
AbSciCon
2019

1
00:00:00,790 --> 00:00:07,320
[Music]

2
00:00:12,320 --> 00:00:09,200
[Applause]

3
00:00:15,110 --> 00:00:12,330
thank you so much Madeline thank you for

4
00:00:16,340 --> 00:00:15,120
staying this late and we're going to

5
00:00:19,250 --> 00:00:16,350
talk about the evolution of complex

6
00:00:21,500 --> 00:00:19,260
genetics as a part of a larger program

7
00:00:23,740 --> 00:00:21,510
looking at the evolution of complexity

8
00:00:25,760 --> 00:00:23,750
throughout life a couple of previous

9
00:00:27,290 --> 00:00:25,770
speakers have talked about this I'm not

10
00:00:31,220 --> 00:00:27,300
going to talk about the transitions

11
00:00:33,260 --> 00:00:31,230
leading to complex life we look at

12
00:00:34,640 --> 00:00:33,270
complex function particularly how the

13
00:00:37,400 --> 00:00:34,650

complex organism such as ourselves

14
00:00:39,170 --> 00:00:37,410
evolve and not in terms of specific at

15
00:00:41,180 --> 00:00:39,180
that amaze you know has the human

16
00:00:44,090 --> 00:00:41,190
eyeball evolved but how does things like

17
00:00:46,490 --> 00:00:44,100
imaging vision evolve an imaging vision

18
00:00:48,620 --> 00:00:46,500
is evolved multiple times and through

19
00:00:51,500 --> 00:00:48,630
different independent routes things like

20
00:00:53,530 --> 00:00:51,510
powered flight also these complex

21
00:00:55,340 --> 00:00:53,540
functions have evolved independently

22
00:00:57,139 --> 00:00:55,350
multiple times once you've reached

23
00:01:00,260 --> 00:00:57,149
multicellularity which is a completely

24
00:01:02,900 --> 00:01:00,270
different issue so we want to know why

25
00:01:05,150 --> 00:01:02,910
is this happen and and therefore could

26
00:01:06,740 --> 00:01:05,160
it happen on other worlds with a life

27
00:01:09,139 --> 00:01:06,750
capable of doing that what's the

28
00:01:11,330 --> 00:01:09,149
requirements in order to happen life

29
00:01:13,789 --> 00:01:11,340
with complex function you have to be

30
00:01:16,010 --> 00:01:13,799
able to have differentiated cells cells

31
00:01:18,670 --> 00:01:16,020
that can perform lots of functions have

32
00:01:20,990 --> 00:01:18,680
specialized molecules within them and

33
00:01:23,270 --> 00:01:21,000
they have to be put in place in the

34
00:01:24,740 --> 00:01:23,280
right place and at the right time and in

35
00:01:26,300 --> 00:01:24,750
developmental programs after then we

36
00:01:27,679 --> 00:01:26,310
take it away again at the right place in

37
00:01:29,929 --> 00:01:27,689
the right time so this is a very

38
00:01:31,609 --> 00:01:29,939

complicated control process and that

39
00:01:34,039 --> 00:01:31,619
speaks to there being a complex genetic

40
00:01:36,500 --> 00:01:34,049
program underneath that to make that all

41
00:01:38,810 --> 00:01:36,510
work so we've been asking is this the

42
00:01:40,730 --> 00:01:38,820
evolution of this complex genetic

43
00:01:43,190 --> 00:01:40,740
program the capability to evolve a

44
00:01:47,270 --> 00:01:43,200
complex program itself a bottleneck in

45
00:01:49,550 --> 00:01:47,280
the evolution of complex life complex

46
00:01:51,350 --> 00:01:49,560
genetics is pretty much restricted to

47
00:01:54,219 --> 00:01:51,360
the eukaryotes of this sort of

48
00:01:56,870 --> 00:01:54,229
complexity so eukaryotes tend to have

49
00:01:58,940 --> 00:01:56,880
big genomes this is quite an old plot

50
00:02:01,609 --> 00:01:58,950
this is all the completely sequence

51
00:02:03,440 --> 00:02:01,619
genomes I could find our database these

52
00:02:06,020 --> 00:02:03,450
green guys here the eukaryotes along

53
00:02:10,850 --> 00:02:06,030
here we have gene and size in log scale

54
00:02:12,380 --> 00:02:10,860
and mega bases rafted Giga bases here is

55
00:02:14,539 --> 00:02:12,390
a huge amounts of information and

56
00:02:16,130 --> 00:02:14,549
eukaryotes can develop and have

57
00:02:18,020 --> 00:02:16,140
developed independently in multiple

58
00:02:20,960 --> 00:02:18,030
times very complex developmental

59
00:02:22,530 --> 00:02:20,970
programs by contrast prokaryotes and I

60
00:02:24,179 --> 00:02:22,540
don't apologize for lumping every

61
00:02:25,740 --> 00:02:24,189
else together in prokaryotes and it's

62
00:02:27,660 --> 00:02:25,750
old-fashioned there you go

63
00:02:29,759 --> 00:02:27,670

tend to have much smaller genomes and

64
00:02:32,309 --> 00:02:29,769
when they do develop multicellularity it

65
00:02:34,589 --> 00:02:32,319
is simple this is limited a very limited

66
00:02:36,660 --> 00:02:34,599
number of cell types and it is quite

67
00:02:38,130 --> 00:02:36,670
often faculty to do so they can switch

68
00:02:40,830 --> 00:02:38,140
between unicellular and multicellular

69
00:02:45,149 --> 00:02:40,840
states you and I can't do that

70
00:02:47,369 --> 00:02:45,159
well I can't with the genomes do overlap

71
00:02:49,229 --> 00:02:47,379
in sizes there's there's not a distinct

72
00:02:50,699 --> 00:02:49,239
size gap where you have to jump from

73
00:02:52,649 --> 00:02:50,709
prokaryote to eukaryote and they also

74
00:02:54,449 --> 00:02:52,659
overlap in chemistry and this is

75
00:02:56,879 --> 00:02:54,459
something we spent rather too much time

76
00:02:58,440 --> 00:02:56,889
illustrating but I'll try to be brief if

77
00:03:01,140 --> 00:02:58,450
you look at all the chemistry that's

78
00:03:03,780 --> 00:03:01,150
involved in how genes are controlled in

79
00:03:06,119 --> 00:03:03,790
an organism so you have DNA binding to

80
00:03:08,129 --> 00:03:06,129
RNA RNA binding to proteins proteins

81
00:03:10,710 --> 00:03:08,139
binding to DNA there's a whole zoo of

82
00:03:12,659 --> 00:03:10,720
interactions and you can modulate any

83
00:03:14,250 --> 00:03:12,669
one of these by reacting them with

84
00:03:16,530 --> 00:03:14,260
different chemicals and then each of

85
00:03:19,740 --> 00:03:16,540
them can control different steps in the

86
00:03:21,539 --> 00:03:19,750
in the process of activating a gene so

87
00:03:22,770 --> 00:03:21,549
from the initial transcription all the

88
00:03:24,270 --> 00:03:22,780

way down here to breaking down the

89
00:03:26,719 --> 00:03:24,280
protein in the end of the process and

90
00:03:29,219 --> 00:03:26,729
you ask what combinations actually occur

91
00:03:31,409 --> 00:03:29,229
okay so you can have a targeting moiety

92
00:03:32,640 --> 00:03:31,419
a protein molecule over here and it

93
00:03:34,259 --> 00:03:32,650
binds to something else and then

94
00:03:36,240 --> 00:03:34,269
something happens so you've got

95
00:03:37,830 --> 00:03:36,250
targeting targeted and what happens and

96
00:03:39,809 --> 00:03:37,840
if you look at what happens in the

97
00:03:43,110 --> 00:03:39,819
eukaryotes the answer is pretty much

98
00:03:45,479 --> 00:03:43,120
everything and you look at what happens

99
00:03:48,240 --> 00:03:45,489
in the prokaryotes and the answer is

100
00:03:49,740 --> 00:03:48,250
pretty much everything all the

101
00:03:52,289 --> 00:03:49,750
combinations are there there isn't a

102
00:03:54,390 --> 00:03:52,299
distinct chemical pattern of this the

103
00:03:56,280 --> 00:03:54,400
chemistry is similar in all the major

104
00:03:58,259 --> 00:03:56,290
domains differently distributed

105
00:03:59,879 --> 00:03:58,269
different emphasis but there's not a

106
00:04:01,979 --> 00:03:59,889
fundamental chemical difference and

107
00:04:03,449 --> 00:04:01,989
these independently evolved they are not

108
00:04:04,710 --> 00:04:03,459
the same and we know that independent

109
00:04:06,420 --> 00:04:04,720
evolved because they're completely

110
00:04:09,030 --> 00:04:06,430
different sequences different structures

111
00:04:11,789 --> 00:04:09,040
and so on and so how our hypothesis

112
00:04:16,229 --> 00:04:11,799
which we developed is that the inherent

113
00:04:20,129 --> 00:04:16,239

difference is the control logic and that

114
00:04:22,860 --> 00:04:20,139
eukaryotes us have a gene structure that

115
00:04:25,020 --> 00:04:22,870
is by default off if you add a gene to a

116
00:04:26,610 --> 00:04:25,030
eukaryotic genome it will by default not

117
00:04:29,820 --> 00:04:26,620
do anything you have to do a lot of work

118
00:04:32,490 --> 00:04:29,830
and activity to turn it on prokaryotes

119
00:04:34,560 --> 00:04:32,500
have a default on logic if you add a bit

120
00:04:36,980 --> 00:04:34,570
of DNA to a prokaryotic genome it will

121
00:04:38,970 --> 00:04:36,990
tend to do something

122
00:04:40,470 --> 00:04:38,980
there's lots of everything I won't go

123
00:04:42,120 --> 00:04:40,480
into this we've written it very long and

124
00:04:45,540 --> 00:04:42,130
to be honest rather boring paper about

125
00:04:47,190 --> 00:04:45,550
this I should I should emphasize this is

126
00:04:50,010 --> 00:04:47,200
a style of control

127
00:04:52,170 --> 00:04:50,020
this isn't absolute okay see course you

128
00:04:53,910 --> 00:04:52,180
can you find examples of gene systems in

129
00:04:55,590 --> 00:04:53,920
prokaryotes that default off of course

130
00:04:58,770 --> 00:04:55,600
you can find ones in eukaryotes that are

131
00:05:00,630 --> 00:04:58,780
default on but it is a general style of

132
00:05:02,970 --> 00:05:00,640
control and our hypothesis which I'm not

133
00:05:07,350 --> 00:05:02,980
going to justify today extent of time is

134
00:05:09,810 --> 00:05:07,360
that this is related to why eukaryotic

135
00:05:12,120 --> 00:05:09,820
genomes can be amplified through gene

136
00:05:16,040 --> 00:05:12,130
duplication and made more complex in in

137
00:05:20,250 --> 00:05:16,050
it pre adapts you to be able to do that

138
00:05:21,690 --> 00:05:20,260

okay so is that true while we can't go

139
00:05:22,980 --> 00:05:21,700
back in with a time machine you know the

140
00:05:25,590 --> 00:05:22,990
snowflake thing is great but we'd have

141
00:05:27,090 --> 00:05:25,600
to run the snowflakes a yeast snowflake

142
00:05:29,010 --> 00:05:27,100
experience experiment for a million

143
00:05:30,750 --> 00:05:29,020
years to to test this that's not really

144
00:05:34,110 --> 00:05:30,760
practical with a five year project grant

145
00:05:36,390 --> 00:05:34,120
so we are so we we thought we model it

146
00:05:38,520 --> 00:05:36,400
and wanted as a model of the sort of

147
00:05:40,730 --> 00:05:38,530
complexity that occurs in genetic

148
00:05:43,710 --> 00:05:40,740
systems and evolving genetic systems

149
00:05:45,090 --> 00:05:43,720
which are spaghetti code and if you look

150
00:05:46,650 --> 00:05:45,100
at the sort of control maps that

151
00:05:48,360 --> 00:05:46,660
biochemists like to put up on their wall

152
00:05:50,760 --> 00:05:48,370
to show how much are they are they are

153
00:05:52,350 --> 00:05:50,770
completely spaghetti code control

154
00:05:55,700 --> 00:05:52,360
systems there's no layers there's no

155
00:05:57,630 --> 00:05:55,710
modularity so we wanted a model app and

156
00:05:59,490 --> 00:05:57,640
there's all the feedback between gene

157
00:06:01,560 --> 00:05:59,500
products and gene expression so this

158
00:06:03,120 --> 00:06:01,570
sort of highly abstract Network

159
00:06:05,370 --> 00:06:03,130
formulation didn't really work for us

160
00:06:07,530 --> 00:06:05,380
but equally we didn't want a model of

161
00:06:09,240 --> 00:06:07,540
detail chemistry this sort of thing down

162
00:06:11,190 --> 00:06:09,250
here proteins behind the DNA because if

163
00:06:13,620 --> 00:06:11,200

you try to do that on a on a cell basis

164
00:06:15,390 --> 00:06:13,630
then well that would be too complicated

165
00:06:17,130 --> 00:06:15,400
we actually I wanted something that

166
00:06:19,260 --> 00:06:17,140
looked like the classic undergraduate

167
00:06:21,570 --> 00:06:19,270
textbook diagram of a jakob and mono

168
00:06:24,030 --> 00:06:21,580
control model for those who biochemists

169
00:06:27,870 --> 00:06:24,040
some a it looks like this okay so that's

170
00:06:30,840 --> 00:06:27,880
what we built so this is the model and

171
00:06:32,670 --> 00:06:30,850
it's conceptually quite simple I'll tell

172
00:06:35,520 --> 00:06:32,680
you about the coding later that's a

173
00:06:37,590 --> 00:06:35,530
disaster so we have a population of

174
00:06:39,300 --> 00:06:37,600
organisms in yes we do have five

175
00:06:42,870 --> 00:06:39,310
organisms at the moment in the model and

176
00:06:44,960 --> 00:06:42,880
those organisms contain genes and those

177
00:06:48,060 --> 00:06:44,970
genes can be active and produce

178
00:06:49,470 --> 00:06:48,070
transcripts we need selection in this

179
00:06:52,650 --> 00:06:49,480
this is an eeveelution area mode

180
00:06:55,140 --> 00:06:52,660
so for a organism to be fit those

181
00:06:57,270 --> 00:06:55,150
transcripts must match an environment

182
00:06:59,180 --> 00:06:57,280
and there are positive aspects of that

183
00:07:01,800 --> 00:06:59,190
environment things it must do and

184
00:07:04,110 --> 00:07:01,810
negative aspects things it must not do

185
00:07:05,340 --> 00:07:04,120
if it's just positive then you just

186
00:07:09,480 --> 00:07:05,350
select something that makes everything

187
00:07:10,950 --> 00:07:09,490
and then you is one ok how do you tell

188
00:07:13,530 --> 00:07:10,960

whether you've got a transcript what the

189
00:07:16,230 --> 00:07:13,540
transcription of each gene is controlled

190
00:07:17,790 --> 00:07:16,240
by regulatory elements and you've got

191
00:07:20,100 --> 00:07:17,800
positive regulatory elements and

192
00:07:23,010 --> 00:07:20,110
negative regulator elements positive

193
00:07:24,720 --> 00:07:23,020
turn it on negative turn it on a note

194
00:07:26,310 --> 00:07:24,730
we've only got one type of sequence here

195
00:07:28,380 --> 00:07:26,320
we haven't got translation built into

196
00:07:30,690 --> 00:07:28,390
this I thought of doing that and then

197
00:07:32,610 --> 00:07:30,700
thought it's too complicated at this

198
00:07:34,650 --> 00:07:32,620
stage so if you wonder if you believe in

199
00:07:38,970 --> 00:07:34,660
the RNA world hypothesis this is a sort

200
00:07:40,220 --> 00:07:38,980
of RNA world type Gina how do you

201
00:07:42,990 --> 00:07:40,230
control the genes

202
00:07:44,880 --> 00:07:43,000
well pretty straightforward we've got

203
00:07:47,040 --> 00:07:44,890
positive elements and negative elements

204
00:07:48,840 --> 00:07:47,050
if the number of active positive

205
00:07:51,570 --> 00:07:48,850
elements outweighs the number of active

206
00:07:54,360 --> 00:07:51,580
negative elements it's on if it doesn't

207
00:07:56,190 --> 00:07:54,370
it's off keep it is want to be mean by

208
00:07:57,960 --> 00:07:56,200
active we want feedback between what the

209
00:08:01,260 --> 00:07:57,970
genome does and how it controls its

210
00:08:04,530 --> 00:08:01,270
genes so a genome produces a sequence

211
00:08:07,530 --> 00:08:04,540
here a transcript and we ask for each

212
00:08:09,840 --> 00:08:07,540
regulatory element here does that match

213
00:08:14,430 --> 00:08:09,850

a transcript that's being made at the

214
00:08:16,290 --> 00:08:14,440
moment so here's a gene two on one off

215
00:08:18,720 --> 00:08:16,300
this one is produced let's say the

216
00:08:21,570 --> 00:08:18,730
transcript of this sequence here's

217
00:08:24,530 --> 00:08:21,580
another gene this regulatory element has

218
00:08:28,200 --> 00:08:24,540
sequence ACA that matches that bit there

219
00:08:30,990 --> 00:08:28,210
so that's all this one is see see see

220
00:08:34,409 --> 00:08:31,000
that doesn't match anything here that's

221
00:08:36,390 --> 00:08:34,419
not on this is BC that's not on this is

222
00:08:38,909 --> 00:08:36,400
see that matches that one there so

223
00:08:41,370 --> 00:08:38,919
that's on and you do this for all these

224
00:08:44,400 --> 00:08:41,380
regulator elements in all the genes

225
00:08:46,080 --> 00:08:44,410
matching all the transcripts and that

226
00:08:49,260 --> 00:08:46,090
tells you whether that particular gene

227
00:08:52,500 --> 00:08:49,270
is on or not two things to note this is

228
00:08:53,820 --> 00:08:52,510
a model of regulatory control of genes

229
00:08:55,290 --> 00:08:53,830
is not structural they're not saying

230
00:08:57,270 --> 00:08:55,300
this is an enzyme or something does

231
00:09:01,950 --> 00:08:57,280
something clever so it's a model of gene

232
00:09:03,280 --> 00:09:01,960
regulation and secondly this is a bit

233
00:09:05,110 --> 00:09:03,290
obvious but but

234
00:09:07,210 --> 00:09:05,120
said anyway because it's late in the day

235
00:09:10,660 --> 00:09:07,220
and you know your blood glucose levels

236
00:09:13,780 --> 00:09:10,670
are dropping if your regulatory sequence

237
00:09:17,439 --> 00:09:13,790
here is short it is more likely to match

238
00:09:19,569 --> 00:09:17,449

something here so if you have a short

239
00:09:24,610 --> 00:09:19,579
regulatory sequence it's more likely to

240
00:09:26,920 --> 00:09:24,620
be active okay lots of variables in

241
00:09:28,240 --> 00:09:26,930
incomes of it my goodness you can have

242
00:09:29,949 --> 00:09:28,250
fun with this so we're not going to go

243
00:09:31,960 --> 00:09:29,959
through all that we're going to look to

244
00:09:33,879 --> 00:09:31,970
some summary statistics of what the

245
00:09:36,730 --> 00:09:33,889
output looks at and there are two things

246
00:09:39,939 --> 00:09:36,740
we're going to focus on one in

247
00:09:41,379 --> 00:09:39,949
particular if you have a lot of these

248
00:09:42,970 --> 00:09:41,389
negative elements and you could lose

249
00:09:45,249 --> 00:09:42,980
elements through selection they can be

250
00:09:46,960 --> 00:09:45,259
deleted so if an organism decides it

251
00:09:48,370 --> 00:09:46,970
only needs one regulatory element that's

252
00:09:51,100 --> 00:09:48,380
absolutely fine eventually it'll get

253
00:09:52,689 --> 00:09:51,110
that lots of negative elements compared

254
00:09:55,449 --> 00:09:52,699
to positive suggests you've got us of

255
00:09:58,030 --> 00:09:55,459
default off state now be more likely to

256
00:09:59,920 --> 00:09:58,040
be off than on and lots of short

257
00:10:01,480 --> 00:09:59,930
negative elements will mean you'll have

258
00:10:09,629 --> 00:10:01,490
a default off state they're more likely

259
00:10:13,269 --> 00:10:11,680
there are lots of outputs you can

260
00:10:15,340 --> 00:10:13,279
collect from this so you can measure

261
00:10:17,769 --> 00:10:15,350
average this is just an example run

262
00:10:19,240 --> 00:10:17,779
measure average dream gene length which

263
00:10:20,889 --> 00:10:19,250

turns out to be quite interesting but I

264
00:10:22,840 --> 00:10:20,899
won't go into why you can measure the

265
00:10:24,160 --> 00:10:22,850
number of express genes number producing

266
00:10:27,639 --> 00:10:24,170
transcripts which turns out to be quite

267
00:10:29,319 --> 00:10:27,649
dull and I won't say why you start out

268
00:10:31,990 --> 00:10:29,329
with an entirely random genome and it

269
00:10:33,250 --> 00:10:32,000
adapts fairly rapidly and then sort of

270
00:10:36,730 --> 00:10:33,260
either plateaus or

271
00:10:38,050 --> 00:10:36,740
or slowly increases adaptation and you

272
00:10:39,910 --> 00:10:38,060
can look at the fitness and that turns

273
00:10:42,040 --> 00:10:39,920
out to be really confusing but I went to

274
00:10:45,250 --> 00:10:42,050
a why what I got to focus on is is this

275
00:10:47,350 --> 00:10:45,260
thing here this is for each gene pay

276
00:10:49,650 --> 00:10:47,360
attention now this is complicated the

277
00:10:52,470 --> 00:10:49,660
each gene you look at the shortest

278
00:10:54,819 --> 00:10:52,480
positive and the shortest negative

279
00:10:58,629 --> 00:10:54,829
regulatory element so that's the one

280
00:11:00,370 --> 00:10:58,639
most likely to be active and then you

281
00:11:02,170 --> 00:11:00,380
say for all the genes in the genome what

282
00:11:06,490 --> 00:11:02,180
is the average length of those short

283
00:11:09,129 --> 00:11:06,500
positive and short negative elements if

284
00:11:11,050 --> 00:11:09,139
the short negative ones are shorter on

285
00:11:13,449 --> 00:11:11,060
average they're more likely to be on

286
00:11:15,550 --> 00:11:13,459
that means the genes on average are more

287
00:11:17,490 --> 00:11:15,560
likely to be suppressed off you've got a

288
00:11:21,120 --> 00:11:17,500

default negative

289
00:11:22,860 --> 00:11:21,130
default off-mode if the Shh if the

290
00:11:24,870 --> 00:11:22,870
positive one said to be shorter they're

291
00:11:31,520 --> 00:11:24,880
more likely to be on and so you have a

292
00:11:34,020 --> 00:11:31,530
default on mode so in this example

293
00:11:36,120 --> 00:11:34,030
here's the average minimum regulatory

294
00:11:38,910 --> 00:11:36,130
element length blah blah blah rate is

295
00:11:42,360 --> 00:11:38,920
negative blue is positive the red is a

296
00:11:44,010 --> 00:11:42,370
little bit longer than the blue there's

297
00:11:46,410 --> 00:11:44,020
a lot of noise on that because I had the

298
00:11:47,790 --> 00:11:46,420
mutation turned up too high this is a

299
00:11:49,350 --> 00:11:47,800
great thing about doing life in

300
00:11:51,450 --> 00:11:49,360
computers you can turn mutation up and

301
00:11:54,420 --> 00:11:51,460
down and things like that so this is yes

302
00:11:57,840 --> 00:11:54,430
this is very weakly a prokaryotic

303
00:12:00,690 --> 00:11:57,850
default on mo but he's pretty weak it's

304
00:12:02,760 --> 00:12:00,700
pretty messy I mean yeah you wouldn't

305
00:12:06,650 --> 00:12:02,770
call that a good transit signal in it

306
00:12:09,060 --> 00:12:06,660
and a exoplanet good lecture with you so

307
00:12:10,770 --> 00:12:09,070
here's a few more is another one - doing

308
00:12:12,480 --> 00:12:10,780
the same thing I just dialed down the

309
00:12:15,180 --> 00:12:12,490
mutation rate so it's less noisy but

310
00:12:17,160 --> 00:12:15,190
it's it's weakly default on I wouldn't

311
00:12:19,020 --> 00:12:17,170
say anything this one can't decide what

312
00:12:20,670 --> 00:12:19,030
it's doing it's just wobbling all over

313
00:12:22,740 --> 00:12:20,680

the place but these two are more

314
00:12:25,650 --> 00:12:22,750
interesting here we've started to evolve

315
00:12:29,760 --> 00:12:25,660
a clear pattern a signal of some sort

316
00:12:31,590 --> 00:12:29,770
this one the positive ones these are the

317
00:12:33,750 --> 00:12:31,600
elements that turn things on tend to be

318
00:12:37,230 --> 00:12:33,760
shorter suggest you go the prokaryotic

319
00:12:38,910 --> 00:12:37,240
type default on genetics evolving and

320
00:12:43,290 --> 00:12:38,920
this is the other way around you've got

321
00:12:45,180 --> 00:12:43,300
the negative evolving and you do this a

322
00:12:47,430 --> 00:12:45,190
lot for different combinations of

323
00:12:49,740 --> 00:12:47,440
parameters and different in our days of

324
00:12:51,960 --> 00:12:49,750
the week and things and you try to

325
00:12:54,330 --> 00:12:51,970
correlate are you getting default on or

326
00:12:56,220 --> 00:12:54,340
default off with all the different

327
00:12:59,130 --> 00:12:56,230
inputs and this is what we've got so far

328
00:13:00,990 --> 00:12:59,140
this is really preliminary this is just

329
00:13:03,540 --> 00:13:01,000
straightforward correlation coefficient

330
00:13:06,420 --> 00:13:03,550
greater than I think 32 is significant

331
00:13:09,150 --> 00:13:06,430
to 95% and those are colored in red here

332
00:13:11,010 --> 00:13:09,160
and one really surprising thing that

333
00:13:12,780 --> 00:13:11,020
left out I thought you know it's going

334
00:13:14,220 --> 00:13:12,790
to be you put complicated environments

335
00:13:15,780 --> 00:13:14,230
in the genomes gonna have to be

336
00:13:17,580 --> 00:13:15,790
complicated to respond to each and

337
00:13:20,070 --> 00:13:17,590
you'll have eukaryotic and you'll have

338
00:13:24,330 --> 00:13:20,080

kangaroos evolving it didn't care what

339
00:13:27,570 --> 00:13:24,340
the environment was wow that really

340
00:13:29,460 --> 00:13:27,580
annoyed me it was all about genome

341
00:13:30,780 --> 00:13:29,470
structure the correlations that were

342
00:13:31,380 --> 00:13:30,790
more significant was about genome

343
00:13:34,290 --> 00:13:31,390
structure

344
00:13:36,389 --> 00:13:34,300
and the thing that correlative would

345
00:13:38,190 --> 00:13:36,399
default off evolving that sort of thing

346
00:13:39,750 --> 00:13:38,200
where you've got a nice curve separating

347
00:13:44,699 --> 00:13:39,760
like this saying you've got a default of

348
00:13:46,259 --> 00:13:44,709
wealth having a lot of genes you could

349
00:13:47,970 --> 00:13:46,269
also solve your world alter the number

350
00:13:51,540 --> 00:13:47,980
of bases in your DNA in this - you could

351
00:13:56,759 --> 00:13:51,550
say let's have a seven base genome let's

352
00:13:59,130 --> 00:13:56,769
see Steve better do that few bases

353
00:14:01,740 --> 00:13:59,140
simpler genome correlated with default

354
00:14:04,290 --> 00:14:01,750
off short genes so this is simple

355
00:14:08,060 --> 00:14:04,300
genomes it's the number of genes that

356
00:14:10,199 --> 00:14:08,070
correlated weird so this isn't

357
00:14:13,500 --> 00:14:10,209
inconsistent with the original idea that

358
00:14:15,480 --> 00:14:13,510
having a lot of genes means it's or

359
00:14:18,210 --> 00:14:15,490
favors evolution of this Vieux Carre

360
00:14:22,079 --> 00:14:18,220
optic type default off genetics it's not

361
00:14:25,230 --> 00:14:22,089
entirely supportive of it it's about

362
00:14:27,449 --> 00:14:25,240
many regulatory elements not their size

363
00:14:28,860 --> 00:14:27,459

or complexity and this is this sort of

364
00:14:30,329 --> 00:14:28,870
this is why I mentioned the RNA world

365
00:14:31,680 --> 00:14:30,339
you know because this sort of looks like

366
00:14:35,490 --> 00:14:31,690
that doesn't it lots and lots and lots

367
00:14:38,310 --> 00:14:35,500
of RNAs a little short ones and this

368
00:14:42,090 --> 00:14:38,320
hints that default off was the standard

369
00:14:45,630 --> 00:14:42,100
for first prototype so caveat it's only

370
00:14:47,100 --> 00:14:45,640
a model okay I wish more astrobiology

371
00:14:49,019 --> 00:14:47,110
talks would say it's only a model

372
00:14:52,680 --> 00:14:49,029
because quite often it is it's only a

373
00:14:57,600 --> 00:14:52,690
model in Excel I did this in Excel okay

374
00:14:59,220 --> 00:14:57,610
this is a bad idea and I'm very

375
00:15:00,870 --> 00:14:59,230
fortunate then Enrico Borriello it's

376
00:15:02,490 --> 00:15:00,880
Sarah Walker's groupid AC who is now

377
00:15:05,310 --> 00:15:02,500
recoding this into something this

378
00:15:07,920 --> 00:15:05,320
actually looks like code so conclusion

379
00:15:09,300 --> 00:15:07,930
genetics start ends differ I'll be

380
00:15:11,100 --> 00:15:09,310
developed a model of that that's not

381
00:15:13,530 --> 00:15:11,110
completely abstract and yet not

382
00:15:16,530 --> 00:15:13,540
chemically specific initial results hint

383
00:15:18,449 --> 00:15:16,540
and this is just a hint that more genes

384
00:15:20,880 --> 00:15:18,459
mean default off style

385
00:15:23,400 --> 00:15:20,890
it's about regulatory group number not

386
00:15:24,350 --> 00:15:23,410
about genome complexity and that hints

387
00:15:26,460 --> 00:15:24,360
that this might be a primitive

388
00:15:30,710 --> 00:15:26,470

characteristic in other words you and I

389
00:15:35,720 --> 00:15:30,720
are more similar to Luca than e.coli

390
00:15:38,040 --> 00:15:35,730
your thoughts welcome even yours and

391
00:15:39,689 --> 00:15:38,050
because this is really pretty Murray and

392
00:15:42,389 --> 00:15:39,699
we want your thoughts and input before

393
00:15:45,060 --> 00:15:42,399
we go on to the next stage those are the

394
00:15:49,230 --> 00:15:45,070
papers and one last call out

395
00:15:50,520 --> 00:15:49,240
mdps life I've published in life several

396
00:15:53,580 --> 00:15:50,530
times and served one of my colleagues

397
00:15:54,840 --> 00:15:53,590
and we published in another journal at

398
00:15:56,460 --> 00:15:54,850
all you'll be familiar with that I

399
00:15:59,360 --> 00:15:56,470
wouldn't name that we publishing quite a

400
00:16:03,360 --> 00:15:59,370
bit in this context and these guys are

401
00:16:07,500 --> 00:16:03,370
professional faster cheaper and open

402
00:16:09,720 --> 00:16:07,510
access so I'd have a look at em DPI's

403
00:16:12,210 --> 00:16:09,730
life as a potential and output for your

404
00:16:13,350 --> 00:16:12,220
next paper I think you'll be pleasantly

405
00:16:15,600 --> 00:16:13,360
surprised by the level of

406
00:16:18,210 --> 00:16:15,610
professionalism and speed with which

407
00:16:19,610 --> 00:16:18,220
they can get your stuff out thank you

408
00:16:19,950 --> 00:16:19,620
very much

409
00:16:21,870 --> 00:16:19,960
[Applause]