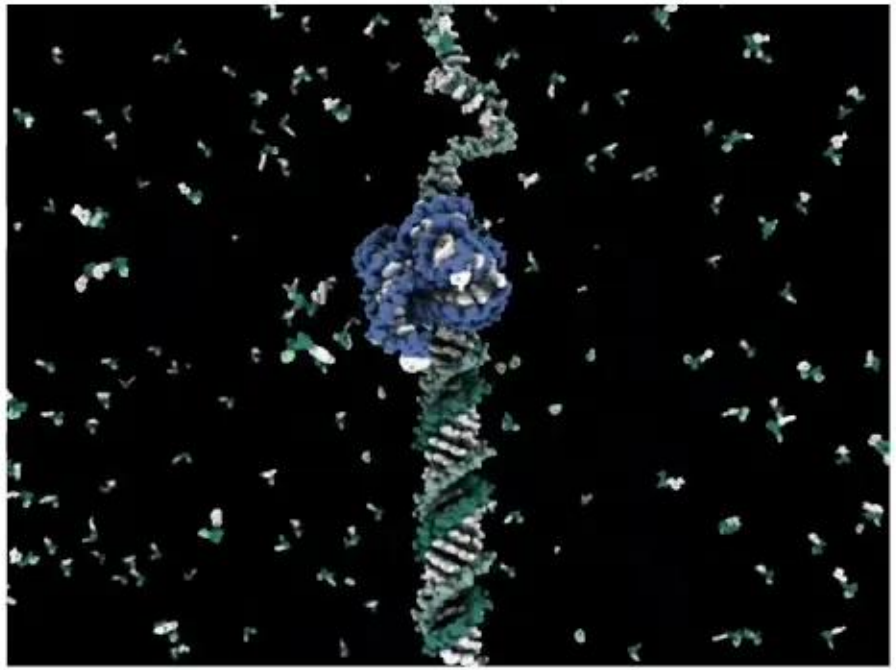
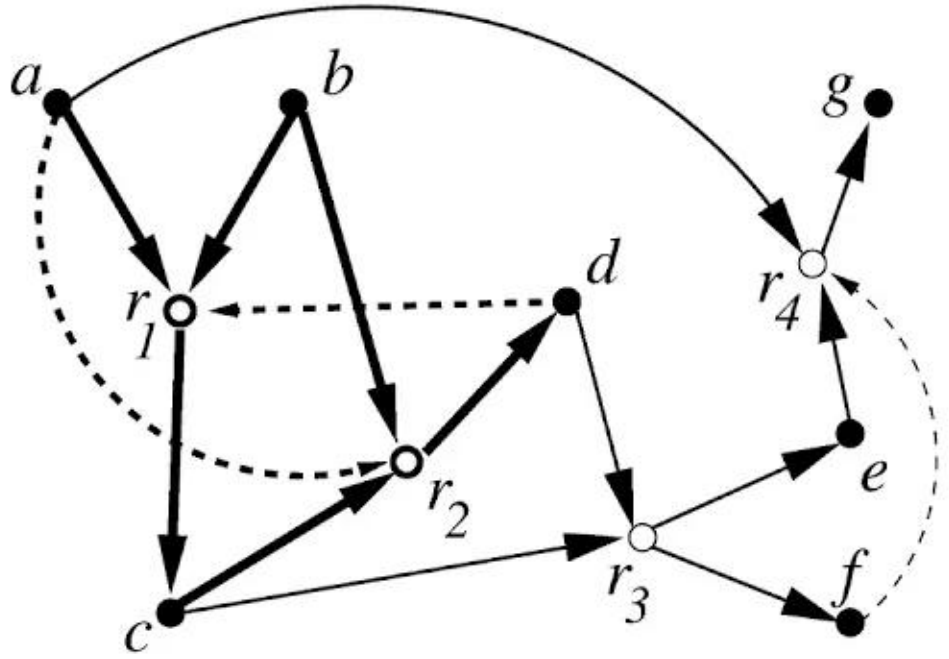
The first replicators were not cellular

A genetic replicator



Janet Iwasa, exploringorigins.org

An autocatalytic network



Hordjik *et al.*(2010) Entropy 12:1733-1742

1
00:00:05,190 --> 00:00:03,189
hello my name is yuta a takagi

2
00:00:06,950 --> 00:00:05,200
and this presentation will be on using

3
00:00:09,190 --> 00:00:06,960
digital life simulations

4
00:00:11,749 --> 00:00:09,200
to examine the evolutionary origin of

5
00:00:13,910 --> 00:00:11,759
cellularity and metabolism

6
00:00:16,550 --> 00:00:13,920
in this research we designed and coded a

7
00:00:18,390 --> 00:00:16,560
digital life system to study evolution

8
00:00:21,590 --> 00:00:18,400
and used the simulation to study the

9
00:00:24,150 --> 00:00:21,600
origin of cellularity

10
00:00:26,310 --> 00:00:24,160
first a bit of background we know that

11
00:00:27,670 --> 00:00:26,320
the last universal common ancestor was

12
00:00:29,589 --> 00:00:27,680
already cellular

13
00:00:31,349 --> 00:00:29,599

because cell membrane proteins such as

14
00:00:33,910 --> 00:00:31,359
signal receptors and the

15
00:00:37,510 --> 00:00:33,920
atp synthase motor are universally

16
00:00:39,430 --> 00:00:37,520
present across the known tree of life

17
00:00:41,990 --> 00:00:39,440
however almost all origin of life

18
00:00:42,869 --> 00:00:42,000
hypotheses suggest a non-cellular

19
00:00:45,110 --> 00:00:42,879
beginning

20
00:00:47,029 --> 00:00:45,120
such as the rna world hypothesis where

21
00:00:47,990 --> 00:00:47,039
the first entities were self-replicating

22
00:00:49,910 --> 00:00:48,000
rnas

23
00:00:51,029 --> 00:00:49,920
or in the auto catalytic network

24
00:00:52,869 --> 00:00:51,039
hypothesis

25
00:00:55,670 --> 00:00:52,879
which more closely resembles complex

26
00:00:58,869 --> 00:00:57,510
this necessarily means that at some

27
00:01:00,869 --> 00:00:58,879
point in early life

28
00:01:02,950 --> 00:01:00,879
there was a transition from non-cellular

29
00:01:05,750 --> 00:01:02,960
to cellular entities

30
00:01:07,109 --> 00:01:05,760
our work was motivated by the fact that

31
00:01:09,350 --> 00:01:07,119
even though there has been a lot of

32
00:01:10,789 --> 00:01:09,360
previous research on the biochemistry of

33
00:01:12,310 --> 00:01:10,799
early cellularity

34
00:01:14,550 --> 00:01:12,320
there has been very little research

35
00:01:17,830 --> 00:01:14,560
examining the evolutionary causes for

36
00:01:22,230 --> 00:01:20,469
i will now give a simplified explanation

37
00:01:24,390 --> 00:01:22,240
of the simulation

38
00:01:26,230 --> 00:01:24,400

the simulation consists of a population

39
00:01:28,469 --> 00:01:26,240
of replicators referred to as

40
00:01:30,149 --> 00:01:28,479
organisms in a shared environment

41
00:01:32,310 --> 00:01:30,159
containing food

42
00:01:34,789 --> 00:01:32,320
the environment is capped at a certain

43
00:01:37,350 --> 00:01:34,799
population size

44
00:01:40,870 --> 00:01:37,360
each organism has a genome composed of

45
00:01:43,510 --> 00:01:40,880
metabolic and cellularity genes

46
00:01:44,469 --> 00:01:43,520
metabolic genes represent logical not

47
00:01:47,270 --> 00:01:44,479
and gates

48
00:01:49,030 --> 00:01:47,280
that denote a boolean logic gate network

49
00:01:51,270 --> 00:01:49,040
the food in the simulations

50
00:01:52,630 --> 00:01:51,280
are boolean logic puzzles that give

51
00:01:54,550 --> 00:01:52,640
energy as a reward

52
00:01:56,789 --> 00:01:54,560
depending on how well they are solved by

53
00:01:58,870 --> 00:01:56,799
the genome's logic gate network

54
00:02:02,630 --> 00:01:58,880
a property we have termed metabolic

55
00:02:06,950 --> 00:02:04,870
energy is stored in the organism and is

56
00:02:09,270 --> 00:02:06,960
used to read through the genome

57
00:02:11,350 --> 00:02:09,280
where each step corresponds to one gene

58
00:02:13,270 --> 00:02:11,360
denoting a calculation in the boolean

59
00:02:15,110 --> 00:02:13,280
logic gate network

60
00:02:16,869 --> 00:02:15,120
the genome can read values from the

61
00:02:18,630 --> 00:02:16,879
puzzle's input register

62
00:02:20,070 --> 00:02:18,640
process the values through its logic

63
00:02:21,910 --> 00:02:20,080

gate network

64
00:02:24,070 --> 00:02:21,920
and rate the solutions to an output

65
00:02:25,830 --> 00:02:24,080
register which is then compared to the

66
00:02:27,830 --> 00:02:25,840
puzzle's solution register

67
00:02:30,710 --> 00:02:27,840
to determine the energy reward based on

68
00:02:32,869 --> 00:02:30,720
the number of correctly solved digits

69
00:02:35,750 --> 00:02:32,879
if stored energy runs out the organism

70
00:02:39,430 --> 00:02:37,910
cellularity genes are logically neutral

71
00:02:41,270 --> 00:02:39,440
in the boolean gate network

72
00:02:44,070 --> 00:02:41,280
but instead determine an organism's

73
00:02:46,309 --> 00:02:44,080
level of cellularity

74
00:02:47,190 --> 00:02:46,319
cellularity is defined as a value

75
00:02:49,270 --> 00:02:47,200
between zero

76
00:02:50,309 --> 00:02:49,280
and one with a zero representing

77
00:02:52,869 --> 00:02:50,319
non-cellular

78
00:02:56,070 --> 00:02:52,879
and one representing completely cellular

79
00:02:58,309 --> 00:02:56,080
in other words entirely closed off

80
00:03:00,229 --> 00:02:58,319
cellularity genes confer cellularity

81
00:03:01,430 --> 00:03:00,239
according to the equation one minus

82
00:03:03,350 --> 00:03:01,440
one-half to the x

83
00:03:04,949 --> 00:03:03,360
where x is the number of cellularity

84
00:03:06,949 --> 00:03:04,959
genes in the genome

85
00:03:10,149 --> 00:03:06,959
asymptotically approaching one with

86
00:03:12,550 --> 00:03:10,159
higher numbers of cellularity genes

87
00:03:14,470 --> 00:03:12,560
organisms with low levels of cellularity

88
00:03:16,390 --> 00:03:14,480

are more open to the environment

89
00:03:17,589 --> 00:03:16,400
and are subject to losing or gaining

90
00:03:19,670 --> 00:03:17,599
genetic material

91
00:03:21,990 --> 00:03:19,680
or free energy from the environment

92
00:03:25,430 --> 00:03:22,000
while organisms with high cellularity

93
00:03:29,589 --> 00:03:27,670
the organisms replicate their genomes

94
00:03:30,390 --> 00:03:29,599
and split into two new identical

95
00:03:32,550 --> 00:03:30,400
entities

96
00:03:33,670 --> 00:03:32,560
when a certain level of excess energy is

97
00:03:35,270 --> 00:03:33,680
reached

98
00:03:37,110 --> 00:03:35,280
with each new daughter organism

99
00:03:40,630 --> 00:03:37,120
receiving half of their parents energy

100
00:03:44,229 --> 00:03:43,030
organisms are constantly subject to low

101
00:03:46,229 --> 00:03:44,239
levels of mutation

102
00:03:48,710 --> 00:03:46,239
in the form of small alterations to the

103
00:03:50,949 --> 00:03:48,720
genome through insertion or deletion of

104
00:03:52,949 --> 00:03:50,959
metabolic or cellularity genes

105
00:03:56,869 --> 00:03:52,959
or through small reconfigurations of the

106
00:04:00,630 --> 00:03:59,190
organisms are also subject to loss or

107
00:04:02,630 --> 00:04:00,640
gain of potentially large

108
00:04:04,390 --> 00:04:02,640
sections of genetic material through

109
00:04:07,670 --> 00:04:04,400
horizontal gene transfer

110
00:04:09,830 --> 00:04:07,680
depending on their level of cellularity

111
00:04:11,750 --> 00:04:09,840
this form of mutation can cause drastic

112
00:04:13,589 --> 00:04:11,760
changes to the metabolic network

113
00:04:15,589 --> 00:04:13,599

and we have found that this is almost

114
00:04:20,229 --> 00:04:15,599
always deleterious to the metabolic

115
00:04:23,670 --> 00:04:22,230
the simulation is quite far abstracted

116
00:04:25,830 --> 00:04:23,680
from biological life

117
00:04:27,270 --> 00:04:25,840
but it does represent a system capable

118
00:04:28,950 --> 00:04:27,280
of real evolution

119
00:04:32,150 --> 00:04:28,960
where cellularity and metabolic

120
00:04:34,230 --> 00:04:32,160
proficiency are dependent variables

121
00:04:37,350 --> 00:04:34,240
allowing us to empirically address broad

122
00:04:39,110 --> 00:04:37,360
evolutionary questions

123
00:04:41,350 --> 00:04:39,120
after development of the software was

124
00:04:43,350 --> 00:04:41,360
completed we needed to narrow down the

125
00:04:45,350 --> 00:04:43,360
parameter space to conditions that

126
00:04:46,629 --> 00:04:45,360
allowed for organisms to survive and

127
00:04:48,310 --> 00:04:46,639
evolve

128
00:04:50,790 --> 00:04:48,320
during this stage we noticed that the

129
00:04:53,990 --> 00:04:50,800
amount of free energy in the environment

130
00:04:56,070 --> 00:04:54,000
had an effect on cellularity

131
00:04:57,909 --> 00:04:56,080
this observation led to our primary

132
00:05:00,390 --> 00:04:57,919
hypothesis

133
00:05:01,510 --> 00:05:00,400
that non-cellularity will be selectively

134
00:05:03,749 --> 00:05:01,520
advantageous

135
00:05:05,749 --> 00:05:03,759
when simple sources of energy are freely

136
00:05:08,310 --> 00:05:05,759
available in the environment

137
00:05:10,469 --> 00:05:08,320
and conversely that high cellularity

138
00:05:12,629 --> 00:05:10,479

will be selectively advantageous

139
00:05:17,270 --> 00:05:12,639
when only complex food sources are

140
00:05:21,670 --> 00:05:19,909
in order to test these hypotheses we ran

141
00:05:23,110 --> 00:05:21,680
a series of selection experiments

142
00:05:25,270 --> 00:05:23,120
altering the organism's

143
00:05:27,909 --> 00:05:25,280
starting cellularity and the amount of

144
00:05:30,469 --> 00:05:27,919
free energy in the environment

145
00:05:32,390 --> 00:05:30,479
simulation sets a and b began with zero

146
00:05:34,790 --> 00:05:32,400
cellularity genes giving initial

147
00:05:37,110 --> 00:05:34,800
cellularity values of zero percent

148
00:05:39,990 --> 00:05:37,120
while such c and d began with three

149
00:05:41,430 --> 00:05:40,000
cellularity genes giving 87.5 percent

150
00:05:43,749 --> 00:05:41,440
cellularity

151
00:05:44,710 --> 00:05:43,759
simulation such a and c were provided

152
00:05:47,510 --> 00:05:44,720
environments with

153
00:05:48,310 --> 00:05:47,520
unlimited food and unlimited energy

154
00:05:51,350 --> 00:05:48,320
while sets b

155
00:05:52,870 --> 00:05:51,360
and d had unlimited food but limited

156
00:05:54,710 --> 00:05:52,880
energy

157
00:05:56,070 --> 00:05:54,720
each condition was run with three

158
00:05:58,230 --> 00:05:56,080
replicas each

159
00:05:59,749 --> 00:05:58,240
under three maximum population settings

160
00:06:02,950 --> 00:05:59,759
of five hundred one thousand

161
00:06:07,270 --> 00:06:04,629
we found that regardless of the starting

162
00:06:09,510 --> 00:06:07,280
cellularity organisms consistently

163
00:06:11,270 --> 00:06:09,520

evolved towards low cellularity

164
00:06:13,749 --> 00:06:11,280
under high environmental energy

165
00:06:15,990 --> 00:06:13,759
conditions and high cellularity under

166
00:06:19,350 --> 00:06:16,000
low environmental energy conditions

167
00:06:22,950 --> 00:06:21,350
we further saw a corresponding trend

168
00:06:24,230 --> 00:06:22,960
towards low and high metabolic

169
00:06:26,550 --> 00:06:24,240
proficiency

170
00:06:28,629 --> 00:06:26,560
with poor solutions evolving under high

171
00:06:29,510 --> 00:06:28,639
environmental energy conditions and good

172
00:06:32,830 --> 00:06:29,520
solutions

173
00:06:34,710 --> 00:06:32,840
evolving under low environmental energy

174
00:06:36,469 --> 00:06:34,720
conditions

175
00:06:38,550 --> 00:06:36,479
we statistically confirmed this

176
00:06:39,909 --> 00:06:38,560
correlation between cellularity and

177
00:06:42,390 --> 00:06:39,919
metabolism

178
00:06:44,309 --> 00:06:42,400
showing that they are indeed co-evolving

179
00:06:45,590 --> 00:06:44,319
with the population average values for

180
00:06:48,309 --> 00:06:45,600
simulation set b

181
00:06:50,629 --> 00:06:48,319
taken at 250 step intervals showing a

182
00:06:54,950 --> 00:06:50,639
linear regression with a p-value

183
00:06:58,309 --> 00:06:56,870
this seems to make sense mechanistically

184
00:07:00,550 --> 00:06:58,319
if you think about it

185
00:07:02,390 --> 00:07:00,560
when easy-to-use resources are plentiful

186
00:07:04,230 --> 00:07:02,400
in the environment it makes

187
00:07:05,670 --> 00:07:04,240
sense to be more open to that

188
00:07:07,670 --> 00:07:05,680

environment and there is

189
00:07:09,110 --> 00:07:07,680
little selective advantage for complex

190
00:07:11,270 --> 00:07:09,120
metabolism

191
00:07:13,110 --> 00:07:11,280
moreover large inconsistent genomic

192
00:07:13,830 --> 00:07:13,120
alterations prevent the darwinian

193
00:07:17,350 --> 00:07:13,840
evolution

194
00:07:19,749 --> 00:07:17,360
of complex metabolism on the other hand

195
00:07:21,749 --> 00:07:19,759
in environments lacking such resources

196
00:07:23,510 --> 00:07:21,759
there is a high selective advantage for

197
00:07:26,309 --> 00:07:23,520
complex metabolisms

198
00:07:28,150 --> 00:07:26,319
capable of utilizing difficult to use

199
00:07:29,990 --> 00:07:28,160
resources

200
00:07:31,670 --> 00:07:30,000
this provides selective pressure for a

201
00:07:33,589 --> 00:07:31,680
more closed-off entity

202
00:07:36,150 --> 00:07:33,599
both in order to retain the metabolic

203
00:07:38,309 --> 00:07:36,160
products of that complex metabolism

204
00:07:43,029 --> 00:07:38,319
and to maintain genetic fidelity by

205
00:07:46,869 --> 00:07:45,189
i believe this also presents a mechanism

206
00:07:48,950 --> 00:07:46,879
for a darwinian transition

207
00:07:51,110 --> 00:07:48,960
where the availability of free energy

208
00:07:53,189 --> 00:07:51,120
drives a change in cellularity that is

209
00:07:55,110 --> 00:07:53,199
coupled with replication fidelity

210
00:07:57,270 --> 00:07:55,120
which is a necessary condition for the

211
00:07:58,869 --> 00:07:57,280
evolution of metabolism

212
00:08:00,550 --> 00:07:58,879
although full disclosure this

213
00:08:02,710 --> 00:08:00,560

interpretation is still a point of

214
00:08:06,070 --> 00:08:02,720
debate between my research collaborators

215
00:08:10,830 --> 00:08:08,869
so to summarize most scholars agree that

216
00:08:13,270 --> 00:08:10,840
life began in a geochemically rich

217
00:08:13,749 --> 00:08:13,280
environment whether that be hydrothermal

218
00:08:16,550 --> 00:08:13,759
vents

219
00:08:18,150 --> 00:08:16,560
tidal pools or any manner of prebiotic

220
00:08:20,150 --> 00:08:18,160
soup scenario

221
00:08:21,350 --> 00:08:20,160
our research adds to the understanding

222
00:08:23,749 --> 00:08:21,360
of life's origins

223
00:08:24,790 --> 00:08:23,759
by showing that cellularity likely arose

224
00:08:27,749 --> 00:08:24,800
in response to a

225
00:08:29,909 --> 00:08:27,759
change towards a less rich environment

226
00:08:31,749 --> 00:08:29,919
perhaps through some geological event

227
00:08:36,630 --> 00:08:31,759
or perhaps simply through the depletion

228
00:08:41,110 --> 00:08:38,949
our simulation research suggests that

229
00:08:41,589 --> 00:08:41,120
any geochemically rich origin of life

230
00:08:43,430 --> 00:08:41,599
setting

231
00:08:45,190 --> 00:08:43,440
would have prevented the evolution of

232
00:08:47,269 --> 00:08:45,200
cellular life

233
00:08:49,430 --> 00:08:47,279
replicators with access to plentiful

234
00:08:51,829 --> 00:08:49,440
prebiotic precursor molecules

235
00:08:53,110 --> 00:08:51,839
have an evolutionary incentive not to

236
00:08:56,870 --> 00:08:53,120
partition themselves

237
00:09:01,269 --> 00:08:59,030
we also found that cellularity and

238
00:09:03,990 --> 00:09:01,279

metabolism would have co-evolved

239
00:09:06,230 --> 00:09:04,000
in response to resource limitation

240
00:09:08,310 --> 00:09:06,240
cellularity provides a replicator

241
00:09:10,550 --> 00:09:08,320
greater protection from mutation

242
00:09:13,269 --> 00:09:10,560
allowing for a complex metabolism

243
00:09:15,509 --> 00:09:13,279
to evolve giving it access to previously

244
00:09:17,750 --> 00:09:15,519
non-useful resources

245
00:09:19,509 --> 00:09:17,760
cellularity also prevents the products

246
00:09:22,710 --> 00:09:19,519
of metabolic processing

247
00:09:24,949 --> 00:09:22,720
to be lost to the environment this gives

248
00:09:27,030 --> 00:09:24,959
a strong evolutionary incentive for

249
00:09:29,190 --> 00:09:27,040
complex metabolism to evolve

250
00:09:30,550 --> 00:09:29,200
when easy to utilize resources are

251
00:09:32,470 --> 00:09:30,560
exhausted

252
00:09:34,829 --> 00:09:32,480
which in turn applies an evolutionary

253
00:09:37,750 --> 00:09:34,839
pressure towards concurrently evolving

254
00:09:39,670 --> 00:09:37,760
cellularity

255
00:09:40,790 --> 00:09:39,680
these mechanisms suggest that the

256
00:09:43,269 --> 00:09:40,800
evolution of both

257
00:09:44,630 --> 00:09:43,279
metabolism and cellularity in early

258
00:09:46,790 --> 00:09:44,640
terrestrial life

259
00:09:48,470 --> 00:09:46,800
would have occurred in response to a

260
00:09:49,750 --> 00:09:48,480
change in the geochemical setting of

261
00:09:51,990 --> 00:09:49,760
life's origin

262
00:09:53,430 --> 00:09:52,000
from an environment rich in prebiotic

263
00:09:56,550 --> 00:09:53,440

precursor molecules

264
00:09:57,990 --> 00:09:56,560
to a poorer environment while any number

265
00:09:59,670 --> 00:09:58,000
of mechanisms

266
00:10:01,910 --> 00:09:59,680
could explain a change in the

267
00:10:04,069 --> 00:10:01,920
geochemical setting of early life

268
00:10:05,910 --> 00:10:04,079
a simple scenario suggests itself where

269
00:10:08,150 --> 00:10:05,920
in an initially rich environment

270
00:10:09,430 --> 00:10:08,160
becomes depleted as replicators grow in

271
00:10:11,350 --> 00:10:09,440
population

272
00:10:12,710 --> 00:10:11,360
driving the shift towards more complex

273
00:10:16,870 --> 00:10:12,720
metabolic processes

274
00:10:21,910 --> 00:10:19,350
lastly i propose a mechanism for a shift

275
00:10:24,710 --> 00:10:21,920
from a non-darwinian to darwinian system

276
00:10:27,110 --> 00:10:24,720
a so-called darwinian transition

277
00:10:29,190 --> 00:10:27,120
cellularity provides a replicator with

278
00:10:30,949 --> 00:10:29,200
greater protection from mutation

279
00:10:33,030 --> 00:10:30,959
with low mutation rates being a

280
00:10:35,990 --> 00:10:33,040
necessary condition for darwinian

281
00:10:38,389 --> 00:10:36,000
evolution of the genome to occur

282
00:10:40,550 --> 00:10:38,399
as such the very capacity for the genome

283
00:10:42,310 --> 00:10:40,560
to evolve in a darwinian manner

284
00:10:43,750 --> 00:10:42,320
may be subject to selection through

285
00:10:45,509 --> 00:10:43,760
cellularity

286
00:10:48,069 --> 00:10:45,519
and by extension the plenitude or

287
00:10:50,310 --> 00:10:48,079
scarcity of resources in the environment

288
00:10:53,350 --> 00:10:50,320

we present a driving mechanism for a

289
00:10:55,910 --> 00:10:53,360
darwinian transition

290
00:10:57,430 --> 00:10:55,920
in conclusion we found that cellularity

291
00:10:59,509 --> 00:10:57,440
and complex metabolism

292
00:11:00,630 --> 00:10:59,519
are not an inevitable outcome of

293
00:11:02,630 --> 00:11:00,640
evolution

294
00:11:04,550 --> 00:11:02,640
but rather they are linked traits that

295
00:11:06,710 --> 00:11:04,560
are reversibly selected for

296
00:11:08,150 --> 00:11:06,720
by the environmental availability of

297
00:11:10,470 --> 00:11:08,160
resources

298
00:11:11,590 --> 00:11:10,480
cellularity and complex metabolism in

299
00:11:13,990 --> 00:11:11,600
early earth-based

300
00:11:16,069 --> 00:11:14,000
life likely evolved in response to a

301
00:11:18,069 --> 00:11:16,079
changing geochemical environment

302
00:11:19,990 --> 00:11:18,079
and facilitated life's capacity for

303
00:11:21,430 --> 00:11:20,000
darwinian evolution

304
00:11:24,230 --> 00:11:21,440
thank you for your attention and

305
00:11:25,990 --> 00:11:24,240
interest in this work

306
00:11:28,069 --> 00:11:26,000
if you are further interested this

307
00:11:30,710 --> 00:11:28,079
research was published in greater detail

308
00:11:32,069 --> 00:11:30,720
in the journal of molecular evolution if

309
00:11:33,750 --> 00:11:32,079
you are interested in using the

310
00:11:34,470 --> 00:11:33,760
simulation or contributing to the

311
00:11:37,269 --> 00:11:34,480
software

312
00:11:37,829 --> 00:11:37,279
it is publicly available on github if

313
00:11:40,310 --> 00:11:37,839

you have

314
00:11:42,790 --> 00:11:40,320
any further questions or comments please

315
00:11:45,829 --> 00:11:42,800
feel free to contact me at utaatakagi

316
00:11:47,430 --> 00:11:45,839
gmail.com i would like to acknowledge my

317
00:11:50,150 --> 00:11:47,440
collaborators in this research

318
00:11:51,590 --> 00:11:50,160
diaper new yen dr tom wexler and dr

319
00:11:53,509 --> 00:11:51,600
aaron goldman

320
00:11:55,269 --> 00:11:53,519
this research was enabled by the oberlin

321
00:11:56,949 --> 00:11:55,279
college department of biology and

322
00:11:59,269 --> 00:11:56,959
funding was provided by the national

323
00:12:02,150 --> 00:11:59,279
aeronautics and space administration