the name goes

895

00:35:07,030 --> 00:35:05,119

the mature they sort of

896

00:35:08,630 --> 00:35:07,040

you can imagine them as maturing the

897

00:35:10,069 --> 00:35:08,640

metal for the nitrogen they're sort of

898

00:35:11,030 --> 00:35:10,079

feeding them

899

00:35:15,270 --> 00:35:11,040

in

900

00:35:17,349 --> 00:35:15,280

so we looked at their partners sometimes

901

00:35:18,630 --> 00:35:17,359

what you're looking for is right

902

00:35:20,790 --> 00:35:18,640

under your nose

903

00:35:22,710 --> 00:35:20,800

and the person

904

00:35:24,310 --> 00:35:22,720

or in this case the protein that is get

905

00:35:27,190 --> 00:35:24,320

doing the most amount of the majority of

906

00:35:30,950 --> 00:35:27,200

the work may not get the credits

907

00:35:32,950 --> 00:35:30,960

so nitrogenases we find through

908

00:35:35,910 --> 00:35:32,960

creating a tree a joint tree of

909

00:35:38,390 --> 00:35:35,920

nitrogenases and match raises we we did

910

00:35:40,710 --> 00:35:38,400

a lot of tree reconstructions made

911

00:35:43,670 --> 00:35:40,720

different models we submitted this paper

912

00:35:45,829 --> 00:35:43,680

to the society of the of the molecular

913

00:35:48,950 --> 00:35:45,839

evolution we want to make sure that we

914

00:35:51,030 --> 00:35:48,960

get the the best reviewer feedback we

915

00:35:53,829 --> 00:35:51,040

can get and we did it was a great

916

00:35:57,829 --> 00:35:56,150

after we reconstructed the ancestor and

917

00:35:59,990 --> 00:35:57,839

studying it we found out that the

918

00:36:02,550 --> 00:36:00,000

ancestor was more like the match rays

919

00:36:07,510 --> 00:36:02,560

and not nitrogenase

920

00:36:10,069 --> 00:36:07,520

so they may have evolved from their body

921

00:36:12,550 --> 00:36:10,079

meaning the ancestor wasn't

922

00:36:15,270 --> 00:36:12,560

perhaps even fixing nitro nitrogen the

923

00:36:17,670 --> 00:36:15,280

way today's nitrogenases are doing these

924

00:36:20,230 --> 00:36:17,680

are all models so i cannot assert uh

925

00:36:23,030 --> 00:36:20,240

these outcomes just yet

926  
00:36:25,430 --> 00:36:23,040  
but if this is true

927  
00:36:28,470 --> 00:36:25,440  
what may might be the case then is that

928  
00:36:31,750 --> 00:36:28,480  
the molecular architecture

929  
00:36:33,670 --> 00:36:31,760  
required for the birth of this

930  
00:36:36,310 --> 00:36:33,680  
nitrogenase enzyme

931  
00:36:38,710 --> 00:36:36,320  
was unlikely shaped by the same

932  
00:36:41,109 --> 00:36:38,720  
environmental drivers that are

933  
00:36:43,030 --> 00:36:41,119  
implicated in the evolution of nitrogen

934  
00:36:45,829 --> 00:36:43,040  
fixation

935  
00:36:48,150 --> 00:36:45,839  
so this is not a

936  
00:36:49,990 --> 00:36:48,160  
judging the book by its cover situation

937  
00:36:52,230 --> 00:36:50,000  
just because you have

938  
00:36:55,109 --> 00:36:52,240

certain elements or metal in the

939

00:36:56,630 --> 00:36:55,119

environment doesn't mean life will

940

00:36:58,950 --> 00:36:56,640

life adopted that just because it was in

941

00:37:01,589 --> 00:36:58,960

the environment there's definitely

942

00:37:03,190 --> 00:37:01,599

we have to consider evolutionary forces

943

00:37:04,710 --> 00:37:03,200

and the laws of biology then we

944

00:37:06,310 --> 00:37:04,720

interpret

945

00:37:08,710 --> 00:37:06,320

the timing and the origins of such

946

00:37:11,030 --> 00:37:08,720

events

947

00:37:12,870 --> 00:37:11,040

we moved on to our third question

948

00:37:14,790 --> 00:37:12,880

all right so how did this evolve it's

949

00:37:15,990 --> 00:37:14,800

great to do computational work but we

950

00:37:17,990 --> 00:37:16,000

have to

951  
00:37:20,950 --> 00:37:18,000  
get into the lab so we rolled up our

952  
00:37:25,109 --> 00:37:20,960  
sleeves and back back to our

953  
00:37:27,270 --> 00:37:25,119  
paleo phenotype assessment methodology

954  
00:37:29,430 --> 00:37:27,280  
we have the reconstructed ancestor let's

955  
00:37:30,710 --> 00:37:29,440  
engineer this inside a modern microbe

956  
00:37:31,910 --> 00:37:30,720  
and let's be

957  
00:37:33,829 --> 00:37:31,920  
calm here

958  
00:37:35,190 --> 00:37:33,839  
we have to work with an organism that is

959  
00:37:37,109 --> 00:37:35,200  
well studied because we don't want to

960  
00:37:38,550 --> 00:37:37,119  
reinvent an entire engineering pipeline

961  
00:37:39,829 --> 00:37:38,560  
for a new organism and we are very

962  
00:37:42,150 --> 00:37:39,839  
grateful that a lot of amazing

963  
00:37:44,630 --> 00:37:42,160

scientists are doing that

964

00:37:47,910 --> 00:37:44,640

let's engineer this ancient dna inside a

965

00:37:53,030 --> 00:37:47,920

modern organism and then study what

966

00:37:58,310 --> 00:37:56,069

we've picked azodo bacteria this is the

967

00:38:01,430 --> 00:37:58,320

prime nitrogen-fixing bacteria it's a

968

00:38:03,030 --> 00:38:01,440

huge focus for a lot of studies that

969

00:38:04,630 --> 00:38:03,040

focus on

970

00:38:06,950 --> 00:38:04,640

it even goes to climate change

971

00:38:10,470 --> 00:38:06,960

implications and crop engineering and

972

00:38:13,349 --> 00:38:10,480

plants biologists love this enzyme

973

00:38:14,950 --> 00:38:13,359

but yet again uh and due to these

974

00:38:16,630 --> 00:38:14,960

there's really

975

00:38:19,030 --> 00:38:16,640

there was a lot of development in the

976

00:38:21,030 --> 00:38:19,040

engineering of this organism

977

00:38:23,190 --> 00:38:21,040

at the genomic level primarily developed

978

00:38:25,109 --> 00:38:23,200

by dennis steen at virginia tech

979

00:38:27,030 --> 00:38:25,119

so we visited his lab to learn these

980

00:38:29,750 --> 00:38:27,040

skills we brought them back to the lab

981

00:38:32,470 --> 00:38:29,760

we tweaked them for our own purposes

982

00:38:35,990 --> 00:38:32,480

so we did the following you are looking

983

00:38:38,230 --> 00:38:36,000

at the the purple and the the blue uh

984

00:38:40,790 --> 00:38:38,240

part of the screen just focus on that

985

00:38:42,870 --> 00:38:40,800

for now what we did is that we wanted

986

00:38:44,950 --> 00:38:42,880

nitrogenase is encoded by

987

00:38:46,310 --> 00:38:44,960

three different genes if hdk at least

988

00:38:48,710 --> 00:38:46,320

these are the ones that we focused for

989

00:38:51,270 --> 00:38:48,720

this study we wanted to know

990

00:38:52,710 --> 00:38:51,280

whether we can swap the entire

991

00:38:55,589 --> 00:38:52,720

genomic

992

00:38:57,190 --> 00:38:55,599

architecture with ancestral dna or we

993

00:38:59,750 --> 00:38:57,200

need whether we should follow

994

00:39:01,589 --> 00:38:59,760

a stepwise modularity so we have to do

995

00:39:03,349 --> 00:39:01,599

these things in order to set the system

996

00:39:05,349 --> 00:39:03,359

up because we have to understand the

997

00:39:06,710 --> 00:39:05,359

limits of what we're dealing with too

998

00:39:09,109 --> 00:39:06,720

and of course the majority of the

999

00:39:10,710 --> 00:39:09,119

question is that how ancient is this

1000

00:39:12,950 --> 00:39:10,720

it is going to be as ancient as we can

1001

00:39:16,710 --> 00:39:12,960

make it at this point

1002

00:39:18,310 --> 00:39:16,720

so on on one level amanda engineered the

1003

00:39:21,190 --> 00:39:18,320

nif d

1004

00:39:23,750 --> 00:39:21,200

with its ancestor that's one gene

1005

00:39:25,349 --> 00:39:23,760

out of the cluster and then in parallel

1006

00:39:28,150 --> 00:39:25,359

she engineered the entire cluster with

1007

00:39:30,069 --> 00:39:28,160

the entire ancient gene coming from the

1008

00:39:31,750 --> 00:39:30,079

tree and i'm pointing out where this is

1009

00:39:34,470 --> 00:39:31,760

coming from we are not yet at the

1010

00:39:36,550 --> 00:39:34,480

deepest part of the tree we are not yet

1011

00:39:37,910 --> 00:39:36,560

at the deep archaean or

1012

00:39:40,069 --> 00:39:37,920

not even there

1013

00:39:42,470 --> 00:39:40,079

we are still within uh

1014

00:39:45,109 --> 00:39:42,480

roughly a billion year maybe if i'm

1015

00:39:47,589 --> 00:39:45,119

generous so we are not traveling deep

1016

00:39:49,510 --> 00:39:47,599

time yet yet i mean a billion is pretty

1017

00:39:52,390 --> 00:39:49,520

long time

1018

00:39:55,990 --> 00:39:52,400

so we want to test the effects

1019

00:39:57,349 --> 00:39:56,000

of the ancestor h and ancestral subunits

1020

00:39:59,829 --> 00:39:57,359

and i'm very happy to share these

1021

00:40:01,990 --> 00:39:59,839

results i've actually just posted this

1022

00:40:03,589 --> 00:40:02,000

uh it will be online today so these are

1023

00:40:05,829 --> 00:40:03,599

the precious results that you're seeing

1024

00:40:08,230 --> 00:40:05,839

it's my first time showing them that we

1025

00:40:11,030 --> 00:40:08,240

resurrected these nitrogenases inside

1026  
00:40:13,430 --> 00:40:11,040  
modern bacteria on you're looking at the

1027  
00:40:15,910 --> 00:40:13,440  
growth curve the optical density plot on

1028  
00:40:18,710 --> 00:40:15,920  
the y-axis and towards time you're

1029  
00:40:19,750 --> 00:40:18,720  
basically assessing the growth of these

1030  
00:40:21,670 --> 00:40:19,760  
bugs

1031  
00:40:23,270 --> 00:40:21,680  
and i must say that

1032  
00:40:25,750 --> 00:40:23,280  
one thing we have in universities are

1033  
00:40:28,069 --> 00:40:25,760  
these amazing undergraduates big shout

1034  
00:40:30,790 --> 00:40:28,079  
out to arizona space program and brook

1035  
00:40:33,510 --> 00:40:30,800  
characters who spent her entire freshman

1036  
00:40:36,470 --> 00:40:33,520  
and sophomore year optimizing a growth

1037  
00:40:38,470 --> 00:40:36,480  
curve a plot she's just fantastic and

1038  
00:40:40,390 --> 00:40:38,480

she did this we tried

1039

00:40:42,309 --> 00:40:40,400

batch cultures we tried test tubes we

1040

00:40:43,990 --> 00:40:42,319

tried um

1041

00:40:46,710 --> 00:40:44,000

multi-valve plates

1042

00:40:48,710 --> 00:40:46,720

uh and what worked was

1043

00:40:50,550 --> 00:40:48,720

the oldest most traditional method is

1044

00:40:53,030 --> 00:40:50,560

what ended up working

1045

00:40:55,030 --> 00:40:53,040

so we have we followed the growth of

1046

00:40:57,030 --> 00:40:55,040

these engineered strains

1047

00:40:58,470 --> 00:40:57,040

uh it it worked fine because that's how

1048

00:40:59,990 --> 00:40:58,480

long it took us to engineer these

1049

00:41:02,470 --> 00:41:00,000

strains anyway so by the time we

1050

00:41:04,390 --> 00:41:02,480

perfected the or at least optimized the

1051

00:41:06,390 --> 00:41:04,400

growth assessments our

1052

00:41:07,750 --> 00:41:06,400

engineered strains were ready

1053

00:41:10,230 --> 00:41:07,760

on the black line you're looking at the

1054

00:41:12,710 --> 00:41:10,240

wild type that's the sort of modern

1055

00:41:13,990 --> 00:41:12,720

native unengineered unperturbed

1056

00:41:16,550 --> 00:41:14,000

bacteria

1057

00:41:18,710 --> 00:41:16,560

delta represents deletion that we do we

1058

00:41:20,870 --> 00:41:18,720

delete these these one of these genes

1059

00:41:23,670 --> 00:41:20,880

inside out of the chromosome just to

1060

00:41:26,230 --> 00:41:23,680

make sure uh just as a control

1061

00:41:28,150 --> 00:41:26,240

and all the other blue and purple lines

1062

00:41:31,030 --> 00:41:28,160

represents the strains with the

1063

00:41:33,349 --> 00:41:31,040

ancestral nitrogenases

1064

00:41:36,069 --> 00:41:33,359

after we engineered these strains we've

1065

00:41:37,270 --> 00:41:36,079

also looked at the relative organismal

1066

00:41:39,990 --> 00:41:37,280

activity

1067

00:41:43,270 --> 00:41:40,000

uh that and by looking at the acetylene

1068

00:41:44,630 --> 00:41:43,280

ethylene reduction so it we are still

1069

00:41:48,069 --> 00:41:44,640

assessing the

1070

00:41:50,230 --> 00:41:48,079

behavior of the entire organism

1071

00:41:52,710 --> 00:41:50,240

and you're looking at the reduction rate

1072

00:41:56,309 --> 00:41:52,720

on the y-axis we normalize this to total

1073

00:41:59,430 --> 00:41:56,319

protein and different uh strains on the

1074

00:42:00,390 --> 00:41:59,440

x-axis and this was performed by alexa

1075

00:42:01,750 --> 00:42:00,400

va

1076  
00:42:05,349 --> 00:42:01,760  
and together with of course amanda

1077  
00:42:07,750 --> 00:42:05,359  
garcia and azul pinochet barros

1078  
00:42:10,150 --> 00:42:07,760  
so you might be seeing what i see

1079  
00:42:12,550 --> 00:42:10,160  
that these bacteria that contain the

1080  
00:42:15,190 --> 00:42:12,560  
ancient dna are active

1081  
00:42:17,270 --> 00:42:15,200  
and it's they're not

1082  
00:42:18,950 --> 00:42:17,280  
vast differences at least based on our

1083  
00:42:20,550 --> 00:42:18,960  
experience since we've done these with

1084  
00:42:23,190 --> 00:42:20,560  
other systems as well

1085  
00:42:26,390 --> 00:42:23,200  
fairly comparable reduction rates

1086  
00:42:27,750 --> 00:42:26,400  
compared to the modern

1087  
00:42:29,750 --> 00:42:27,760  
we didn't

1088  
00:42:31,349 --> 00:42:29,760

stop there and just

1089

00:42:34,630 --> 00:42:31,359

quick shout out amanda will be talking

1090

00:42:37,430 --> 00:42:34,640

about this tomorrow in detail

1091

00:42:39,349 --> 00:42:37,440

as i said these were still inside the

1092

00:42:41,990 --> 00:42:39,359

organism right so we are still inside

1093

00:42:44,470 --> 00:42:42,000

the genome and and yet by the way we do

1094

00:42:46,069 --> 00:42:44,480

sequence the entire genome to make sure

1095

00:42:49,190 --> 00:42:46,079

we don't have any other

1096

00:42:51,349 --> 00:42:49,200

uh y type fragments left anywhere

1097

00:42:53,990 --> 00:42:51,359

that we ensure that there's no other

1098

00:42:56,309 --> 00:42:54,000

modern nitrogenase dna

1099

00:42:59,190 --> 00:42:56,319

that is remaining in this system to give

1100

00:43:01,910 --> 00:42:59,200

us any false positives

1101  
00:43:03,750 --> 00:43:01,920  
moving on we wanted to understand what

1102  
00:43:06,470 --> 00:43:03,760  
does this

1103  
00:43:08,390 --> 00:43:06,480  
enzyme what would it do outside of the

1104  
00:43:10,550 --> 00:43:08,400  
cell environment so we extracted this

1105  
00:43:12,870 --> 00:43:10,560  
protein from a zodobacter after

1106  
00:43:14,790 --> 00:43:12,880  
overexpressing it and i ran this through

1107  
00:43:17,349 --> 00:43:14,800  
proof by column and this was done by

1108  
00:43:19,589 --> 00:43:17,359  
derek harris also a remarkable postdoc

1109  
00:43:23,270 --> 00:43:19,599  
who spent a lot of time purifying these

1110  
00:43:27,190 --> 00:43:24,950  
followed by

1111  
00:43:29,750 --> 00:43:27,200  
enzymatic kinetics to understand the

1112  
00:43:31,670 --> 00:43:29,760  
catalytic function of these enzymes

1113  
00:43:33,270 --> 00:43:31,680

on the um you're looking at the nitrogen

1114

00:43:36,150 --> 00:43:33,280

reduction rates

1115

00:43:39,190 --> 00:43:36,160

when different uh gas head spaces were

1116

00:43:41,349 --> 00:43:39,200

used and and derek has published a paper

1117

00:43:43,829 --> 00:43:41,359

previously to assess

1118

00:43:46,309 --> 00:43:43,839

a mechanism to understand the modern

1119

00:43:47,910 --> 00:43:46,319

nitrogenous mechanisms and and that

1120

00:43:50,230 --> 00:43:47,920

there's a reductive elimination

1121

00:43:53,030 --> 00:43:50,240

mechanism for the nitrogen binding being

1122

00:43:55,109 --> 00:43:53,040

the the core sort of characteristics for

1123

00:43:57,030 --> 00:43:55,119

modern nitrogenases

1124

00:43:58,950 --> 00:43:57,040

but what he is finding is that the

1125

00:44:02,150 --> 00:43:58,960

efficiency is similar at the catalytic

1126

00:44:04,710 --> 00:44:02,160

level to modern nitrogenases as well

1127

00:44:06,870 --> 00:44:04,720

and you're looking at the structures uh

1128

00:44:10,870 --> 00:44:06,880

these are the ratios

1129

00:44:13,109 --> 00:44:10,880

for the the mechanism output and and

1130

00:44:14,470 --> 00:44:13,119

that gives us a clue about the catalytic

1131

00:44:16,870 --> 00:44:14,480

efficiency

1132

00:44:19,109 --> 00:44:16,880

of the extent molybdenum and

1133

00:44:20,950 --> 00:44:19,119

nitrogenases uh compared to the

1134

00:44:23,109 --> 00:44:20,960

ancestral ones

1135

00:44:25,190 --> 00:44:23,119

so it's pretty good

1136

00:44:27,270 --> 00:44:25,200

and i've been working on nitrogenous

1137

00:44:29,510 --> 00:44:27,280

experts i am now in a department where

1138

00:44:31,270 --> 00:44:29,520

these enzymes were even discovered and

1139

00:44:34,470 --> 00:44:31,280

as other bacteria was discovered as an

1140

00:44:37,430 --> 00:44:34,480

organism and they tell me that it's very

1141

00:44:39,910 --> 00:44:37,440

easy to break these enzymes or at least

1142

00:44:42,390 --> 00:44:39,920

they're not as hard as i think

1143

00:44:45,270 --> 00:44:42,400

so the fact that we have in some cases

1144

00:44:46,950 --> 00:44:45,280

up to 20 25 percent difference in amino

1145

00:44:51,270 --> 00:44:46,960

acid and we're still able to recover

1146

00:44:55,910 --> 00:44:53,510

just to wrap up this part the ancient

1147

00:44:58,790 --> 00:44:55,920

environmental conditions do not fully

1148

00:45:01,030 --> 00:44:58,800

explain the timing

1149

00:45:03,349 --> 00:45:01,040

of critical and singular biogeochemical

1150

00:45:05,910 --> 00:45:03,359

innovations

1151  
00:45:07,910 --> 00:45:05,920  
alone at least and these experimental

1152  
00:45:10,230 --> 00:45:07,920  
studies i think will be very critical to

1153  
00:45:12,790 --> 00:45:10,240  
understand the molecular innovations

1154  
00:45:16,150 --> 00:45:12,800  
of life's earliest traces

1155  
00:45:18,390 --> 00:45:16,160  
and we have done some work as

1156  
00:45:21,109 --> 00:45:18,400  
frank mentioned where we resurrected

1157  
00:45:22,630 --> 00:45:21,119  
rubisco that is also to be thought to be

1158  
00:45:23,430 --> 00:45:22,640  
one of those major

1159  
00:45:26,069 --> 00:45:23,440  
uh

1160  
00:45:27,910 --> 00:45:26,079  
enzymes impacting carbon isotope a

1161  
00:45:30,470 --> 00:45:27,920  
fractionation record throughout earth's

1162  
00:45:34,069 --> 00:45:30,480  
history uh and because that work is

1163  
00:45:36,470 --> 00:45:34,079

published a few a couple weeks ago

1164

00:45:39,670 --> 00:45:36,480

i opted to present the latest one so

1165

00:45:43,829 --> 00:45:41,990

but moving on where are we going with

1166

00:45:45,670 --> 00:45:43,839

this of course this is all laboratory

1167

00:45:48,309 --> 00:45:45,680

artificial life there's a lot of

1168

00:45:50,230 --> 00:45:48,319

molecular component here

1169

00:45:52,390 --> 00:45:50,240

and keep in mind that integration of

1170

00:45:53,750 --> 00:45:52,400

evolution with biochemistry itself is

1171

00:45:56,790 --> 00:45:53,760

still pretty

1172

00:45:59,349 --> 00:45:56,800

new grounds for for biochemists

1173

00:46:01,430 --> 00:45:59,359

that that we miss so much opportunity we

1174

00:46:03,670 --> 00:46:01,440

just when we don't understand

1175

00:46:05,670 --> 00:46:03,680

where life came from but not just for

1176  
00:46:07,829 --> 00:46:05,680  
astrobiologists but when we study what

1177  
00:46:10,150 --> 00:46:07,839  
we have it's true that we have to

1178  
00:46:11,270 --> 00:46:10,160  
understand the past to make sense of the

1179  
00:46:13,510 --> 00:46:11,280  
present

1180  
00:46:15,910 --> 00:46:13,520  
we we must do this and and we should not

1181  
00:46:18,550 --> 00:46:15,920  
miss this opportunity to harvest all the

1182  
00:46:20,550 --> 00:46:18,560  
sequin space and why life and understand

1183  
00:46:22,230 --> 00:46:20,560  
why life narrowed it down to a subset of

1184  
00:46:24,470 --> 00:46:22,240  
what we study today

1185  
00:46:27,589 --> 00:46:24,480  
but this is where muse comes in because

1186  
00:46:29,109 --> 00:46:27,599  
my desire was to connect these things to

1187  
00:46:30,470 --> 00:46:29,119  
actual earth

1188  
00:46:31,670 --> 00:46:30,480

real life

1189

00:46:34,390 --> 00:46:31,680

to

1190

00:46:36,950 --> 00:46:34,400

sites where people go and study the

1191

00:46:38,630 --> 00:46:36,960

analogues for early earth environments

1192

00:46:41,589 --> 00:46:38,640

where they call the culture these

1193

00:46:43,670 --> 00:46:41,599

microbes if possible and study them so

1194

00:46:44,390 --> 00:46:43,680

that was my desire to be have to work

1195

00:46:48,710 --> 00:46:44,400

with

1196

00:46:51,670 --> 00:46:48,720

very important i think for astrobiology

1197

00:46:54,550 --> 00:46:51,680

because when we assess the habitability

1198

00:46:55,430 --> 00:46:54,560

or um or really when we analyze

1199

00:47:00,710 --> 00:46:55,440

different

1200

00:47:02,870 --> 00:47:00,720

or whatever we are visiting we are

1201  
00:47:05,270 --> 00:47:02,880  
focusing on a problem that is existing

1202  
00:47:06,950 --> 00:47:05,280  
largely in an organic chemistry space

1203  
00:47:08,390 --> 00:47:06,960  
but we need to integrate an

1204  
00:47:10,230 --> 00:47:08,400  
understanding of the elemental

1205  
00:47:15,829 --> 00:47:10,240  
composition

1206  
00:47:17,670 --> 00:47:15,839  
habitability as well so we are looking

1207  
00:47:20,230 --> 00:47:17,680  
and working towards finding ways to

1208  
00:47:22,470 --> 00:47:20,240  
transform this into a problem that can

1209  
00:47:25,990 --> 00:47:22,480  
be studied in inorganic chemistry space

1210  
00:47:31,190 --> 00:47:28,790  
and and perhaps uh

1211  
00:47:32,630 --> 00:47:31,200  
you might be wondering this is great and

1212  
00:47:33,349 --> 00:47:32,640  
what you're what you're studying is at

1213  
00:47:35,750 --> 00:47:33,359

the

1214

00:47:37,670 --> 00:47:35,760

time scale that is relevant for earth

1215

00:47:38,630 --> 00:47:37,680

but how do we make this an observation

1216

00:47:40,870 --> 00:47:38,640

problem

1217

00:47:43,510 --> 00:47:40,880

how can i as an astronomer study this

1218

00:47:45,750 --> 00:47:43,520

what is the use sorry for exoplanets or

1219

00:47:47,750 --> 00:47:45,760

different observation

1220

00:47:49,589 --> 00:47:47,760

so in the remainder of time for

1221

00:47:50,950 --> 00:47:49,599

quickly i want to share one of our

1222

00:47:53,750 --> 00:47:50,960

latest papers

1223

00:47:55,349 --> 00:47:53,760

we thought about this we are largely at

1224

00:47:57,750 --> 00:47:55,359

the geochemistry and evolutionary

1225

00:48:00,150 --> 00:47:57,760

biochemistry space and the blue and the

1226  
00:48:01,910 --> 00:48:00,160  
green circles but what we need is a more

1227  
00:48:03,349 --> 00:48:01,920  
integration of the stellar atmospheric

1228  
00:48:06,870 --> 00:48:03,359  
chemistry

1229  
00:48:09,670 --> 00:48:06,880  
so the idea was then can we find

1230  
00:48:12,069 --> 00:48:09,680  
molecules proteins that not only impact

1231  
00:48:15,910 --> 00:48:12,079  
the earth and fill in all the criteria

1232  
00:48:19,270 --> 00:48:15,920  
that i just presented but also may be

1233  
00:48:22,309 --> 00:48:19,280  
maybe of use throughout integration of

1234  
00:48:24,470 --> 00:48:22,319  
say surface irradiance studies or

1235  
00:48:27,030 --> 00:48:24,480  
more generally

1236  
00:48:28,870 --> 00:48:27,040  
observation of gases by a signature

1237  
00:48:29,829 --> 00:48:28,880  
level assessment so with that we turned

1238  
00:48:31,829 --> 00:48:29,839

into

1239

00:48:34,150 --> 00:48:31,839

pigments

1240

00:48:37,349 --> 00:48:34,160

our pursuit was to tie together

1241

00:48:39,670 --> 00:48:37,359

bioinformatic analysis geochemical and

1242

00:48:41,990 --> 00:48:39,680

solar spectral models

1243

00:48:44,309 --> 00:48:42,000

atmospheric attenuation in particular

1244

00:48:46,309 --> 00:48:44,319

here radiative transfer calculations and

1245

00:48:48,069 --> 00:48:46,319

enzymatic observation

1246

00:48:50,390 --> 00:48:48,079

simple

1247

00:48:52,309 --> 00:48:50,400

characterize the coupling between the

1248

00:48:55,190 --> 00:48:52,319

internal this could be as internal as

1249

00:48:58,150 --> 00:48:55,200

the inside in cellular

1250

00:49:00,630 --> 00:48:58,160

or population and external conditions

1251  
00:49:03,829 --> 00:49:00,640  
and infer biosignature characteristics

1252  
00:49:05,430 --> 00:49:03,839  
generated by enzymes that have changed

1253  
00:49:07,430 --> 00:49:05,440  
over time

1254  
00:49:09,670 --> 00:49:07,440  
one of the first

1255  
00:49:11,270 --> 00:49:09,680  
enzymes we focused is rhodopsin these

1256  
00:49:12,710 --> 00:49:11,280  
are interesting

1257  
00:49:14,230 --> 00:49:12,720  
they're pigment

1258  
00:49:16,710 --> 00:49:14,240  
they're really

1259  
00:49:18,390 --> 00:49:16,720  
actually fascinating they look simple

1260  
00:49:20,470 --> 00:49:18,400  
but they are not simpler

1261  
00:49:22,950 --> 00:49:20,480  
they are tuned to the peak photon

1262  
00:49:24,549 --> 00:49:22,960  
frequencies generated by the sun

1263  
00:49:26,870 --> 00:49:24,559

and and this is really fascinating

1264

00:49:29,670 --> 00:49:26,880

especially for biologists i think that

1265

00:49:31,430 --> 00:49:29,680

when i'm from black sea so i was always

1266

00:49:33,750 --> 00:49:31,440

fascinated by the depths of black sea

1267

00:49:35,190 --> 00:49:33,760

and how organisms deep that and by the

1268

00:49:37,190 --> 00:49:35,200

way it really does look black if you

1269

00:49:39,589 --> 00:49:37,200

haven't visited it it's just remarkable

1270

00:49:42,630 --> 00:49:39,599

that these microbes what they can do to

1271

00:49:44,630 --> 00:49:42,640

harvest the limited of the environmental

1272

00:49:46,950 --> 00:49:44,640

energy that they got and they come up

1273

00:49:48,790 --> 00:49:46,960

with the most innovative ways to make

1274

00:49:51,030 --> 00:49:48,800

use of the harsh conditions that they

1275

00:49:53,030 --> 00:49:51,040

reside in a phenomena that i think most

1276

00:49:54,390 --> 00:49:53,040

of us in this room can relate to or some

1277

00:49:55,589 --> 00:49:54,400

of us

1278

00:49:57,430 --> 00:49:55,599

so we have to come up with these

1279

00:50:00,470 --> 00:49:57,440

innovative ways to survive life does

1280

00:50:03,109 --> 00:50:00,480

this so here is rhodopsin catherine

1281

00:50:04,630 --> 00:50:03,119

cephas this was a perfect student she

1282

00:50:07,510 --> 00:50:04,640

was an undergrad

1283

00:50:10,630 --> 00:50:07,520

at the astronomy department who also

1284

00:50:13,190 --> 00:50:10,640

loved biology so she came to me saying i

1285

00:50:16,549 --> 00:50:13,200

always want to be a biologist but i also

1286

00:50:17,349 --> 00:50:16,559

always want to be an astronomer i said

1287

00:50:19,750 --> 00:50:17,359

well

1288

00:50:21,190 --> 00:50:19,760

i have the perfect project for you

1289

00:50:24,630 --> 00:50:21,200

so

1290

00:50:26,870 --> 00:50:24,640

what she started with was to

1291

00:50:28,150 --> 00:50:26,880

create phylogeny of these rhodopsin

1292

00:50:30,630 --> 00:50:28,160

proteins

1293

00:50:32,230 --> 00:50:30,640

again we have this amazing amount of

1294

00:50:33,750 --> 00:50:32,240

information that is in literature this

1295

00:50:35,910 --> 00:50:33,760

vast amount of sequence data that

1296

00:50:38,470 --> 00:50:35,920

constantly keeps coming from this

1297

00:50:39,990 --> 00:50:38,480

wondrous metagenomic studies every year

1298

00:50:41,829 --> 00:50:40,000

it's really difficult to keep up with

1299

00:50:44,309 --> 00:50:41,839

which is very exciting

1300

00:50:45,349 --> 00:50:44,319

she updated a rhodopsin tree created it

1301

00:50:47,589 --> 00:50:45,359

here you're looking at different

1302

00:50:49,990 --> 00:50:47,599

topologies all these letters on the top

1303

00:50:52,950 --> 00:50:50,000

refer to a different evolutionary model

1304

00:50:55,109 --> 00:50:52,960

that we try to understand the best fits

1305

00:50:57,589 --> 00:50:55,119

of our topology so we go through this

1306

00:50:59,750 --> 00:50:57,599

iteration it takes us months we want to

1307

00:51:01,829 --> 00:50:59,760

understand by studying the sequence

1308

00:51:04,069 --> 00:51:01,839

again the ancestral functionality what

1309

00:51:06,230 --> 00:51:04,079

did this enzyme do

1310

00:51:08,630 --> 00:51:06,240

can we relate to any modern function to

1311

00:51:10,790 --> 00:51:08,640

the past we know that there is a proton

1312

00:51:13,109 --> 00:51:10,800

pump function or cholera uh pump

1313

00:51:15,190 --> 00:51:13,119

function for this enzyme today but what

1314

00:51:16,710 --> 00:51:15,200

did happen in the past so after going

1315

00:51:18,549 --> 00:51:16,720

through these inferences

1316

00:51:20,630 --> 00:51:18,559

we benefited from

1317

00:51:22,390 --> 00:51:20,640

these two studies that came from

1318

00:51:24,710 --> 00:51:22,400

university of tokyo

1319

00:51:26,790 --> 00:51:24,720

where they developed

1320

00:51:29,030 --> 00:51:26,800

back to back they published these papers

1321

00:51:33,190 --> 00:51:29,040

machine learning algorithms

1322

00:51:35,990 --> 00:51:33,200

to predict the valence spectra of the

1323

00:51:38,470 --> 00:51:36,000

rhodopsin proteins specifically and they

1324

00:51:40,630 --> 00:51:38,480

tested these the accuracies of these

1325

00:51:42,069 --> 00:51:40,640

training models in multiple ways not

1326

00:51:44,230 --> 00:51:42,079

only through

1327

00:51:46,309 --> 00:51:44,240

experiments that that they do in the

1328

00:51:48,309 --> 00:51:46,319

biochemical level but also they

1329

00:51:51,190 --> 00:51:48,319

generated multiple libraries and they

1330

00:51:52,630 --> 00:51:51,200

had a lot of experience in this studies

1331

00:51:55,910 --> 00:51:52,640

anyway and collaborated with marine

1332

00:51:57,510 --> 00:51:55,920

biologists to test the accuracy of the

1333

00:51:58,630 --> 00:51:57,520

predictions of their machine learning

1334

00:52:00,549 --> 00:51:58,640

algorithm

1335

00:52:02,549 --> 00:52:00,559

which was also great because due to

1336

00:52:05,670 --> 00:52:02,559

covet shutdown we had limited access to

1337

00:52:07,589 --> 00:52:05,680

lab anyway so really thankful for any

1338

00:52:08,870 --> 00:52:07,599

computational researcher who makes their

1339

00:52:11,349 --> 00:52:08,880

resources

1340

00:52:13,109 --> 00:52:11,359

available in open access and they were

1341

00:52:14,950 --> 00:52:13,119

able to answer all of our questions

1342

00:52:18,069 --> 00:52:14,960

throughout this process

1343

00:52:20,470 --> 00:52:18,079

doctoral student even fair she

1344

00:52:21,670 --> 00:52:20,480

took this method and and she wanted to

1345

00:52:23,670 --> 00:52:21,680

understand

1346

00:52:26,309 --> 00:52:23,680

what sort of different libraries we can

1347

00:52:28,630 --> 00:52:26,319

generate for our rhodopsins to implement

1348

00:52:30,549 --> 00:52:28,640

this machine learning methodology so we

1349

00:52:33,270 --> 00:52:30,559

weren't only interested in a handful of

1350

00:52:35,990 --> 00:52:33,280

ancestors what we can do computationally

1351  
00:52:37,430 --> 00:52:36,000  
that we can't do in the lab is to try

1352  
00:52:40,230 --> 00:52:37,440  
all kinds of variants we can do

1353  
00:52:41,030 --> 00:52:40,240  
thousands at once it's really remarkable

1354  
00:52:43,270 --> 00:52:41,040  
so

1355  
00:52:45,430 --> 00:52:43,280  
she created alternative sequences for

1356  
00:52:47,750 --> 00:52:45,440  
every ancestor that we generate we

1357  
00:52:50,630 --> 00:52:47,760  
assigned a threshold for what we define

1358  
00:52:52,630 --> 00:52:50,640  
as our confidence level and also the

1359  
00:52:54,950 --> 00:52:52,640  
threshold for the problems probabilistic

1360  
00:52:57,829 --> 00:52:54,960  
or most probabilistic ancestor that we

1361  
00:52:59,109 --> 00:52:57,839  
created and and we identified the sites

1362  
00:53:01,190 --> 00:52:59,119  
in the protein that are actually

1363  
00:53:03,030 --> 00:53:01,200

functionally important and generated

1364

00:53:05,270 --> 00:53:03,040

these libraries and

1365

00:53:07,430 --> 00:53:05,280

performed wavelength prediction

1366

00:53:09,270 --> 00:53:07,440

and through the suggestion of an

1367

00:53:11,349 --> 00:53:09,280

excellent reviewer we also tested the

1368

00:53:15,030 --> 00:53:11,359

robustness of the machine learning

1369

00:53:18,390 --> 00:53:15,040

algorithm to our uh to the training data

1370

00:53:23,190 --> 00:53:20,950

and for that we created the wavelength

1371

00:53:25,750 --> 00:53:23,200

prediction values for all these

1372

00:53:27,589 --> 00:53:25,760

rhodopsin proteins in our tree that we

1373

00:53:28,790 --> 00:53:27,599

created using our kill and bacterial

1374

00:53:30,790 --> 00:53:28,800

sequences

1375

00:53:33,510 --> 00:53:30,800

what we find is that the spectral

1376  
00:53:36,230 --> 00:53:33,520  
properties of these ancestral adoptions

1377  
00:53:39,270 --> 00:53:36,240  
indicate a green ship

1378  
00:53:42,950 --> 00:53:39,280  
as we go further back in time that that

1379  
00:53:46,390 --> 00:53:42,960  
the lambda max is in a range between

1380  
00:53:48,950 --> 00:53:46,400  
480 and 585

1381  
00:53:51,510 --> 00:53:48,960  
so what does this mean now that we have

1382  
00:53:53,829 --> 00:53:51,520  
an ancestor and we can detect its

1383  
00:53:56,710 --> 00:53:53,839  
wavelength spectra we went to our

1384  
00:53:58,309 --> 00:53:56,720  
friends at nexus and kudos to these

1385  
00:54:00,230 --> 00:53:58,319  
research coordination networks and this

1386  
00:54:03,510 --> 00:54:00,240  
is really why they are there so that you

1387  
00:54:07,109 --> 00:54:03,520  
have your 1 800 astronomer that you can

1388  
00:54:09,349 --> 00:54:07,119

go and talk to and and

1389

00:54:10,950 --> 00:54:09,359

perhaps collaborate with to make sense

1390

00:54:13,270 --> 00:54:10,960

of the biological data that you're

1391

00:54:15,430 --> 00:54:13,280

creating so eddie schwederman from uc

1392

00:54:17,750 --> 00:54:15,440

riverside came to the rescue having

1393

00:54:19,829 --> 00:54:17,760

worked on this with us sarma shilda

1394

00:54:22,549 --> 00:54:19,839

sarma for a long time

1395

00:54:25,109 --> 00:54:22,559

she he modeled the arcane

1396

00:54:27,589 --> 00:54:25,119

environment surface irradiance profile

1397

00:54:29,109 --> 00:54:27,599

independently and then we compared the

1398

00:54:31,190 --> 00:54:29,119

data that we generate with the

1399

00:54:32,390 --> 00:54:31,200

environmental the surface conditions of

1400

00:54:34,630 --> 00:54:32,400

early earth

1401

00:54:36,950 --> 00:54:34,640

what we think is happening is is that

1402

00:54:39,349 --> 00:54:36,960

the green wavelength absorption is

1403

00:54:42,150 --> 00:54:39,359

parsimonious with the spectral futures

1404

00:54:44,470 --> 00:54:42,160

of the early earth environment so we are

1405

00:54:46,549 --> 00:54:44,480

looking for a proton pumping protein

1406

00:54:48,630 --> 00:54:46,559

that may have residing in a microbe that

1407

00:54:52,150 --> 00:54:48,640

is swimming in the shallow waters

1408

00:54:53,270 --> 00:54:52,160

or perhaps was underneath some biofilm

1409

00:54:55,829 --> 00:54:53,280

and and

1410

00:54:57,430 --> 00:54:55,839

due to shallow coastal marine niches

1411

00:55:00,309 --> 00:54:57,440

which is maybe a great environment

1412

00:55:04,950 --> 00:55:00,319

abundant nutrients metabolically useful

1413

00:55:07,030 --> 00:55:04,960

light availability a tuned phototoxic

1414

00:55:09,190 --> 00:55:07,040

uv radiation so this may be a

1415

00:55:13,030 --> 00:55:09,200

nice little shelter for them

1416

00:55:15,670 --> 00:55:13,040

and we also just published this work

1417

00:55:18,309 --> 00:55:15,680

to summarize the histories of the

1418

00:55:21,190 --> 00:55:18,319

enzymes they can be amenable to

1419

00:55:24,309 --> 00:55:21,200

informatic and evolutionary construction

1420

00:55:25,109 --> 00:55:24,319

and such reconstructions in my opinion

1421

00:55:27,589 --> 00:55:25,119

are

1422

00:55:29,910 --> 00:55:27,599

opening entirely new possibilities for

1423

00:55:32,069 --> 00:55:29,920

the study of early life including the

1424

00:55:36,069 --> 00:55:32,079

fine resolution probing of geochemical

1425

00:55:38,230 --> 00:55:36,079

conditions and stellar properties

1426

00:55:40,470 --> 00:55:38,240

and let's go back to our major question

1427

00:55:42,789 --> 00:55:40,480

we want to know whether what we see here

1428

00:55:43,910 --> 00:55:42,799

is a particular case or a general

1429

00:55:45,990 --> 00:55:43,920

category

1430

00:55:47,510 --> 00:55:46,000

and as my friend and colleague sanjoy

1431

00:55:49,990 --> 00:55:47,520

likes to put it

1432

00:55:51,750 --> 00:55:50,000

exoplanets are one of the

1433

00:55:54,950 --> 00:55:51,760

early earth is one of the best exoplanet

1434

00:55:58,789 --> 00:55:57,270

but there's a lot of unknowns

1435

00:56:00,950 --> 00:55:58,799

and i know some of you are feeling this

1436

00:56:06,230 --> 00:56:00,960

bingo so i'll just help you out

1437

00:56:09,589 --> 00:56:08,870

and and study of early life

1438

00:56:11,510 --> 00:56:09,599

is

1439

00:56:13,349 --> 00:56:11,520

all hands on that problem we need

1440

00:56:15,109 --> 00:56:13,359

everybody in this endeavor because we

1441

00:56:17,270 --> 00:56:15,119

cannot find what we're looking for if we

1442

00:56:20,470 --> 00:56:17,280

don't understand it here and we owe it

1443

00:56:23,109 --> 00:56:20,480

to this life on this planet i think

1444

00:56:24,950 --> 00:56:23,119

and with that i will end uh first and

1445

00:56:27,349 --> 00:56:24,960

foremost my lab

1446

00:56:29,190 --> 00:56:27,359

i'm no longer by myself in this pursuit

1447

00:56:31,910 --> 00:56:29,200

it's amazing but

1448

00:56:33,829 --> 00:56:31,920

how many uh people are excited about

1449

00:56:35,670 --> 00:56:33,839

this in my group it's just one to bet i

1450

00:56:37,750 --> 00:56:35,680

created a lab that i wanted to be a part

1451

00:56:40,309 --> 00:56:37,760

of so this is just really makes me very

1452

00:56:42,150 --> 00:56:40,319

happy to work with them every day our

1453

00:56:44,470 --> 00:56:42,160

collaborators for different project

1454

00:56:46,309 --> 00:56:44,480

projects and funding

1455

00:56:49,589 --> 00:56:46,319

comes from nasa national science

1456

00:56:51,670 --> 00:56:49,599

foundation john templeton foundation nih

1457

00:56:53,910 --> 00:56:51,680

and also human frontiers and science

1458

00:56:55,910 --> 00:56:53,920

young investigator grants as well as

1459

00:56:57,510 --> 00:56:55,920

university of arizona and wisconsin

1460

00:57:00,390 --> 00:56:57,520

foundation

1461

00:57:01,750 --> 00:57:00,400

and also the spanish government that is

1462

00:57:04,470 --> 00:57:01,760

funding

1463

00:57:07,190 --> 00:57:04,480

one of our postdocs here and also our

1464

00:57:09,829 --> 00:57:07,200

cohort we just started our icard about

1465

00:57:11,910 --> 00:57:09,839

six months ago i'm already just

1466

00:57:13,430 --> 00:57:11,920

it's we've been doing a lot of cool work

1467

00:57:15,829 --> 00:57:13,440

i think we created a really great

1468

00:57:17,109 --> 00:57:15,839

community that is tackling this problem

1469

00:57:19,589 --> 00:57:17,119

of uh

1470

00:57:21,990 --> 00:57:19,599

understanding whether we can

1471

00:57:24,549 --> 00:57:22,000

integrate in organic thinking into

1472

00:57:27,510 --> 00:57:24,559

habitability from very different angles

1473

00:57:29,670 --> 00:57:27,520

i'm very grateful to work with them as

1474

00:57:31,829 --> 00:57:29,680

well and the names in yellow are

1475

00:57:34,309 --> 00:57:31,839

presenting here the asterix have talks

1476

00:57:36,069 --> 00:57:34,319

or posters

1477

00:57:38,150 --> 00:57:36,079

and finally we are launching a really

1478

00:57:39,990 --> 00:57:38,160

cool astrobiology program at university

1479

00:57:41,670 --> 00:57:40,000

of wisconsin-madison so i hope the

1480

00:57:44,390 --> 00:57:41,680

students interested in these problems

1481

00:57:59,190 --> 00:57:44,400

will check it out and join us

1482

00:58:03,829 --> 00:58:02,470

thank you very much bethel

1483

00:58:06,309 --> 00:58:03,839

um

1484

00:58:09,829 --> 00:58:06,319

we're close to the bottom of the hour oh

1485

00:58:11,190 --> 00:58:09,839

sorry but i would uh invite everyone to

1486

00:58:13,670 --> 00:58:11,200

stick around

1487

00:58:15,829 --> 00:58:13,680

who can do without coffee for a few

1488

00:58:18,230 --> 00:58:15,839

minutes so that we can enjoy

1489

00:58:20,069 --> 00:58:18,240

a question and answered period and i'm

1490

00:58:21,270 --> 00:58:20,079

sure there's going to be plenty of

1491

00:58:23,430 --> 00:58:21,280

questions

1492

00:58:26,549 --> 00:58:23,440

and i think we have somebody who can

1493

00:58:31,190 --> 00:58:28,390

first of all one of our guests last

1494

00:58:43,349 --> 00:58:31,200

night here is gideon ask the question

1495

00:58:47,270 --> 00:58:44,870

first off thanks

1496

00:58:48,630 --> 00:58:47,280

an amazing talk

1497

00:58:50,309 --> 00:58:48,640

so happy to have you

1498

00:58:54,309 --> 00:58:50,319

wesley swingley from northern illinois

1499

00:58:56,549 --> 00:58:54,319

university kind of a topical question um

1500

00:58:58,150 --> 00:58:56,559

sorry it's cutting it a philosophical

1501  
00:58:59,670 --> 00:58:58,160  
question you mentioned that these things

1502  
00:59:02,950 --> 00:58:59,680  
aren't

1503  
00:59:04,950 --> 00:59:02,960  
deterministic or inevitable

1504  
00:59:06,789 --> 00:59:04,960  
what is the what do you think that is

1505  
00:59:09,109 --> 00:59:06,799  
between these singular

1506  
00:59:11,190 --> 00:59:09,119  
systems like nitrogen or oxygenic

1507  
00:59:13,030 --> 00:59:11,200  
photosynthesis versus things that are

1508  
00:59:14,870 --> 00:59:13,040  
big buckets of stuff like carbon

1509  
00:59:16,630 --> 00:59:14,880  
fixation or pigments and things like

1510  
00:59:19,349 --> 00:59:16,640  
that

1511  
00:59:20,950 --> 00:59:19,359  
in terms of the like the molecular level

1512  
00:59:23,510 --> 00:59:20,960  
just likewise

1513  
00:59:25,910 --> 00:59:23,520

why do we have one nitrogenase and

1514

00:59:27,990 --> 00:59:25,920

10 different carbon fixation pathways

1515

00:59:30,549 --> 00:59:28,000

that that's a very good question i think

1516

00:59:33,510 --> 00:59:30,559

maybe what determines the adaptation

1517

00:59:35,430 --> 00:59:33,520

isn't just the enzyme itself and i think

1518

00:59:37,670 --> 00:59:35,440

it goes back to

1519

00:59:39,829 --> 00:59:37,680

our lack of understanding

1520

00:59:42,069 --> 00:59:39,839

of what how earth the environment

1521

00:59:44,630 --> 00:59:42,079

impacted these molecular innovations so

1522

00:59:47,510 --> 00:59:44,640

i don't quite have a good answer

1523

00:59:49,670 --> 00:59:47,520

when it comes to evolutionary biology

1524

00:59:51,109 --> 00:59:49,680

because that's not even i think a

1525

00:59:53,589 --> 00:59:51,119

question that many evolutionary

1526

00:59:56,309 --> 00:59:53,599

biologists are tackling right now but

1527

00:59:58,630 --> 00:59:56,319

it's the uh so we have a big question

1528

01:00:00,789 --> 00:59:58,640

mark as to why there are certain

1529

01:00:03,109 --> 01:00:00,799

constraints

1530

01:00:08,549 --> 01:00:03,119

thanks

1531

01:00:12,030 --> 01:00:09,910

so

1532

01:00:15,109 --> 01:00:12,040

can you apply this technique to

1533

01:00:17,910 --> 01:00:15,119

chemosynthetic metabolisms and estimate

1534

01:00:20,390 --> 01:00:17,920

the likelihood in time for the origin of

1535

01:00:23,109 --> 01:00:20,400

them maybe methanogenesis

1536

01:00:25,589 --> 01:00:23,119

or iron reducing bacteria can we

1537

01:00:27,589 --> 01:00:25,599

translate this information to estimation

1538

01:00:30,150 --> 01:00:27,599

of timing of origin

1539

01:00:32,069 --> 01:00:30,160

for similar possible mechanisms on other

1540

01:00:34,710 --> 01:00:32,079

worlds

1541

01:00:35,670 --> 01:00:34,720

yeah i think if we if we have to be very

1542

01:00:39,750 --> 01:00:35,680

careful

1543

01:00:42,390 --> 01:00:39,760

um i think in when we set up the system

1544

01:00:43,270 --> 01:00:42,400

iv outlined the first criteria

1545

01:00:46,789 --> 01:00:43,280

and

1546

01:00:48,950 --> 01:00:46,799

i published that in 2017 i'm writing an

1547

01:00:51,190 --> 01:00:48,960

updated one as to what sort of criteria

1548

01:00:53,589 --> 01:00:51,200

i think is necessary for reconstruction

1549

01:00:56,630 --> 01:00:53,599

of ancient metabolisms when it comes to

1550

01:00:58,710 --> 01:00:56,640

the timing of origin i think we are now

1551

01:01:00,390 --> 01:00:58,720

able to not only reconstruct a single

1552

01:01:02,710 --> 01:01:00,400

gene but we can engineer entire

1553

01:01:04,870 --> 01:01:02,720

communities so uh there's nothing that

1554

01:01:06,789 --> 01:01:04,880

stops us from engineering a community

1555

01:01:08,870 --> 01:01:06,799

and and maybe that goes back to the

1556

01:01:10,710 --> 01:01:08,880

business question that if we understand

1557

01:01:12,230 --> 01:01:10,720

how these communities are interacting we

1558

01:01:13,750 --> 01:01:12,240

might be able to

1559

01:01:16,470 --> 01:01:13,760

understand whether

1560

01:01:18,230 --> 01:01:16,480

there is factors of determinism

1561

01:01:20,549 --> 01:01:18,240

so so yes i think it's definitely a

1562

01:01:22,710 --> 01:01:20,559

possibility but just be mindful is how

1563

01:01:25,510 --> 01:01:22,720

you pick your system

1564

01:01:27,349 --> 01:01:25,520

and be very careful with the design

1565

01:01:29,190 --> 01:01:27,359

happy to help thank you and those those

1566

01:01:32,710 --> 01:01:29,200

questions were from sean domicile goldman

1567

01:01:34,069 --> 01:01:32,720

and kendall lynch oh hi sean

1568

01:01:36,309 --> 01:01:34,079

next question

1569

01:01:39,190 --> 01:01:36,319

hypothetical beautiful talk um steven

1570

01:01:42,150 --> 01:01:39,200

freed from johns hopkins um this is a

1571

01:01:44,549 --> 01:01:42,160

question coming from a fellow biochemist

1572

01:01:46,710 --> 01:01:44,559

i was curious about so nitrogenases use

1573

01:01:48,630 --> 01:01:46,720

lots of chaperones to get their

1574

01:01:51,109 --> 01:01:48,640

biogenesis so i was curious if you were

1575

01:01:53,670 --> 01:01:51,119

surprised that the

1576

01:01:55,910 --> 01:01:53,680

nitrogenase was able to use the extant

1577

01:01:59,029 --> 01:01:55,920

as zodobacterous chaperones and did that

1578

01:02:00,789 --> 01:01:59,039

map with where the mutations were in the

1579

01:02:02,069 --> 01:02:00,799

ancestral node like

1580

01:02:04,069 --> 01:02:02,079

do you know where on the surface they

1581

01:02:07,109 --> 01:02:04,079

were and did that surprise you or not

1582

01:02:09,029 --> 01:02:07,119

surprise you yeah so uh well

1583

01:02:10,870 --> 01:02:09,039

so i i'm in this point where everything

1584

01:02:12,470 --> 01:02:10,880

is surprising me and nothing is

1585

01:02:14,230 --> 01:02:12,480

surprising me at the same time it's a

1586

01:02:17,029 --> 01:02:14,240

very quantum state

1587

01:02:19,829 --> 01:02:17,039

of existence but uh it's we have not

1588

01:02:21,430 --> 01:02:19,839

done an in-depth network analysis uh i

1589

01:02:23,029 --> 01:02:21,440

think our postdoc bruno here he's

1590

01:02:24,230 --> 01:02:23,039

interested in that exact question we

1591

01:02:26,390 --> 01:02:24,240

actually just talked about that in one

1592

01:02:28,230 --> 01:02:26,400

of our group meetings uh so yes we're

1593

01:02:30,549 --> 01:02:28,240

definitely interested in understanding

1594

01:02:33,510 --> 01:02:30,559

at the more uh larger level systems

1595

01:02:34,950 --> 01:02:33,520

level what what might be happening uh i

1596

01:02:38,309 --> 01:02:34,960

think it's fascinating that we're even

1597

01:02:39,670 --> 01:02:38,319

able to swap these highly constrained

1598

01:02:43,990 --> 01:02:39,680

systems

1599

01:02:49,910 --> 01:02:45,190

i had a

1600

01:02:53,430 --> 01:02:49,920

regarding the dobson you indicated that

1601  
01:02:56,390 --> 01:02:53,440  
the change in uh frequency optimization

1602  
01:02:59,430 --> 01:02:56,400  
of the rhodopsin suggested

1603  
01:03:00,950 --> 01:02:59,440  
of that that this particular adoption

1604  
01:03:03,349 --> 01:03:00,960  
involved in

1605  
01:03:06,150 --> 01:03:03,359  
could it also indicate a change in the

1606  
01:03:07,109 --> 01:03:06,160  
peak output of the sun a billion years

1607  
01:03:08,870 --> 01:03:07,119  
ago and

1608  
01:03:10,470 --> 01:03:08,880  
so it's actually reflecting physical

1609  
01:03:12,309 --> 01:03:10,480  
something happening at the sun by the

1610  
01:03:14,230 --> 01:03:12,319  
surface of the earth

1611  
01:03:16,870 --> 01:03:14,240  
uh yeah so the what

1612  
01:03:19,349 --> 01:03:16,880  
eddie has done was independent than from

1613  
01:03:21,990 --> 01:03:19,359

our calculations so he did look at the

1614

01:03:24,230 --> 01:03:22,000

irradiance profile of the sun itself

1615

01:03:27,589 --> 01:03:24,240

um independent of uh

1616

01:03:29,910 --> 01:03:27,599

before we integrated and interpreted the

1617

01:03:32,390 --> 01:03:29,920

phenotype that we've predicted that

1618

01:03:34,870 --> 01:03:32,400

belongs to these ancestral proteins and

1619

01:03:36,309 --> 01:03:34,880

then we sort of mapped them as to this

1620

01:03:39,349 --> 01:03:36,319

is the environment

1621

01:03:41,990 --> 01:03:39,359

uh and it matched uh it in the water

1622

01:03:43,990 --> 01:03:42,000

column where the the host organism might

1623

01:03:45,430 --> 01:03:44,000

be residing but we haven't done any

1624

01:03:47,670 --> 01:03:45,440

direct uh

1625

01:03:50,150 --> 01:03:47,680

tests that focuses on the sun or the

1626

01:03:51,910 --> 01:03:50,160

star itself although

1627

01:03:53,510 --> 01:03:51,920

i think eddie is definitely interested

1628

01:04:01,349 --> 01:03:53,520

in

1629

01:04:04,710 --> 01:04:01,359

was stronger the local environment or

1630

01:04:07,349 --> 01:04:04,720

the or the solar radiance profile

1631

01:04:09,670 --> 01:04:07,359

that's so i'm i am very cautious when it

1632

01:04:11,670 --> 01:04:09,680

comes to interpreting the early

1633

01:04:14,230 --> 01:04:11,680

environments and and what sort of

1634

01:04:16,789 --> 01:04:14,240

possibilities there might be if we sort

1635

01:04:19,109 --> 01:04:16,799

of really conservatively propose that

1636

01:04:19,990 --> 01:04:19,119

perhaps they were either deeper that

1637

01:04:21,750 --> 01:04:20,000

they were

1638

01:04:25,750 --> 01:04:21,760

where they could be located in the water

1639

01:04:28,230 --> 01:04:25,760

column or maybe there was a biofilm um

1640

01:04:30,230 --> 01:04:28,240

that was covering these organisms but i

1641

01:04:31,029 --> 01:04:30,240

i want to be very careful since we don't

1642

01:04:33,349 --> 01:04:31,039

know

1643

01:04:34,789 --> 01:04:33,359

um anything really about those

1644

01:04:36,470 --> 01:04:34,799

environments so

1645

01:04:38,309 --> 01:04:36,480

i won't speculate

1646

01:04:40,549 --> 01:04:38,319

therefore it may be more likely i think

1647

01:04:41,270 --> 01:04:40,559

there's more information about the from

1648

01:04:43,349 --> 01:04:41,280

the

1649

01:04:45,270 --> 01:04:43,359

astronomy and anyway when it comes to

1650

01:04:47,190 --> 01:04:45,280

understanding these early models

1651

01:04:49,990 --> 01:04:47,200

thank you

1652

01:04:51,670 --> 01:04:50,000

thanks next question please um yeah your

1653

01:04:53,670 --> 01:04:51,680

story about nitrogenase was quite

1654

01:04:54,950 --> 01:04:53,680

inspiring as an outsider to the field i

1655

01:04:56,150 --> 01:04:54,960

was really impressed with what you could

1656

01:04:57,910 --> 01:04:56,160

do with it now i'm going to ask you

1657

01:05:00,789 --> 01:04:57,920

something really unfair

1658

01:05:02,390 --> 01:05:00,799

what about rubisco how far back do you

1659

01:05:04,309 --> 01:05:02,400

think that goes

1660

01:05:06,630 --> 01:05:04,319

and why

1661

01:05:09,190 --> 01:05:06,640

so uh rubisco is this a very fair

1662

01:05:10,870 --> 01:05:09,200

question first of all so rubisco is a

1663

01:05:12,230 --> 01:05:10,880

bit more obviously a different type of

1664

01:05:15,270 --> 01:05:12,240

beast

1665

01:05:17,430 --> 01:05:15,280

if you attribute the presence of early

1666

01:05:18,710 --> 01:05:17,440

carbon fixation

1667

01:05:20,549 --> 01:05:18,720

records to

1668

01:05:23,109 --> 01:05:20,559

rubisco or something like rubisco then

1669

01:05:23,910 --> 01:05:23,119

you can stretch its presence as early as

1670

01:05:24,710 --> 01:05:23,920

the

1671

01:05:27,029 --> 01:05:24,720

first

1672

01:05:30,309 --> 01:05:27,039

reliable data of the carbon isotope

1673

01:05:32,150 --> 01:05:30,319

biosignature right but

1674

01:05:34,150 --> 01:05:32,160

the simple answer to your question is

1675

01:05:35,910 --> 01:05:34,160

that nobody knows how old rubisco is and

1676

01:05:38,069 --> 01:05:35,920

we make this assumption that something

1677

01:05:41,510 --> 01:05:38,079

like it or itself was there

1678

01:05:43,829 --> 01:05:41,520

uh when the carbon fixation um

1679

01:05:45,109 --> 01:05:43,839

the biological fixation was

1680

01:05:47,910 --> 01:05:45,119

emerged

1681

01:05:50,710 --> 01:05:47,920

what we've done uh is to after we

1682

01:05:52,870 --> 01:05:50,720

created the phylogeny we included all

1683

01:05:54,710 --> 01:05:52,880

forms of rubisco from today

1684

01:05:56,950 --> 01:05:54,720

and we simply wanted to know whether we

1685

01:05:59,990 --> 01:05:56,960

can use the sequence information of this

1686

01:06:01,349 --> 01:06:00,000

enzyme to understand first whether we

1687

01:06:03,349 --> 01:06:01,359

can

1688

01:06:06,950 --> 01:06:03,359

find signatures of

1689

01:06:08,710 --> 01:06:06,960

rise of oxygen on this enzyme

1690

01:06:10,789 --> 01:06:08,720

everything that i know about molecular

1691

01:06:11,750 --> 01:06:10,799

evolution tells me that this enzyme

1692

01:06:13,430 --> 01:06:11,760

really

1693

01:06:14,710 --> 01:06:13,440

there will be variations

1694

01:06:16,950 --> 01:06:14,720

right but the assumption is that there's

1695

01:06:19,029 --> 01:06:16,960

some sort of uniformity and and we

1696

01:06:22,230 --> 01:06:19,039

looked at where the enzyme is binding to

1697

01:06:25,109 --> 01:06:22,240

oxygen to carbon dioxide uh and in fact

1698

01:06:26,950 --> 01:06:25,119

found that some parts of the enzyme was

1699

01:06:29,270 --> 01:06:26,960

evolving more rapidly than the others

1700

01:06:32,950 --> 01:06:29,280

that corresponded to the oxygen carbon

1701

01:06:34,309 --> 01:06:32,960

dioxide binding that also coincidentally

1702

01:06:35,910 --> 01:06:34,319

belonged to the ancestors of

1703

01:06:38,549 --> 01:06:35,920

cyanobacteria

1704

01:06:41,270 --> 01:06:38,559

so we published this in geobiology

1705

01:06:43,510 --> 01:06:41,280

saying we can constrain goe

1706

01:06:44,549 --> 01:06:43,520

using the rubisco sequence

1707

01:06:47,029 --> 01:06:44,559

perhaps

1708

01:06:48,549 --> 01:06:47,039

by analyzing the sequences itself we are

1709

01:06:50,150 --> 01:06:48,559

not timing anything but we're just

1710

01:06:51,510 --> 01:06:50,160

saying that there is a signature on this

1711

01:06:53,510 --> 01:06:51,520

enzyme that is telling us something

1712

01:06:55,430 --> 01:06:53,520

about what happened on earth

1713

01:06:59,270 --> 01:06:55,440

after we engineered this inside

1714

01:07:01,990 --> 01:06:59,280

cyanobacteria we did the isotope

1715

01:07:04,150 --> 01:07:02,000

studies and we we didn't go as far we

1716

01:07:05,109 --> 01:07:04,160

didn't go pre-goe yet that's in the

1717

01:07:06,150 --> 01:07:05,119

works

1718

01:07:08,069 --> 01:07:06,160

we think we have a bit more

1719

01:07:09,829 --> 01:07:08,079

understanding the two about under

1720

01:07:11,670 --> 01:07:09,839

evolution of this enzyme before we go to

1721

01:07:13,270 --> 01:07:11,680

that engineering but at the first

1722

01:07:15,190 --> 01:07:13,280

attempt our

1723

01:07:17,430 --> 01:07:15,200

biosignature reconstruction

1724

01:07:19,349 --> 01:07:17,440

was in agreement with what the

1725

01:07:21,190 --> 01:07:19,359

paleobiologists have been thinking about

1726

01:07:22,630 --> 01:07:21,200

or measuring

1727

01:07:24,309 --> 01:07:22,640

fantastic i'm going to have to go to the

1728

01:07:26,470 --> 01:07:24,319

library but um i have a follow-up

1729

01:07:29,029 --> 01:07:26,480

question carbonic anhydrase is similar

1730

01:07:30,789 --> 01:07:29,039

to rubisco and its fractionation ability

1731

01:07:32,950 --> 01:07:30,799

in your phylogeny did you find any

1732

01:07:34,390 --> 01:07:32,960

relationship between carbonic anhydrase

1733

01:07:35,990 --> 01:07:34,400

and rubisco yes so i actually

1734

01:07:37,990 --> 01:07:36,000

reconstructed the

1735

01:07:39,750 --> 01:07:38,000

history of carbonic anhydrase as well

1736

01:07:41,589 --> 01:07:39,760

and in the field transit paper we

1737

01:07:43,349 --> 01:07:41,599

published it there

1738

01:07:45,750 --> 01:07:43,359

actually funny you asked that so

1739

01:07:48,789 --> 01:07:45,760

carbonic anhydrase is uh sort of the

1740

01:07:50,789 --> 01:07:48,799

worst example so if anyone's wondering

1741

01:07:52,710 --> 01:07:50,799

that's how not to do it because it's a

1742

01:07:55,270 --> 01:07:52,720

very uh as you know we even have it in

1743

01:07:56,710 --> 01:07:55,280

our blood right so this enzyme is

1744

01:07:59,990 --> 01:07:56,720

everywhere

1745

01:08:01,430 --> 01:08:00,000

and uh in some form and uh it is very

1746

01:08:03,109 --> 01:08:01,440

hard to track it because of the high

1747

01:08:05,270 --> 01:08:03,119

horizontal gene transfer rate of this

1748

01:08:06,950 --> 01:08:05,280

enzyme so we reconstructed it and i

1749

01:08:09,109 --> 01:08:06,960

thought actually you know what we really

1750

01:08:10,549 --> 01:08:09,119

should not use this as a probe because

1751

01:08:12,710 --> 01:08:10,559

it is all over the place and i cannot

1752

01:08:14,230 --> 01:08:12,720

constrain anything so i included that as

1753

01:08:16,309 --> 01:08:14,240

an example of what

1754

01:08:18,630 --> 01:08:16,319

kind of protein does not fit the build

1755

01:08:21,430 --> 01:08:18,640

for deep time reconstruction

1756

01:08:23,669 --> 01:08:21,440

fantastic i'm off to the library

1757

01:08:26,070 --> 01:08:23,679

thank you very much i believe that we're

1758

01:08:28,229 --> 01:08:26,080

out of time now we've been out of time

1759

01:08:30,390 --> 01:08:28,239

but uh please uh

1760

01:08:33,590 --> 01:08:30,400

discuss your your questions with batul

1761

01:08:39,669 --> 01:08:36,390

and with that i will thank our speaker