EARTH-LIFE
SCIENCE
INSTITUTE

Loren Williams

RNA & Protein: A match made in the Hadean

1
00:00:00,160 --> 00:00:13,870
[Music]

2
00:00:18,980 --> 00:00:16,790
first I I want to introduce some people

3
00:00:21,849 --> 00:00:18,990
in my lab who did the work

4
00:00:27,380 --> 00:00:21,859
Nick HUD he's one of my best postdocs

5
00:00:31,009 --> 00:00:27,390
and Nick is my close collaborator and

6
00:00:32,900 --> 00:00:31,019
and a lot of what I'm doing is it has to

7
00:00:35,930 --> 00:00:32,910
do with our interaction but I'm gonna

8
00:00:36,590 --> 00:00:35,940
talk about work from Anton Petrov second

9
00:00:40,819 --> 00:00:36,600
right there

10
00:00:43,700 --> 00:00:40,829
Nick Kovacs Katherine Lanier Jessica

11
00:00:46,520 --> 00:00:43,710
Bowman and then my collaborators I have

12
00:00:49,250 --> 00:00:46,530
a some very nice collaborators at

13
00:00:52,580 --> 00:00:49,260

Georgia Tech Nick HUD Roger work tell

14
00:00:56,119 --> 00:00:52,590
and then George Fox who said Houston and

15
00:00:59,240 --> 00:00:56,129
I'd like to thank the organizers Irina

16
00:01:00,560 --> 00:00:59,250
wherever you are for inviting me here

17
00:01:02,779 --> 00:01:00,570
and running such a great meeting and

18
00:01:05,600 --> 00:01:02,789
keeping it so organized of course this

19
00:01:07,940 --> 00:01:05,610
is Japan so of course it's organized and

20
00:01:10,280 --> 00:01:07,950
I like to say hello to my friends in

21
00:01:10,850 --> 00:01:10,290
California at the end the I good

22
00:01:21,890 --> 00:01:10,860
afternoon

23
00:01:24,590 --> 00:01:21,900
so we're what we do is we try to

24
00:01:28,370 --> 00:01:24,600
understand the origin of life from

25
00:01:31,310 --> 00:01:28,380
extant biology and we look at the tree

26
00:01:33,080 --> 00:01:31,320
of life or the circle of life I guess if

27
00:01:35,600 --> 00:01:33,090
this thing and we look all around this

28
00:01:37,490 --> 00:01:35,610
thing and we try to use that to walk

29
00:01:41,060 --> 00:01:37,500
backwards in time and there's a lot of

30
00:01:44,870 --> 00:01:41,070
people who do this kind of thing but we

31
00:01:46,429 --> 00:01:44,880
I think probably go back further and one

32
00:01:48,889 --> 00:01:46,439
of the one of the ways we can do that is

33
00:01:50,030 --> 00:01:48,899
that we focus on structure on

34
00:01:53,450 --> 00:01:50,040
three-dimensional structure there is

35
00:01:55,130 --> 00:01:53,460
this axiom I guess you could say that

36
00:01:57,380 --> 00:01:55,140
structure is more conserved than

37
00:01:59,929 --> 00:01:57,390
sequence and if you want to look far far

38
00:02:02,389 --> 00:01:59,939

back in time you use structure instead

39
00:02:05,899 --> 00:02:02,399
of sequence so you can have especially

40
00:02:08,419 --> 00:02:05,909
with nucleic acids you can have you can

41
00:02:11,180 --> 00:02:08,429
have RNAs in which the the sequence is

42
00:02:13,320 --> 00:02:11,190
essentially scrambled between two RNAs

43
00:02:17,110 --> 00:02:13,330
and yet the structures are

44
00:02:18,790 --> 00:02:17,120
essentially it's in the experiment so so

45
00:02:20,410 --> 00:02:18,800
you can have variation in sequence but

46
00:02:22,839 --> 00:02:20,420
structure is conserved if you want to

47
00:02:28,020 --> 00:02:22,849
look far back in time you use structure

48
00:02:32,740 --> 00:02:30,790
okay so this is basically what we're

49
00:02:36,759 --> 00:02:32,750
after this is the Tree of Life we're

50
00:02:38,290 --> 00:02:36,769
interested in Luca and what was Luca

51
00:02:40,089 --> 00:02:38,300
we're sort of interested in that and

52
00:02:41,920 --> 00:02:40,099
then where did Luca come from basically

53
00:02:45,100 --> 00:02:41,930
this is the process and we're focused on

54
00:02:47,170 --> 00:02:45,110
the translation system and sort of for

55
00:02:49,360 --> 00:02:47,180
background you should know that the

56
00:02:52,660 --> 00:02:49,370
translation system was done at Luca so

57
00:02:54,430 --> 00:02:52,670
when you talk about the origin and

58
00:02:56,590 --> 00:02:54,440
evolution of the translation system

59
00:02:59,170 --> 00:02:56,600
you're really talking about things that

60
00:03:00,610 --> 00:02:59,180
happened before Luca so if we can if we

61
00:03:02,259 --> 00:03:00,620
can understand the origin of the

62
00:03:05,020 --> 00:03:02,269
translation system we are we are looking

63
00:03:06,580 --> 00:03:05,030

really far back in time you know 3.8 or

64
00:03:11,259 --> 00:03:06,590
4 billion years ago it depending on when

65
00:03:12,490 --> 00:03:11,269
life so of course this is what everybody

66
00:03:13,990 --> 00:03:12,500
does when they look at the origin of

67
00:03:17,970 --> 00:03:14,000
life and this is sort of various models

68
00:03:20,890 --> 00:03:17,980
people have there's RNA world clay world

69
00:03:22,990 --> 00:03:20,900
metabolism world vents and membranes

70
00:03:27,910 --> 00:03:23,000
which we've heard some about some of

71
00:03:29,289 --> 00:03:27,920
this at this meeting and I probably

72
00:03:31,300 --> 00:03:29,299
haven't really done justice to these

73
00:03:32,770 --> 00:03:31,310
cartoon representations but these are

74
00:03:33,940 --> 00:03:32,780
these are probably about all the models

75
00:03:38,229 --> 00:03:33,950
but this is kind of a survey of the

76
00:03:42,360 --> 00:03:38,239
model and then we have the data and this

77
00:03:45,160 --> 00:03:42,370
is the data that we focus on which is

78
00:03:47,770 --> 00:03:45,170
what I call the universal gene set

79
00:03:50,289 --> 00:03:47,780
things that everything alive has and

80
00:03:52,270 --> 00:03:50,299
it's actually there's several amazing

81
00:03:54,160 --> 00:03:52,280
characteristics of this universal gene

82
00:03:57,819 --> 00:03:54,170
set and this is this is a from

83
00:03:59,800 --> 00:03:57,829
Doolittle's paper in 2004 but pace has

84
00:04:02,140 --> 00:03:59,810
looked at this and Coonan and there is a

85
00:04:03,640 --> 00:04:02,150
nice consensus pretty much on what the

86
00:04:06,789 --> 00:04:03,650
universal gene set is and these are

87
00:04:09,099 --> 00:04:06,799
basically genes that you can find

88
00:04:11,470 --> 00:04:09,109

orthologues in everything alive so this

89
00:04:12,610 --> 00:04:11,480
is this is what you would say is

90
00:04:14,500 --> 00:04:12,620
universal biology

91
00:04:16,810 --> 00:04:14,510
everything has this and the first thing

92
00:04:18,580 --> 00:04:16,820
about it is it's a very small list it

93
00:04:20,379 --> 00:04:18,590
doesn't have very many genes on it

94
00:04:23,830 --> 00:04:20,389
depends on whose version you have but it

95
00:04:27,520 --> 00:04:23,840
has between 30 to 50 genes it's a small

96
00:04:29,110 --> 00:04:27,530
list and I've color-coded them here by

97
00:04:33,100 --> 00:04:29,120
what they what they're involved in and

98
00:04:34,990 --> 00:04:33,110
the the pink is translation and then we

99
00:04:37,210 --> 00:04:35,000
have replication transcription and

100
00:04:39,280 --> 00:04:37,220
replication and basically that the point

101
00:04:41,890 --> 00:04:39,290
of this thing is that this is dominated

102
00:04:44,500 --> 00:04:41,900
by translation so the universal gene set

103
00:04:46,600 --> 00:04:44,510
of biology is dominated by translation

104
00:04:49,510 --> 00:04:46,610
and this list actually does not have

105
00:04:51,730 --> 00:04:49,520
genes that encode RNAs so there are no

106
00:04:53,140 --> 00:04:51,740
genes four tRNAs on here and ribosomal

107
00:04:54,730 --> 00:04:53,150
RNAs and things like that if you

108
00:04:56,620 --> 00:04:54,740
actually included those on this list it

109
00:04:59,080 --> 00:04:56,630
would be a longer list and there would

110
00:05:04,510 --> 00:04:59,090
be more pink okay so the universal gene

111
00:05:07,750 --> 00:05:04,520
set of life is dominated by translation

112
00:05:11,110 --> 00:05:07,760
that's point of this and so you have

113
00:05:13,870 --> 00:05:11,120

this problem here that you have the data

114
00:05:15,730 --> 00:05:13,880
that sort of tells us about Luca and

115
00:05:18,130 --> 00:05:15,740
then you have these models and the

116
00:05:20,380 --> 00:05:18,140
models don't fit the data in the sense

117
00:05:22,590 --> 00:05:20,390
that the right let's say already world

118
00:05:25,240 --> 00:05:22,600
for example the RNA world predicts

119
00:05:26,920 --> 00:05:25,250
ribozymes that can replicate themselves

120
00:05:28,480 --> 00:05:26,930
and do metabolism and things like this

121
00:05:32,440 --> 00:05:28,490
and there are no there's nothing like

122
00:05:33,130 --> 00:05:32,450
that here there's no integral membrane

123
00:05:38,980 --> 00:05:33,140
proteins

124
00:05:41,140 --> 00:05:38,990
biosynthesis there's you know that these

125
00:05:46,390 --> 00:05:41,150
models are totally disconnected from

126
00:05:47,920 --> 00:05:46,400
this data and so what should you do if a

127
00:05:50,890 --> 00:05:47,930
model and you have data and they don't

128
00:05:54,520 --> 00:05:50,900
fit generally people would throw away

129
00:05:55,600 --> 00:05:54,530
the data that seems to be sort of what

130
00:05:57,700 --> 00:05:55,610
has happened in the origin of life

131
00:06:00,310 --> 00:05:57,710
people who study the origin of life

132
00:06:03,010 --> 00:06:00,320
basically ignore the universal gene set

133
00:06:07,090 --> 00:06:03,020
because it doesn't seem to relate to

134
00:06:11,320 --> 00:06:07,100
their models so when we look at this

135
00:06:13,510 --> 00:06:11,330
basically our idea is I just ignore all

136
00:06:16,000 --> 00:06:13,520
the models I just want to use the data

137
00:06:17,950 --> 00:06:16,010
and I want to walk the data back and say

138
00:06:20,080 --> 00:06:17,960

what does the data tell me about the

139
00:06:22,210 --> 00:06:20,090
origin of life I don't care about RNA

140
00:06:24,730 --> 00:06:22,220
world I don't care about metabolism

141
00:06:25,330 --> 00:06:24,740
world I just want to know what this

142
00:06:28,060 --> 00:06:25,340
means

143
00:06:31,300 --> 00:06:28,070
number one I want to know why is it so

144
00:06:36,339 --> 00:06:31,310
small why is it dominated by translation

145
00:06:38,110 --> 00:06:36,349
and can we use this to walk back and an

146
00:06:39,850 --> 00:06:38,120
important thing about this list is these

147
00:06:41,530 --> 00:06:39,860
are sequences right these are sequences

148
00:06:43,690 --> 00:06:41,540
that encode

149
00:06:45,610 --> 00:06:43,700
proteins but we know the structures of

150
00:06:47,050 --> 00:06:45,620
these proteins right crystallographers

151
00:06:49,150 --> 00:06:47,060
have determined drivers almost

152
00:06:51,340 --> 00:06:49,160
structures we know this is so so this

153
00:06:53,740 --> 00:06:51,350
isn't just a gene set of life this is

154
00:06:56,260 --> 00:06:53,750
really the sort of structure of life we

155
00:06:59,340 --> 00:06:56,270
know the universal three-dimensional

156
00:07:01,480 --> 00:06:59,350
structure of biology the conserved part

157
00:07:06,870 --> 00:07:01,490
another thing I'll point out about this

158
00:07:11,890 --> 00:07:09,460
you know people focus on things that are

159
00:07:13,990 --> 00:07:11,900
universal in biology but because there

160
00:07:16,150 --> 00:07:14,000
is this disconnect between the data and

161
00:07:18,520 --> 00:07:16,160
the models they tend to focus on small

162
00:07:20,880 --> 00:07:18,530
molecules so you'll hear people talk

163
00:07:23,020 --> 00:07:20,890

about metabolites and small molecules of

164
00:07:24,700 --> 00:07:23,030
biology that are highly conserved and

165
00:07:27,130 --> 00:07:24,710
they'll ascribe importance to these

166
00:07:29,560 --> 00:07:27,140
things but they never look at the

167
00:07:31,180 --> 00:07:29,570
macromolecules which are universal

168
00:07:33,130 --> 00:07:31,190
biology and the reason I think people

169
00:07:35,650 --> 00:07:33,140
don't do that is because this doesn't

170
00:07:38,940 --> 00:07:35,660
fit the models and it's very hard to

171
00:07:41,380 --> 00:07:38,950
take these universal macromolecules and

172
00:07:42,720 --> 00:07:41,390
fit them to these models they just

173
00:07:46,090 --> 00:07:42,730
they're just there's just no

174
00:07:49,360 --> 00:07:46,100
relationship between them okay so this

175
00:07:51,670 --> 00:07:49,370
is what we do we're looking at universal

176
00:07:53,140 --> 00:07:51,680
three-dimensional structures and we do

177
00:07:54,850 --> 00:07:53,150
that as I said because structures more

178
00:07:55,960 --> 00:07:54,860
conservative than sequence we want to

179
00:07:57,850 --> 00:07:55,970
look far back in time we use

180
00:07:59,410 --> 00:07:57,860
three-dimensional structures and we're

181
00:08:02,860 --> 00:07:59,420
asking you know what is the universal

182
00:08:04,300 --> 00:08:02,870
three-dimensional structure of life how

183
00:08:06,370 --> 00:08:04,310
conserved is it you know sequence is

184
00:08:08,380 --> 00:08:06,380
kind of digital essentially you could

185
00:08:09,400 --> 00:08:08,390
say this is a C this is a G but when

186
00:08:10,720 --> 00:08:09,410
you're talking about structure you know

187
00:08:13,600 --> 00:08:10,730
you're talking about something that is

188
00:08:15,400 --> 00:08:13,610

continuous you know you have RMS ease of

189
00:08:17,140 --> 00:08:15,410
atoms or something and it's and it's a

190
00:08:19,510 --> 00:08:17,150
little bit it's actually significantly

191
00:08:21,910 --> 00:08:19,520
more different difficult to sort of be

192
00:08:23,680 --> 00:08:21,920
analytical and quantitative about

193
00:08:26,110 --> 00:08:23,690
structural conservation than it is about

194
00:08:27,580 --> 00:08:26,120
sequence conservation so we have to sort

195
00:08:31,780 --> 00:08:27,590
of work and figure out how to deal with

196
00:08:33,670 --> 00:08:31,790
that we want to know how structure is

197
00:08:35,440 --> 00:08:33,680
elaborated like we have things that are

198
00:08:38,500 --> 00:08:35,450
conserved but then the part that varies

199
00:08:39,790 --> 00:08:38,510
what are the rules of that and then what

200
00:08:41,260 --> 00:08:39,800
are the origins of these conserved

201
00:08:45,100 --> 00:08:41,270
structures that's really that's really

202
00:08:47,560 --> 00:08:45,110
what we're after so what I what I'm sure

203
00:08:49,660 --> 00:08:47,570
you hear are secondary structures of

204
00:08:52,050 --> 00:08:49,670
ribosomal RNAs first thing these are

205
00:08:54,880 --> 00:08:52,060
very large molecules they're huge and

206
00:08:58,290 --> 00:08:54,890
this is e coli yeast

207
00:09:01,269 --> 00:08:58,300
I Inhumans and I want to first focus on

208
00:09:02,590 --> 00:09:01,279
let's just look at this little that

209
00:09:04,420 --> 00:09:02,600
little purple thing right there that

210
00:09:08,920 --> 00:09:04,430
little knob and what I want you to

211
00:09:12,460 --> 00:09:08,930
notice is B coli has it yeast has it

212
00:09:14,920 --> 00:09:12,470
fruit fly has it almost sapiens have it

213
00:09:17,949 --> 00:09:14,930

everything everything in biology has

214
00:09:20,410 --> 00:09:17,959
that okay and we have sequences of many

215
00:09:21,759 --> 00:09:20,420
ribosomal I mean the data is very rich

216
00:09:24,040 --> 00:09:21,769
here and we have three-dimensional

217
00:09:26,220 --> 00:09:24,050
structures and we can say that

218
00:09:31,269 --> 00:09:26,230
everything in biology has that little

219
00:09:32,470 --> 00:09:31,279
knob right there and but what really we

220
00:09:34,060 --> 00:09:32,480
want to look at is the three-dimensional

221
00:09:36,460 --> 00:09:34,070
structures so what I want to do is I

222
00:09:38,230 --> 00:09:36,470
want to show you kind of what what we

223
00:09:39,880 --> 00:09:38,240
mean when we say the universal structure

224
00:09:42,790 --> 00:09:39,890
of biology so I'm just going to take a

225
00:09:43,870 --> 00:09:42,800
part of this thing we have crystal

226
00:09:54,360 --> 00:09:43,880
structures of all these and show you

227
00:10:01,810 --> 00:09:58,650
okay I fixed it so this is this is a

228
00:10:03,880 --> 00:10:01,820
part of the ribosomal RNA from ecoli and

229
00:10:08,019 --> 00:10:03,890
I've stripped out all the bases it's

230
00:10:08,650 --> 00:10:08,029
just a trace of the backbone and it's

231
00:10:12,310 --> 00:10:08,660
beautiful

232
00:10:14,310 --> 00:10:12,320
okay if you look at and this is a global

233
00:10:16,689 --> 00:10:14,320
superimposition of ribosomes

234
00:10:18,280 --> 00:10:16,699
this is archaea okay so these are two

235
00:10:22,000 --> 00:10:18,290
separate branches of the phylogenetic

236
00:10:23,710 --> 00:10:22,010
tree and you can see that these look

237
00:10:25,689 --> 00:10:23,720
pretty similar now the interesting thing

238
00:10:27,819 --> 00:10:25,699

is the sequences are not necessarily the

239
00:10:30,939 --> 00:10:27,829
same right like this is a helix here and

240
00:10:33,460 --> 00:10:30,949
you helix can accommodate many sequences

241
00:10:36,069 --> 00:10:33,470
right so the sequences are not fully

242
00:10:37,480 --> 00:10:36,079
conserved but the structure is so these

243
00:10:39,910 --> 00:10:37,490
are two branches of the Tree of Life

244
00:10:43,600 --> 00:10:39,920
I can look at the other one here's yeast

245
00:10:45,850 --> 00:10:43,610
and basically this is it this is the

246
00:10:49,030 --> 00:10:45,860
diversity in all of biology that you're

247
00:10:51,430 --> 00:10:49,040
looking at this is this is how much this

248
00:10:53,710 --> 00:10:51,440
thing changes and I used to say this

249
00:10:55,810 --> 00:10:53,720
thing is how much it changes over four

250
00:11:00,699 --> 00:10:55,820
billion years of evolution but Greg

251
00:11:03,250 --> 00:11:00,709
Fournier pointed out that I need to

252
00:11:06,189 --> 00:11:03,260
double that because the distance from

253
00:11:08,139 --> 00:11:06,199
Luca to e.coli is four billion years the

254
00:11:10,569 --> 00:11:08,149
distance from Luca to

255
00:11:14,350 --> 00:11:10,579
hey lark EULA is 4 billion years so this

256
00:11:16,509 --> 00:11:14,360
is 8 billion years of evolution and this

257
00:11:20,379 --> 00:11:16,519
is showing you how much the structure

258
00:11:24,009 --> 00:11:20,389
changes and then just to kind of top it

259
00:11:26,170 --> 00:11:24,019
off you know evolution sort of is is on

260
00:11:28,059 --> 00:11:26,180
steroids in organelles and things are

261
00:11:30,809 --> 00:11:28,069
happening much faster in organelles in

262
00:11:34,179 --> 00:11:30,819
in in cytoplasm and so here is the

263
00:11:38,079 --> 00:11:34,189

mitochondrial ribosome from Homo sapiens

264
00:11:40,329 --> 00:11:38,089
and this you know basically this is just

265
00:11:41,619 --> 00:11:40,339
cemented this is okay there's some

266
00:11:45,309 --> 00:11:41,629
geologists here so I'm going to say

267
00:11:49,059 --> 00:11:45,319
something incorrect me this is the most

268
00:11:51,549 --> 00:11:49,069
permanent thing that is not cold in the

269
00:11:55,420 --> 00:11:51,559
universe I don't know if that's true but

270
00:11:58,900 --> 00:11:55,430
it sounds good you can it's but it

271
00:12:01,780 --> 00:11:58,910
really is very very constant okay so

272
00:12:04,509 --> 00:12:01,790
what we have done is this is just part

273
00:12:06,189 --> 00:12:04,519
of the ribosome we have looked at the

274
00:12:08,230 --> 00:12:06,199
entire ribosome and we've defined what

275
00:12:11,650 --> 00:12:08,240
we call the common core so this is the

276
00:12:13,629 --> 00:12:11,660
bacterial this is e.coli ribosome this

277
00:12:16,299 --> 00:12:13,639
is large subunit RNA and small subunit

278
00:12:19,119 --> 00:12:16,309
and the blue parts are the part that are

279
00:12:22,809 --> 00:12:19,129
structurally conserved in everything

280
00:12:24,730 --> 00:12:22,819
alive okay and we've Chad Bernier is a

281
00:12:27,009 --> 00:12:24,740
graduate student who actually graduated

282
00:12:28,929 --> 00:12:27,019
but he spent years this this to work

283
00:12:30,400 --> 00:12:28,939
this out involved superimposition

284
00:12:33,699 --> 00:12:30,410
alignment it was a sort of iterative

285
00:12:36,340 --> 00:12:33,709
process of figuring out what's what and

286
00:12:37,869 --> 00:12:36,350
and so like for example this is this

287
00:12:39,759 --> 00:12:37,879
little blue bump I told you that

288
00:12:42,160 --> 00:12:39,769

everything has it's right there and

289
00:12:44,350 --> 00:12:42,170
right there but then like look at it

290
00:12:46,960 --> 00:12:44,360
let's focus on this little arm here see

291
00:12:49,749 --> 00:12:46,970
e.coli has an arm there and halo ocula

292
00:12:51,910 --> 00:12:49,759
doesn't so that's black so the things

293
00:12:53,439 --> 00:12:51,920
that are common are blue and and really

294
00:12:56,439 --> 00:12:53,449
the point of this slide is that to a

295
00:13:00,009 --> 00:12:56,449
first approximation the e.coli ribosome

296
00:13:03,100 --> 00:13:00,019
is the universal core everything in

297
00:13:06,400 --> 00:13:03,110
biology has something equivalent to the

298
00:13:08,019 --> 00:13:06,410
the bacterial ribosome okay and you can

299
00:13:10,030 --> 00:13:08,029
see it's not exactly true because the

300
00:13:11,610 --> 00:13:10,040
black parts are or where that breaks

301
00:13:15,480 --> 00:13:11,620
down but to a first approximation

302
00:13:21,290 --> 00:13:15,490
everything alive has a ecoli ribosome

303
00:13:28,790 --> 00:13:24,380
all around this tree that's that's 8

304
00:13:30,199 --> 00:13:28,800
billion years of evolution ok so but now

305
00:13:34,940 --> 00:13:30,209
I want to focus on the parts that are

306
00:13:36,620 --> 00:13:34,950
different because you can look here like

307
00:13:38,750 --> 00:13:36,630
look at the you the human ribosome has

308
00:13:41,540 --> 00:13:38,760
these devil horns right look at those

309
00:13:43,130 --> 00:13:41,550
those devil horns and of course bacteria

310
00:13:44,960 --> 00:13:43,140
don't have that and so there are places

311
00:13:47,110 --> 00:13:44,970
of the ribosome varies and what I want

312
00:13:49,250 --> 00:13:47,120
to focus on that for a minute because

313
00:13:51,949 --> 00:13:49,260

what I want to do is figure out how the

314
00:13:53,990 --> 00:13:51,959
ribosome changes over time and it turns

315
00:13:56,750 --> 00:13:54,000
out that the rules are very strict and

316
00:13:59,090 --> 00:13:56,760
there are very specific things that can

317
00:14:01,430 --> 00:13:59,100
happen to the ribosome you know like

318
00:14:02,690 --> 00:14:01,440
nothing can happen there that's just not

319
00:14:05,509 --> 00:14:02,700
allowed ok

320
00:14:06,769 --> 00:14:05,519
but this helix right here that's a lot

321
00:14:12,259 --> 00:14:06,779
of things are allowed there in fact they

322
00:14:13,880 --> 00:14:12,269
have well okay let me just really what

323
00:14:15,530 --> 00:14:13,890
what you have is you have something

324
00:14:17,840 --> 00:14:15,540
called the universal common core and

325
00:14:20,210 --> 00:14:17,850
that's the E coli ribosome basically and

326
00:14:23,509 --> 00:14:20,220
then you have a eukaryotic shell and

327
00:14:27,019 --> 00:14:23,519
those are these sorts of things and then

328
00:14:29,269 --> 00:14:27,029
in metazoans especially in mammals you

329
00:14:34,090 --> 00:14:29,279
have these tentacles that go out so

330
00:14:36,470 --> 00:14:34,100
really there are I'd say an extant

331
00:14:38,540 --> 00:14:36,480
ribosomes there are these sort of three

332
00:14:39,980 --> 00:14:38,550
phases there's the common core that

333
00:14:42,230 --> 00:14:39,990
everything happens then there's a

334
00:14:45,290 --> 00:14:42,240
eukaryotic shell the first shell and

335
00:14:47,210 --> 00:14:45,300
then there are these metazoan expansion

336
00:14:49,610 --> 00:14:47,220
segments so this is this is the large

337
00:14:51,740 --> 00:14:49,620
subunit we're looking at this sort of

338
00:14:54,019 --> 00:14:51,750

summarizes the evolution of the ribosome

339
00:14:56,240 --> 00:14:54,029
and we know this because we have

340
00:14:59,120 --> 00:14:56,250
examples all over the phylogenetic tree

341
00:15:03,199 --> 00:14:59,130
of these things okay now I want to focus

342
00:15:05,030 --> 00:15:03,209
on this helix this is called helix 25 an

343
00:15:07,400 --> 00:15:05,040
e.coli but it's called expansion segment

344
00:15:08,960 --> 00:15:07,410
seven and eukaryotes because it's no

345
00:15:11,269 --> 00:15:08,970
longer a helix once you get that you

346
00:15:13,490 --> 00:15:11,279
eukaryotes it gets really enormous in

347
00:15:16,100 --> 00:15:13,500
fact look at human it has these

348
00:15:19,730 --> 00:15:16,110
tentacles humans has has really long

349
00:15:21,560 --> 00:15:19,740
tentacles the ribosome it's kind of like

350
00:15:23,960 --> 00:15:21,570
kudzu if you actually cuz he was a

351
00:15:26,329 --> 00:15:23,970
Japanese vine that was imported in the

352
00:15:30,170 --> 00:15:26,339
United States and it grows like a foot a

353
00:15:32,949 --> 00:15:30,180
day and the eukaryotic ribosomes kind of

354
00:15:35,030 --> 00:15:32,959
reminds me of kudzu these arms are

355
00:15:36,560 --> 00:15:35,040
changing very rapidly over

356
00:15:40,250 --> 00:15:36,570
evolution and they seem to be getting

357
00:15:41,420 --> 00:15:40,260
longer and longer but what I want to do

358
00:15:42,860 --> 00:15:41,430
is I want to look at these in three

359
00:15:44,360 --> 00:15:42,870
dimensions we know how this grows

360
00:15:45,410 --> 00:15:44,370
everything I'm showing you here is in

361
00:15:47,420 --> 00:15:45,420
two dimensions but we know what this

362
00:15:49,880 --> 00:15:47,430
looks like in three dimensions and so

363
00:15:52,970 --> 00:15:49,890

this is what we call the vishnu basement

364
00:15:54,650 --> 00:15:52,980
this is helix 25 this is the bottom this

365
00:15:58,250 --> 00:15:54,660
is there everywhere it never goes away

366
00:16:00,530 --> 00:15:58,260
and and that's we know that what that is

367
00:16:02,450 --> 00:16:00,540
from bacterial ribosomes and we have

368
00:16:04,240 --> 00:16:02,460
ecoli and thermus thermophilus we have

369
00:16:06,710 --> 00:16:04,250
crystal structures of quite a few

370
00:16:10,010 --> 00:16:06,720
bacterial ribosomes and then when you

371
00:16:12,950 --> 00:16:10,020
look at archaea you can see helix 25 is

372
00:16:17,360 --> 00:16:12,960
still there and then something has grown

373
00:16:19,820 --> 00:16:17,370
out so this allows us to have an

374
00:16:21,800 --> 00:16:19,830
estimate of Luca right because we say

375
00:16:24,050 --> 00:16:21,810
what is common between archaea and

376
00:16:27,620 --> 00:16:24,060
bacteria is our best guess for Luca and

377
00:16:29,390 --> 00:16:27,630
so we say Luke I had helix 25 because

378
00:16:32,680 --> 00:16:29,400
that's what's common between archaea and

379
00:16:34,760 --> 00:16:32,690
bacteria that's sort of a standard way

380
00:16:36,680 --> 00:16:34,770
okay so but we can just keep walking

381
00:16:38,690 --> 00:16:36,690
here here's archaea

382
00:16:41,990 --> 00:16:38,700
here's yeast and what you can see is

383
00:16:43,520 --> 00:16:42,000
that the archaea element is essentially

384
00:16:44,660 --> 00:16:43,530
in yeast and then things have grown out

385
00:16:48,200 --> 00:16:44,670
of it

386
00:16:49,550 --> 00:16:48,210
here's fruit fly and fruit fly looks I

387
00:16:51,190 --> 00:16:49,560
mean really there's common ancestors

388
00:16:54,050 --> 00:16:51,200

here I don't want to give the idea that

389
00:16:56,090 --> 00:16:54,060
fruit fly ball from yeast but that but

390
00:16:58,640 --> 00:16:56,100
essentially the common ancestor looks

391
00:17:01,520 --> 00:16:58,650
like east and then things have grown out

392
00:17:02,960 --> 00:17:01,530
of it and then it's not nice you have a

393
00:17:04,640 --> 00:17:02,970
lot of crystal structures and cryo-em

394
00:17:06,710 --> 00:17:04,650
structures and then you can see the same

395
00:17:08,240 --> 00:17:06,720
thing so that so yes this is really

396
00:17:09,980 --> 00:17:08,250
important to us when we made this

397
00:17:11,870 --> 00:17:09,990
discovery that this is how the ribosome

398
00:17:13,970 --> 00:17:11,880
changes because this is a process called

399
00:17:15,800 --> 00:17:13,980
accretion it's just like geology

400
00:17:18,500 --> 00:17:15,810
it's just the way trees grow at the

401
00:17:21,650 --> 00:17:18,510
ribosome accretes and when of the creats

402
00:17:23,270 --> 00:17:21,660
like this helix 25 it's always there

403
00:17:25,280 --> 00:17:23,280
right when you add things onto the

404
00:17:26,660 --> 00:17:25,290
ribosome you don't mess with the

405
00:17:29,270 --> 00:17:26,670
underlying structure that's what

406
00:17:31,430 --> 00:17:29,280
evolution does and we have lots of

407
00:17:32,810 --> 00:17:31,440
examples over this of this over the I

408
00:17:34,430 --> 00:17:32,820
don't want to take time here but you can

409
00:17:36,320 --> 00:17:34,440
just sort of see it it looks like a

410
00:17:39,410 --> 00:17:36,330
movie of things going okay so the

411
00:17:41,930 --> 00:17:39,420
ribosome the ribosomal RNA I need to be

412
00:17:44,810 --> 00:17:41,940
careful the ribosomal RNA grows by

413
00:17:48,650 --> 00:17:44,820

accretion okay and when it does that

414
00:17:49,070 --> 00:17:48,660
when it when it increases in size it

415
00:17:50,750 --> 00:17:49,080
leaves

416
00:17:52,850 --> 00:17:50,760
fingerprints it's very nice it's like a

417
00:17:54,740 --> 00:17:52,860
tree and when you have a branch going on

418
00:17:56,389 --> 00:17:54,750
a tree there's a knot and so even if you

419
00:17:57,590 --> 00:17:56,399
cut the branch off you can cut into the

420
00:17:59,810 --> 00:17:57,600
tree and you could say yes there was a

421
00:18:01,490 --> 00:17:59,820
branch here once and so this is

422
00:18:04,159 --> 00:18:01,500
something we've identified called an

423
00:18:07,549 --> 00:18:04,169
insertion fingerprint that it's a

424
00:18:11,029 --> 00:18:07,559
specific structural element that allows

425
00:18:12,950 --> 00:18:11,039
us to say this green thing grew out of

426
00:18:14,630 --> 00:18:12,960
that red thing and we can do that

427
00:18:16,730 --> 00:18:14,640
because we have these structures of I

428
00:18:19,700 --> 00:18:16,740
mean we know that's true because this is

429
00:18:21,909 --> 00:18:19,710
answer the the the blue is the

430
00:18:24,730 --> 00:18:21,919
prokaryotic ribosomes and this is the

431
00:18:27,169 --> 00:18:24,740
eukaryotic ribosomes and we know that

432
00:18:29,659 --> 00:18:27,179
that the red was essentially that the

433
00:18:31,789 --> 00:18:29,669
blue was the ancestor of the red and

434
00:18:33,710 --> 00:18:31,799
green so you know this is not modeling

435
00:18:35,450 --> 00:18:33,720
this is structures and and we really

436
00:18:37,070 --> 00:18:35,460
know how the ribosome changes over time

437
00:18:39,470 --> 00:18:37,080
so we have two things we have the

438
00:18:41,269 --> 00:18:39,480

ribosome grows by accretion and when

439
00:18:43,070 --> 00:18:41,279
there are growth steps we don't know

440
00:18:45,950 --> 00:18:43,080
that all of them but many of the growth

441
00:18:48,230 --> 00:18:45,960
steps leave fingerprints so that even if

442
00:18:49,549 --> 00:18:48,240
you only have the human ribosome you

443
00:18:51,620 --> 00:18:49,559
could look and say there was a growth

444
00:18:52,940 --> 00:18:51,630
event here but we don't have to do that

445
00:18:55,940 --> 00:18:52,950
because there have many many ribosomes

446
00:18:58,039 --> 00:18:55,950
of all different sizes and shapes okay

447
00:19:00,440 --> 00:18:58,049
so now I want to use this tree analogy

448
00:19:03,200 --> 00:19:00,450
because the ribosome is very similar to

449
00:19:05,509 --> 00:19:03,210
a tree in that a tree records its

450
00:19:07,039 --> 00:19:05,519
history and it does that because it

451
00:19:08,690 --> 00:19:07,049
grows by accretion if the tree died out

452
00:19:10,759 --> 00:19:08,700
every year and regrew then next year

453
00:19:13,009 --> 00:19:10,769
then it wouldn't record its history

454
00:19:16,370 --> 00:19:13,019
right the reason a tree records history

455
00:19:18,409 --> 00:19:16,380
is that something happens and then that

456
00:19:19,639 --> 00:19:18,419
doesn't change you you add layers and

457
00:19:21,440 --> 00:19:19,649
layers you had twigs and branches and

458
00:19:23,389 --> 00:19:21,450
you you don't perturb the underlying

459
00:19:26,269 --> 00:19:23,399
thing when you have growth processes and

460
00:19:28,129 --> 00:19:26,279
so maybe for from Arizona and you

461
00:19:31,039 --> 00:19:28,139
haven't seen a tree before this wouldn't

462
00:19:34,190 --> 00:19:31,049
make sense to you but for the rest of us

463
00:19:35,769 --> 00:19:34,200

we can look at this tree and we can you

464
00:19:38,240 --> 00:19:35,779
could at very high level of detail

465
00:19:40,430 --> 00:19:38,250
ascribe relative age to everything on

466
00:19:43,370 --> 00:19:40,440
this tree right you could say the oldest

467
00:19:45,620 --> 00:19:43,380
part of this tree is in the center of

468
00:19:48,310 --> 00:19:45,630
that trunk the leaves are the most

469
00:19:50,870 --> 00:19:48,320
recent that small tricked twigs are are

470
00:19:54,830 --> 00:19:50,880
younger than the big twigs etc right and

471
00:19:56,990 --> 00:19:54,840
you can look at this tree and and sort

472
00:19:59,120 --> 00:19:57,000
of wat work out its history and it turns

473
00:20:00,740 --> 00:19:59,130
out I'm going to explain it a little bit

474
00:20:02,870 --> 00:20:00,750
to you but the ribosome is the same

475
00:20:05,690 --> 00:20:02,880
thing okay because it grew by a

476
00:20:08,360 --> 00:20:05,700
and because it leaves telltale signs

477
00:20:09,920 --> 00:20:08,370
when it undergoes a growth event you can

478
00:20:12,110 --> 00:20:09,930
look at the ribosome and you can read

479
00:20:14,600 --> 00:20:12,120
out its history just the same way you

480
00:20:15,800 --> 00:20:14,610
can read out the history of this tree so

481
00:20:18,830 --> 00:20:15,810
what this really means is there is a

482
00:20:21,380 --> 00:20:18,840
history of biology before Luca because

483
00:20:23,630 --> 00:20:21,390
the ribosome was done at Luca and we

484
00:20:26,420 --> 00:20:23,640
have ways of reading out the history of

485
00:20:28,370 --> 00:20:26,430
that ribosome so this is just to be

486
00:20:30,950 --> 00:20:28,380
really clear these are observations the

487
00:20:32,870 --> 00:20:30,960
modern ribosome it grew and it is

488
00:20:34,550 --> 00:20:32,880

growing by accretion and we have we know

489
00:20:36,620 --> 00:20:34,560
that because we have crystal structures

490
00:20:38,270 --> 00:20:36,630
of so many ribosomes across the

491
00:20:43,700 --> 00:20:38,280
phylogenetic tree so that is what we

492
00:20:45,530 --> 00:20:43,710
call an observation we we know that when

493
00:20:46,610 --> 00:20:45,540
growth events occur some of them not all

494
00:20:48,710 --> 00:20:46,620
of them but some of them leave

495
00:20:51,560 --> 00:20:48,720
fingerprints okay so this is what we

496
00:20:54,160 --> 00:20:51,570
know and then these are in inferences so

497
00:20:56,900 --> 00:20:54,170
we are assuming that the common core

498
00:20:59,090 --> 00:20:56,910
also grew by accretion now we don't know

499
00:21:01,130 --> 00:20:59,100
that because we can't watch it the way

500
00:21:03,560 --> 00:21:01,140
we can watch eukaryotic ribosomes but

501
00:21:05,480 --> 00:21:03,570
we're just assuming the process is

502
00:21:07,880 --> 00:21:05,490
continuous and we're also assuming that

503
00:21:09,800 --> 00:21:07,890
in the common core that growth fence

504
00:21:11,810 --> 00:21:09,810
left fingerprints so we're just assuming

505
00:21:14,420 --> 00:21:11,820
a kind of continuity if our assumption

506
00:21:15,410 --> 00:21:14,430
is wrong then it's wrong but I just want

507
00:21:18,440 --> 00:21:15,420
to be clear about what our assumptions

508
00:21:20,390 --> 00:21:18,450
are so based on that we can look at the

509
00:21:23,030 --> 00:21:20,400
common core so this is the thing that

510
00:21:25,670 --> 00:21:23,040
everything in biology has and we can

511
00:21:28,550 --> 00:21:25,680
work out how it grew so just I'll just

512
00:21:30,290 --> 00:21:28,560
focus on right here every time the color

513
00:21:33,110 --> 00:21:30,300

changes there is an insertion

514
00:21:35,000 --> 00:21:33,120
fingerprint here okay so you can say

515
00:21:36,950 --> 00:21:35,010
something there was a growth event that

516
00:21:39,740 --> 00:21:36,960
green and the red got out of that green

517
00:21:41,360 --> 00:21:39,750
got outed etc so we can see these

518
00:21:44,210 --> 00:21:41,370
insertion fingerprints and we're

519
00:21:46,160 --> 00:21:44,220
assuming that it grew by accretion so we

520
00:21:49,310 --> 00:21:46,170
can work out in a very high level of

521
00:21:53,600 --> 00:21:49,320
detail the steps in the growth of the

522
00:21:55,400 --> 00:21:53,610
common core and so see this red thing

523
00:21:57,050 --> 00:21:55,410
right there that that's actually the

524
00:21:59,440 --> 00:21:57,060
secondary structure is misleading that's

525
00:22:02,180 --> 00:21:59,450
a continuous helix this is your mother

526
00:22:03,710 --> 00:22:02,190
okay that is the beginning of the

527
00:22:06,980 --> 00:22:03,720
ribosome that is the first thing that

528
00:22:08,930 --> 00:22:06,990
happened and then there's a series of

529
00:22:10,550 --> 00:22:08,940
growth events that ultimately leave this

530
00:22:13,340 --> 00:22:10,560
but we can work out in a very high level

531
00:22:15,610 --> 00:22:13,350
of detail how the ribosome Grill in fact

532
00:22:16,850 --> 00:22:15,620
it's so much detail that it was just

533
00:22:19,700 --> 00:22:16,860
mind-boggling

534
00:22:21,470 --> 00:22:19,710
and we just kind of left it for a while

535
00:22:23,690 --> 00:22:21,480
and then we said okay let's simplify it

536
00:22:25,130 --> 00:22:23,700
so we group these things into phases we

537
00:22:26,600 --> 00:22:25,140
just kind of did a coarse graining of

538
00:22:29,120 --> 00:22:26,610

this to cut down on the amount of

539
00:22:31,400 --> 00:22:29,130
information so really we have these

540
00:22:33,320 --> 00:22:31,410
steps one two three four five we just

541
00:22:35,840 --> 00:22:33,330
call that we just sort of group them by

542
00:22:38,120 --> 00:22:35,850
color and it was a little you know we

543
00:22:39,980 --> 00:22:38,130
could have so we have we have five

544
00:22:41,600 --> 00:22:39,990
phases here we could have done four or

545
00:22:44,750 --> 00:22:41,610
we could have done six it was that's a

546
00:22:46,220 --> 00:22:44,760
bit art that's arbitrary but we just had

547
00:22:47,750 --> 00:22:46,230
to kind of cut down on the amount of

548
00:22:50,539 --> 00:22:47,760
detail because it was just more than we

549
00:22:52,789 --> 00:22:50,549
could deal with so but so this is sort

550
00:22:54,919 --> 00:22:52,799
of a more code so the blue is the first

551
00:22:59,390 --> 00:22:54,929
thing that happened dark blue light blue

552
00:23:01,310 --> 00:22:59,400
green then all those colors okay and we

553
00:23:03,680 --> 00:23:01,320
have large subunit small subunit and we

554
00:23:04,789 --> 00:23:03,690
have ways of correlating this is public

555
00:23:06,350 --> 00:23:04,799
so I don't to go into it but we have

556
00:23:08,620 --> 00:23:06,360
ways of correlating between the two

557
00:23:11,450 --> 00:23:08,630
subunits so we have a very very detailed

558
00:23:14,510 --> 00:23:11,460
model for the evolution of the ribosome

559
00:23:15,860 --> 00:23:14,520
and in our lab what we do is we make

560
00:23:18,500 --> 00:23:15,870
these things I mean we we have an

561
00:23:20,240 --> 00:23:18,510
experimental lab we we have predictions

562
00:23:22,070 --> 00:23:20,250
on what these molecules should do we can

563
00:23:24,350 --> 00:23:22,080

make any piece of RNA and we just make

564
00:23:25,909 --> 00:23:24,360
them and test the predictions and so

565
00:23:30,169 --> 00:23:25,919
that's really one of the main focuses of

566
00:23:32,150 --> 00:23:30,179
my lab and and this is the nice thing

567
00:23:34,220 --> 00:23:32,160
about RNA is that it tends to be modular

568
00:23:36,440 --> 00:23:34,230
and so because we have this structural

569
00:23:38,480 --> 00:23:36,450
map of what has happened we also have a

570
00:23:40,669 --> 00:23:38,490
functional map okay because the

571
00:23:42,590 --> 00:23:40,679
structure and the function are so easy

572
00:23:44,630 --> 00:23:42,600
to decipher so we have a very detailed

573
00:23:46,940 --> 00:23:44,640
map and all of these you know this is a

574
00:23:49,100 --> 00:23:46,950
hypothesis and we can make all of this

575
00:23:50,419 --> 00:23:49,110
easily we can we can make each one of

576
00:23:51,740 --> 00:23:50,429
these molecules we can test the

577
00:23:53,900 --> 00:23:51,750
predictions about what it should be

578
00:23:56,450 --> 00:23:53,910
doing so this is all experimental II

579
00:23:57,980 --> 00:23:56,460
grounded okay but I want to change

580
00:24:00,350 --> 00:23:57,990
subjects totally cuz so far I've talked

581
00:24:02,090 --> 00:24:00,360
to only about RNA and we've just talked

582
00:24:02,900 --> 00:24:02,100
about the changes of RNA how much time

583
00:24:07,970 --> 00:24:02,910
do I have Petula

584
00:24:11,419 --> 00:24:10,909
I'm fine okay cuz I have 47 more slides

585
00:24:18,020 --> 00:24:11,429
good

586
00:24:19,880 --> 00:24:18,030
she said I'm fine I don't really you

587
00:24:21,260 --> 00:24:19,890
guys so basically everything I've talked

588
00:24:22,789 --> 00:24:21,270

about so far is published but this is

589
00:24:25,010 --> 00:24:22,799
new and this is what I'm really

590
00:24:26,330 --> 00:24:25,020
interested in is the proteins I haven't

591
00:24:29,659 --> 00:24:26,340
talked about proteins yet we've only

592
00:24:31,159 --> 00:24:29,669
talked about RNA so what this is this is

593
00:24:33,140 --> 00:24:31,169
the three-dimensional structure of the

594
00:24:36,320 --> 00:24:33,150
bacterial ribosome and we have color

595
00:24:38,390 --> 00:24:36,330
coded the RNA by those phases so the

596
00:24:40,340 --> 00:24:38,400
oldest thing here is dark blue then

597
00:24:43,970 --> 00:24:40,350
light blue and green cetera okay so that

598
00:24:47,000 --> 00:24:43,980
the RNA is colored by its it's relative

599
00:24:51,530 --> 00:24:47,010
age and then the grey things here are

600
00:24:54,560 --> 00:24:51,540
the ribosomal proteins and what we have

601
00:24:56,240 --> 00:24:54,570
done is we have used the RNA to date the

602
00:25:00,770 --> 00:24:56,250
proteins and we just made kind of a

603
00:25:03,080 --> 00:25:00,780
simple assumption that that if a part of

604
00:25:07,070 --> 00:25:03,090
protein is surrounded by RNA then it's

605
00:25:11,930 --> 00:25:07,080
of the same age and and this is kind of

606
00:25:14,150 --> 00:25:11,940
a test of our model because I'll sort of

607
00:25:15,799 --> 00:25:14,160
show you why but if we can date the

608
00:25:17,570 --> 00:25:15,809
ribosomal RNA and we know what we're

609
00:25:24,360 --> 00:25:17,580
doing there then we should be able to

610
00:25:30,990 --> 00:25:29,370
okay all right okay so so what I have

611
00:25:33,240 --> 00:25:31,000
over here is I'm showing the ribosomal

612
00:25:34,830 --> 00:25:33,250
proteins and I've faded out the RNA and

613
00:25:38,250 --> 00:25:34,840

you can see the Rebels normal proteins

614
00:25:42,210 --> 00:25:38,260
are they're unique in in all of biology

615
00:25:45,840 --> 00:25:42,220
in they have these they're not random I

616
00:25:47,669 --> 00:25:45,850
mean they're they're not random coil

617
00:25:50,490 --> 00:25:47,679
because random coil is in an ensemble

618
00:25:53,730 --> 00:25:50,500
these are in specific States but they're

619
00:25:56,000 --> 00:25:53,740
non canonical so that's a it's a they're

620
00:25:58,440 --> 00:25:56,010
frozen they don't move but they're not

621
00:26:01,799 --> 00:25:58,450
folded in normal ways and you just don't

622
00:26:04,529 --> 00:26:01,809
see proteins doing this in biology and

623
00:26:06,720 --> 00:26:04,539
and what we've been able to do is we we

624
00:26:10,710 --> 00:26:06,730
take a given protein like this one and

625
00:26:13,590 --> 00:26:10,720
we cut it into segments and we cut it

626
00:26:15,899 --> 00:26:13,600
based on the age of the RNA around it so

627
00:26:18,269 --> 00:26:15,909
we can we basically say this part in

628
00:26:20,100 --> 00:26:18,279
fact this this color is one age that's

629
00:26:22,230 --> 00:26:20,110
another age that's at another age and we

630
00:26:24,630 --> 00:26:22,240
do that by the RNA that is surrounding

631
00:26:26,250 --> 00:26:24,640
it okay and that turns out to be pretty

632
00:26:27,840 --> 00:26:26,260
easy and clean there's some places where

633
00:26:32,700 --> 00:26:27,850
it's ambiguous you know but it's a

634
00:26:34,440 --> 00:26:32,710
pretty clean way so so we and there's a

635
00:26:36,060 --> 00:26:34,450
lot of proteins in the ribosome so this

636
00:26:38,010 --> 00:26:36,070
is a lot of data but this is kind of

637
00:26:40,980 --> 00:26:38,020
what it looks like is like so this is a

638
00:26:43,409 --> 00:26:40,990

ribosomal protein and this red part is

639
00:26:46,049 --> 00:26:43,419
the oldest and then this yellow part is

640
00:26:49,919 --> 00:26:46,059
younger and that is younger okay so we

641
00:26:52,740 --> 00:26:49,929
can segment these proteins by age and so

642
00:26:53,570 --> 00:26:52,750
these segments here are all of the same

643
00:26:58,820 --> 00:26:53,580
age

644
00:27:02,669 --> 00:26:58,830
time is moving in this direction right

645
00:27:05,310 --> 00:27:02,679
so and we didn't build this in we didn't

646
00:27:07,169 --> 00:27:05,320
sort of know it's going to be like this

647
00:27:09,990 --> 00:27:07,179
but this is basically a reaction

648
00:27:12,299 --> 00:27:10,000
coordinate of protein folding okay what

649
00:27:14,460 --> 00:27:12,309
we have here is frozen random coil I

650
00:27:15,750 --> 00:27:14,470
have another no okay we have tons of

651
00:27:17,970 --> 00:27:15,760
these proteins right there's a lot of

652
00:27:19,889 --> 00:27:17,980
them we have a lot of data here and I'm

653
00:27:21,779 --> 00:27:19,899
really in this what we call phase three

654
00:27:23,850 --> 00:27:21,789
that's the earliest phase where we see

655
00:27:26,269 --> 00:27:23,860
protein the actual core the ribosome

656
00:27:30,240 --> 00:27:26,279
itself has essentially no protein so

657
00:27:32,940 --> 00:27:30,250
when protein starts to show up in in

658
00:27:34,799 --> 00:27:32,950
phase 3 it's never folded there's no

659
00:27:37,020 --> 00:27:34,809
alpha helixes there's no beta sheet

660
00:27:38,250 --> 00:27:37,030
there's no hydrophobic cores there's

661
00:27:40,620 --> 00:27:38,260
nothing like that

662
00:27:42,990 --> 00:27:40,630
just looks like that and then the second

663
00:27:45,120 --> 00:27:43,000

thing we have is these isolated

664
00:27:47,580 --> 00:27:45,130
secondary structural elements and they

665
00:27:50,280 --> 00:27:47,590
tend to be data sheets not alpha helixes

666
00:27:54,420 --> 00:27:50,290
and then in the next phase we see

667
00:27:58,580 --> 00:27:54,430
collapse of these secondary elements to

668
00:28:01,020 --> 00:27:58,590
globular structures and then and then

669
00:28:02,430 --> 00:28:01,030
then those globular structures elaborate

670
00:28:04,890 --> 00:28:02,440
and the nice thing is I'm only showing

671
00:28:06,600 --> 00:28:04,900
you the prokaryotic ribosomes if we add

672
00:28:08,850 --> 00:28:06,610
the eukaryotic there's that shell and

673
00:28:10,890 --> 00:28:08,860
that shell came after one and a half

674
00:28:13,740 --> 00:28:10,900
billion years so we have another layer

675
00:28:15,300 --> 00:28:13,750
out here where we can say what happened

676
00:28:17,670 --> 00:28:15,310
and you can really see that protein

677
00:28:19,320 --> 00:28:17,680
folding so this is a reaction at

678
00:28:22,050 --> 00:28:19,330
coordinate for protein folding maybe

679
00:28:23,160 --> 00:28:22,060
I'll just okay jump to this I'm just

680
00:28:24,690 --> 00:28:23,170
gonna jump ahead we have a bunch of

681
00:28:26,460 --> 00:28:24,700
statistics so I'm just gonna ignore all

682
00:28:28,440 --> 00:28:26,470
that but this is what it this is what it

683
00:28:31,530 --> 00:28:28,450
looks like this is the ribosomal RNA

684
00:28:32,970 --> 00:28:31,540
colored by phase right so the I've

685
00:28:34,920 --> 00:28:32,980
changed the coloring scheme I know it

686
00:28:36,990 --> 00:28:34,930
yeah so this is young and this is old

687
00:28:38,670 --> 00:28:37,000
and this is ribosomal protein and what

688
00:28:40,410 --> 00:28:38,680

we get is a reaction coordinate and we

689
00:28:44,280 --> 00:28:40,420
think this is the reaction coordinate

690
00:28:46,920 --> 00:28:44,290
for the evolution of protein folding and

691
00:28:48,930 --> 00:28:46,930
what what we can see in this this result

692
00:28:53,570 --> 00:28:48,940
I just going to jump to them getting cut

693
00:28:58,050 --> 00:28:53,580
off here is that protein folding

694
00:28:59,760 --> 00:28:58,060
occurred in a sea of RNA those secondary

695
00:29:02,970 --> 00:28:59,770
structural elements I talked to you

696
00:29:05,190 --> 00:29:02,980
about I showed you are enveloped in RNA

697
00:29:07,860 --> 00:29:05,200
and we think RNA basically chaperoned

698
00:29:09,990 --> 00:29:07,870
protein folding and that protein folding

699
00:29:13,740 --> 00:29:10,000
is an emergent property of interactions

700
00:29:14,760 --> 00:29:13,750
with RNA and I didn't really show you

701
00:29:17,570 --> 00:29:14,770
this but

702
00:29:20,790 --> 00:29:17,580
but the RNA folding changes around the

703
00:29:24,090 --> 00:29:20,800
are basically these two polymers are

704
00:29:27,030 --> 00:29:24,100
changing together the RNA is changing as

705
00:29:29,850 --> 00:29:27,040
protein folds and as protein folds RNA

706
00:29:31,860 --> 00:29:29,860
is it's folding is changing also not as

707
00:29:35,630 --> 00:29:31,870
dramatically but they're definitely more

708
00:29:37,440 --> 00:29:35,640
sophisticated RNA folds around

709
00:29:40,680 --> 00:29:37,450
sophisticated protein folds so that

710
00:29:43,260 --> 00:29:40,690
leads us to our sort of what we believe

711
00:29:46,950 --> 00:29:43,270
is happening is that RNA and protein are

712
00:29:49,860 --> 00:29:46,960
what we call molecular symbols and that

713
00:29:51,510 --> 00:29:49,870

the evolution of protein folding was

714
00:29:53,880 --> 00:29:51,520
chaperoned by RNA and the

715
00:29:55,890 --> 00:29:53,890
the evolution of RNA I'm not saying

716
00:29:58,020 --> 00:29:55,900
folding because you had stem loops and

717
00:30:01,440 --> 00:29:58,030
things but complex RNA assembly was

718
00:30:02,850 --> 00:30:01,450
chaperoned by protein so that kind of

719
00:30:05,550 --> 00:30:02,860
leads us to how I've started to think

720
00:30:08,730 --> 00:30:05,560
about biology in general this is because

721
00:30:10,980 --> 00:30:08,740
my friend Nick HUD grows figs and he

722
00:30:13,320 --> 00:30:10,990
gives me fig jam sometime and so I

723
00:30:15,360 --> 00:30:13,330
started thinking about fig trees and you

724
00:30:17,820 --> 00:30:15,370
know fig there's a symbiotic

725
00:30:20,780 --> 00:30:17,830
relationship between wasps and figs the

726
00:30:23,820 --> 00:30:20,790
wasps pollinate the figs and the figs

727
00:30:25,740 --> 00:30:23,830
feed the wasp embryos and there's this

728
00:30:28,260 --> 00:30:25,750
beautiful what's called a mutualism

729
00:30:30,090 --> 00:30:28,270
relationship and this is all over

730
00:30:31,620 --> 00:30:30,100
biology of course in fact you are

731
00:30:34,740 --> 00:30:31,630
involved in relationships like this with

732
00:30:36,810 --> 00:30:34,750
microbes in your gut and and there is

733
00:30:40,170 --> 00:30:36,820
something about mutualism relationships

734
00:30:42,150 --> 00:30:40,180
in biology that confers advantage and we

735
00:30:44,610 --> 00:30:42,160
basically say that goes down to the

736
00:30:48,780 --> 00:30:44,620
molecular level and that RNA and protein

737
00:30:51,390 --> 00:30:48,790
are molecular symbols all RNA is made by

738
00:30:52,770 --> 00:30:51,400

protein all protein is made by RNA and

739
00:30:54,600 --> 00:30:52,780
you could we could throw in membranes

740
00:30:57,870 --> 00:30:54,610
and other elements to this too but that

741
00:31:00,210 --> 00:30:57,880
basically we define life as a molecular

742
00:31:01,890 --> 00:31:00,220
symbiosis and that definition sort of

743
00:31:05,520 --> 00:31:01,900
comes from what we have learned from the

744
00:31:12,420 --> 00:31:05,530
ribosome so that i think i'm terminated

745
00:31:25,740 --> 00:31:22,590
okay thanks Lauren that was great um I

746
00:31:27,750 --> 00:31:25,750
find myself wondering as you described

747
00:31:29,730 --> 00:31:27,760
the secretion through time do we know

748
00:31:32,390 --> 00:31:29,740
much about the functional consequences

749
00:31:34,200 --> 00:31:32,400
of those modifications of the ribosome

750
00:31:37,590 --> 00:31:34,210
in your carry ox

751
00:31:39,360 --> 00:31:37,600
actually that's some of it has to do

752
00:31:41,400 --> 00:31:39,370
actually charlie was just asking with

753
00:31:43,770 --> 00:31:41,410
ribosome biogenesis I mean if you people

754
00:31:46,620 --> 00:31:43,780
have cut these things off and say what

755
00:31:48,390 --> 00:31:46,630
happens and ribosome biogenesis stalls

756
00:31:50,730 --> 00:31:48,400
out but actually in our lab we've taken

757
00:31:52,320 --> 00:31:50,740
those human expansion segments and we've

758
00:31:54,360 --> 00:31:52,330
done pulldown assays and say what

759
00:31:55,830 --> 00:31:54,370
proteins bind to them and it's really

760
00:31:58,650 --> 00:31:55,840
interesting I mean it's like the

761
00:32:00,600 --> 00:31:58,660
proteasome is associated with the

762
00:32:02,070 --> 00:32:00,610
ribosome it binds to that and so we just

763
00:32:05,190 --> 00:32:02,080

we think they're kind of especially

764
00:32:07,110 --> 00:32:05,200
those tentacles are and those those are

765
00:32:09,600 --> 00:32:07,120
kind of docking sites we think for

766
00:32:11,940 --> 00:32:09,610
things auxilary to translation you know

767
00:32:16,680 --> 00:32:11,950
some protein quality control things like

768
00:32:18,660 --> 00:32:16,690
that darn yeah nice to talk I'm

769
00:32:19,560 --> 00:32:18,670
concerned about one thing and maybe you

770
00:32:21,690 --> 00:32:19,570
can help me out with this

771
00:32:24,810 --> 00:32:21,700
when phylogenetic people are making

772
00:32:26,640 --> 00:32:24,820
sequences some there are short branches

773
00:32:28,290 --> 00:32:26,650
and some of their long branches but

774
00:32:30,930 --> 00:32:28,300
there is no such thing as a branch that

775
00:32:33,690 --> 00:32:30,940
has zero length and it seems to me that

776
00:32:36,860 --> 00:32:33,700
if you're married to this accretion only

777
00:32:40,380 --> 00:32:36,870
model what you've done is put the e coli

778
00:32:42,990 --> 00:32:40,390
ribosome which as some at the bottom

779
00:32:46,620 --> 00:32:43,000
with zero length and then accreted only

780
00:32:47,880 --> 00:32:46,630
but it seems to me that it's I mean when

781
00:32:49,470 --> 00:32:47,890
you're doing phylogenetic analysis you

782
00:32:50,790 --> 00:32:49,480
have deletions and insertions and I

783
00:32:54,300 --> 00:32:50,800
think they weight them pretty much

784
00:32:56,220 --> 00:32:54,310
equally you have waited only insertions

785
00:33:00,750 --> 00:32:56,230
and said there are no deletions is that

786
00:33:05,670 --> 00:33:00,760
right no actually I and you know I used

787
00:33:07,650 --> 00:33:05,680
that es7 and I and I could and really I

788
00:33:10,380 --> 00:33:07,660

should say this clearly that that that

789
00:33:12,630 --> 00:33:10,390
the the best estimate for the universal

790
00:33:15,690 --> 00:33:12,640
common core is the prokaryotic ribosomes

791
00:33:18,600 --> 00:33:15,700
and in some places you have things that

792
00:33:21,120 --> 00:33:18,610
are added on in in bacteria and in other

793
00:33:23,070 --> 00:33:21,130
places in archaea and that and there are

794
00:33:24,990 --> 00:33:23,080
we can see that places there are

795
00:33:25,840 --> 00:33:25,000
deletions I kind of didn't point it out

796
00:33:27,510 --> 00:33:25,850
because I was

797
00:33:29,980 --> 00:33:27,520
well for example you showed the

798
00:33:32,230 --> 00:33:29,990
comparison between ecoli and the Archaea

799
00:33:35,020 --> 00:33:32,240
and I'm thinking why do you assume that

800
00:33:37,390 --> 00:33:35,030
the e.coli is more basic and then

801
00:33:38,890 --> 00:33:37,400
something to the Archaea maybe the

802
00:33:40,960 --> 00:33:38,900
e.coli lost something in the last four

803
00:33:45,400 --> 00:33:40,970
okay here's a here's an example of

804
00:33:47,740 --> 00:33:45,410
exactly that so this is in the small

805
00:33:50,530 --> 00:33:47,750
subunit but you see this blue ecoli has

806
00:33:53,230 --> 00:33:50,540
this and sorry this is thermus

807
00:33:54,700 --> 00:33:53,240
thermophilus and sorry anyway yeah these

808
00:33:56,530 --> 00:33:54,710
are both bacteria so I can't really show

809
00:33:58,930 --> 00:33:56,540
that here but yes we have places where

810
00:34:01,420 --> 00:33:58,940
things get added on and taken off I mean

811
00:34:06,550 --> 00:34:01,430
here you can see a place where thermos

812
00:34:08,860 --> 00:34:06,560
has thermos has loss or the common

813
00:34:10,300 --> 00:34:08,870

ancestor something got added onto e.coli

814
00:34:12,910 --> 00:34:10,310
but not thermos and then something got

815
00:34:14,560 --> 00:34:12,920
add onto thermos but not e.coli so you

816
00:34:19,000 --> 00:34:14,570
know things are getting added on and

817
00:34:20,980 --> 00:34:19,010
taken off and we have yeah you're right

818
00:34:23,560 --> 00:34:20,990
so e.coli is not at the bottom no new

819
00:34:25,600 --> 00:34:23,570
Kiko lies not the bottom it yeah I said

820
00:34:27,310 --> 00:34:25,610
I was just sort of speaking shorthand

821
00:34:29,020 --> 00:34:27,320
it's our approximate I was using it as

822
00:34:31,030 --> 00:34:29,030
an approximation for the bottom but I

823
00:34:36,550 --> 00:34:31,040
could have used hello I could have used

824
00:34:39,880 --> 00:34:36,560
archaea for that actually yeah very

825
00:34:42,910 --> 00:34:39,890
interesting do you think you can deduce

826
00:34:47,620 --> 00:34:42,920
which ii was the lifestyle of luke if

827
00:34:51,250 --> 00:34:47,630
thermophiles meso file secret file no I

828
00:34:54,970 --> 00:34:51,260
don't think so I mean it might be

829
00:34:58,170 --> 00:34:54,980
possible once we know more but you know

830
00:35:00,430 --> 00:34:58,180
I mean what we know about Luca really is

831
00:35:04,930 --> 00:35:00,440
really what we know is the translation

832
00:35:06,430 --> 00:35:04,940
system and and and you know people who

833
00:35:08,980 --> 00:35:06,440
are trying to do you know does it live

834
00:35:11,670 --> 00:35:08,990
in a vent or whatever like actually Sean

835
00:35:15,910 --> 00:35:11,680
talked about that paper I mean people

836
00:35:17,770 --> 00:35:15,920
are sometimes willing to kind of make

837
00:35:20,230 --> 00:35:17,780
assumptions about what Lucas is but when

838
00:35:22,120 --> 00:35:20,240

I see that happening it really seems to

839
00:35:24,010 --> 00:35:22,130
be model dependent like you know if

840
00:35:26,200 --> 00:35:24,020
somebody is a vent person they're gonna

841
00:35:28,600 --> 00:35:26,210
say Luke I lived in a vent I promise you

842
00:35:32,290 --> 00:35:28,610
this and if they're you know I mean that

843
00:35:33,520 --> 00:35:32,300
is just what is going to happen and you

844
00:35:35,650 --> 00:35:33,530
know I just think you got to be very

845
00:35:37,300 --> 00:35:35,660
careful and so that's why we use the

846
00:35:39,490 --> 00:35:37,310
universal gene set I mean that's data

847
00:35:41,110 --> 00:35:39,500
that's there's no we're

848
00:35:42,910 --> 00:35:41,120
trying to make any assumptions about

849
00:35:44,560 --> 00:35:42,920
what was there we're just saying these

850
00:35:47,380 --> 00:35:44,570
are the things if you take ribosomal

851
00:35:50,890 --> 00:35:47,390
protein l-3 from bacteria and you've

852
00:35:52,840 --> 00:35:50,900
searched in yet strangest archaea you

853
00:35:55,240 --> 00:35:52,850
can find you will find it right I mean

854
00:35:57,790 --> 00:35:55,250
this is just that's but people are very

855
00:36:00,670 --> 00:35:57,800
you know commonly trying to infer things

856
00:36:02,410 --> 00:36:00,680
about Luca and I when I look at the

857
00:36:04,720 --> 00:36:02,420
literature it just seems like it they

858
00:36:10,240 --> 00:36:04,730
always get the answer they want so I'm

859
00:36:12,340 --> 00:36:10,250
very suspicious of those things science

860
00:36:15,940 --> 00:36:12,350
a great luck for molecular symbiosis and

861
00:36:19,120 --> 00:36:15,950
I have a question for ability nope short

862
00:36:21,130 --> 00:36:19,130
peptide of ribosomal proteins okay so

863
00:36:24,070 --> 00:36:21,140

before interaction in what are they have

864
00:36:26,230 --> 00:36:24,080
to be distributed in severe form so have

865
00:36:27,970 --> 00:36:26,240
you checked that sort of peptide or so

866
00:36:33,490 --> 00:36:27,980
the peptide process with two beta

867
00:36:35,650 --> 00:36:33,500
strands okay

868
00:36:37,780 --> 00:36:35,660
we actually we've made these peptides

869
00:36:40,660 --> 00:36:37,790
and they're soluble so we actually are

870
00:36:42,700 --> 00:36:40,670
studying the chaperoning of the peptide

871
00:36:46,270 --> 00:36:42,710
folding by RNA they're the peptides are

872
00:36:52,240 --> 00:36:46,280
soluble oh if you can such the secrets

873
00:36:54,370 --> 00:36:52,250
you can estimate the solubility these

874
00:36:56,110 --> 00:36:54,380
are soluble peptides because remember

875
00:36:58,420 --> 00:36:56,120
they're living in a sea of RNA and

876
00:37:01,030 --> 00:36:58,430
they're very highly charged they have a

877
00:37:02,920 --> 00:37:01,040
lottery but in the case of a peptide or

878
00:37:04,960 --> 00:37:02,930
in general peptide yeah okay so our

879
00:37:08,320 --> 00:37:04,970
model for the ribosome is the ribosome

880
00:37:10,570 --> 00:37:08,330
early on was doing nonspecific it was

881
00:37:13,300 --> 00:37:10,580
making and I don't I mean really

882
00:37:14,950 --> 00:37:13,310
nonspecific I mean esters you can make

883
00:37:16,600 --> 00:37:14,960
polyester with the ribosome it doesn't

884
00:37:21,100 --> 00:37:16,610
character so I think that the ribosome

885
00:37:23,980 --> 00:37:21,110
was making rasa mates of ester peptide

886
00:37:25,540 --> 00:37:23,990
it was and some small fraction of what

887
00:37:27,940 --> 00:37:25,550
it was making and it could have been a

888
00:37:29,380 --> 00:37:27,950

very small fraction because there wasn't

889
00:37:31,810 --> 00:37:29,390
a lot of competition in those days right

890
00:37:34,330 --> 00:37:31,820
we're talking about some small fraction

891
00:37:36,550 --> 00:37:34,340
stuck and conferred advantage right if

892
00:37:39,070 --> 00:37:36,560
most of it was insoluble or did nothing

893
00:37:40,780 --> 00:37:39,080
really didn't matter it's a if some

894
00:37:42,610 --> 00:37:40,790
small fraction stuck and conferred

895
00:37:44,290 --> 00:37:42,620
advantage and allowed the ribosome to

896
00:37:46,030 --> 00:37:44,300
fold better and function better and be

897
00:37:48,160 --> 00:37:46,040
more stable or whatever the Selective

898
00:37:51,850 --> 00:37:48,170
pressure was we don't really know but

899
00:37:53,410 --> 00:37:51,860
then then that would lock it in and and

900
00:37:55,150 --> 00:37:53,420
actually we believe

901
00:37:57,069 --> 00:37:55,160
that peptides got longer and the

902
00:37:58,660 --> 00:37:57,079
ribosomal RNA got bigger and that this

903
00:38:01,420 --> 00:37:58,670
was a kind of a bootstrapping process

904
00:38:02,980 --> 00:38:01,430
that cycled in on each other I didn't

905
00:38:05,140 --> 00:38:02,990
talk about the tunnel but we can really

906
00:38:06,520 --> 00:38:05,150
see we can infer sort of what was

907
00:38:08,319 --> 00:38:06,530
happening with the peptides by the

908
00:38:10,930 --> 00:38:08,329
length of the tunnel and it's it's

909
00:38:13,720 --> 00:38:10,940
pretty clear that there was a sort of a

910
00:38:17,410 --> 00:38:13,730
cyclic process peptides got bigger the

911
00:38:23,890 --> 00:38:17,420
ribosome got bigger okay I'm sorry but

912
00:38:27,819 --> 00:38:23,900
I'm I'm too long-winded to two questions

913
00:38:29,380 --> 00:38:27,829

learned one is is there any possibility

914
00:38:32,680 --> 00:38:29,390
of you play with the idea that instead

915
00:38:34,750 --> 00:38:32,690
of the the helix 25 being the ancestral

916
00:38:37,510 --> 00:38:34,760
state that since bacteria reduced in

917
00:38:40,630 --> 00:38:37,520
many ways that that entire arc eel bit

918
00:38:43,329 --> 00:38:40,640
is the primitive local legal for want of

919
00:38:46,030 --> 00:38:43,339
a better word state and that there was a

920
00:38:48,490 --> 00:38:46,040
loss of that upper part and the other

921
00:38:50,799 --> 00:38:48,500
question is have you thought at all

922
00:38:52,359 --> 00:38:50,809
about whether non ribosomal protein

923
00:38:56,680 --> 00:38:52,369
synthesis could have been earlier or

924
00:38:58,240 --> 00:38:56,690
later okay first about yeah I mean we we

925
00:39:01,359 --> 00:38:58,250
sort of are using the assumption that

926
00:39:03,730 --> 00:39:01,369
things that are common are that our best

927
00:39:05,109 --> 00:39:03,740
representation of the ancestor but it's

928
00:39:08,319 --> 00:39:05,119
not necessarily true

929
00:39:10,750 --> 00:39:08,329
so yeah yes you know we really can't

930
00:39:14,440 --> 00:39:10,760
exclude that that longer helix was

931
00:39:17,250 --> 00:39:14,450
ancestral we can't okay now the next

932
00:39:20,530 --> 00:39:17,260
what was the next question

933
00:39:22,750 --> 00:39:20,540
non ribosomal protein synthesis preceded

934
00:39:24,309 --> 00:39:22,760
the ribosome so that you could have well

935
00:39:26,440 --> 00:39:24,319
I guess I guess it depends on what you

936
00:39:28,270 --> 00:39:26,450
mean by non-ribosomal I do believe that

937
00:39:30,010 --> 00:39:28,280
peptides were being synthesized abiotic

938
00:39:32,170 --> 00:39:30,020

ly before the ribosome in fact I think

939
00:39:34,660 --> 00:39:32,180
all the ribosome did was help what was

940
00:39:36,069 --> 00:39:34,670
going on there and I think that possibly

941
00:39:38,200 --> 00:39:36,079
the original function of the ribosome

942
00:39:40,660 --> 00:39:38,210
was just to keep the two ends apart

943
00:39:42,970 --> 00:39:40,670
because and stop the siccolas ation that

944
00:39:44,920 --> 00:39:42,980
kills your polymerization so really all

945
00:39:47,559 --> 00:39:44,930
the ribosome it looks like in the very

946
00:39:50,140 --> 00:39:47,569
early part of the ribosome it was just a

947
00:39:52,299 --> 00:39:50,150
device where two amino acids could bind

948
00:39:54,640 --> 00:39:52,309
and they could cycle eyes and that ends

949
00:39:56,620 --> 00:39:54,650
were not allowed to come together so in

950
00:39:58,329 --> 00:39:56,630
that sense yes there was peptide

951
00:40:02,020 --> 00:39:58,339
synthesis going on before the ribosomes

952
00:40:03,700 --> 00:40:02,030
but you mean biological you know like or

953
00:40:05,410 --> 00:40:03,710
that still occurs I would have a hard

954
00:40:07,220 --> 00:40:05,420
time with that because that requires

955
00:40:08,870 --> 00:40:07,230
complex coded proteins

956
00:40:12,859 --> 00:40:08,880
and I can't imagine where they came from

957
00:40:15,109 --> 00:40:12,869
so yeah I'm not big on that model that

958
00:40:17,000 --> 00:40:15,119
there was that there were intact

959
00:40:20,599 --> 00:40:17,010
functional proteins before the ribosome

960
00:40:25,910 --> 00:40:20,609
I I don't know I I think that's a

961
00:40:27,470 --> 00:40:25,920
difficult model hey Lauren thanks to the

962
00:40:29,690 --> 00:40:27,480
nice talk

963
00:40:32,030 --> 00:40:29,700

I was wondering just how you

964
00:40:34,670 --> 00:40:32,040
conceptualize the gradual buildup of the

965
00:40:36,200 --> 00:40:34,680
ribosome in general there's these

966
00:40:38,180 --> 00:40:36,210
different kind of hierarchical levels

967
00:40:40,640 --> 00:40:38,190
that it seems like this progressive

968
00:40:45,430 --> 00:40:40,650
addition of subunits has occurred so you

969
00:40:47,480 --> 00:40:45,440
showed the cool squiggly green peptides

970
00:40:49,280 --> 00:40:47,490
gradually accumulating secondary

971
00:40:51,560 --> 00:40:49,290
structure and then looks like there's

972
00:40:54,109 --> 00:40:51,570
some tertiary structure coming in and so

973
00:40:55,700 --> 00:40:54,119
I don't there's not any ribosomes today

974
00:40:57,740 --> 00:40:55,710
that only have the green bit is that

975
00:41:00,380 --> 00:40:57,750
right that's correct okay right and so

976
00:41:01,970 --> 00:41:00,390
as that happens and in addition to that

977
00:41:03,620 --> 00:41:01,980
at the same time presumably there's all

978
00:41:05,810 --> 00:41:03,630
these different ribosomal subunits that

979
00:41:07,400 --> 00:41:05,820
are being tacked onto the side and and

980
00:41:10,070 --> 00:41:07,410
so we see all these different subunits

981
00:41:12,440 --> 00:41:10,080
in archaea versus bacteria versus

982
00:41:13,609 --> 00:41:12,450
eukaryotes so would it what's driving

983
00:41:15,890 --> 00:41:13,619
that and what is the functional

984
00:41:19,250 --> 00:41:15,900
consequence of that do you think okay

985
00:41:21,170 --> 00:41:19,260
well what we think is happening is that

986
00:41:23,810 --> 00:41:21,180
you have this core and small things were

987
00:41:25,700 --> 00:41:23,820
added on you know and and it kind of

988
00:41:28,220 --> 00:41:25,710

grew by the addition of small elements

989
00:41:31,550 --> 00:41:28,230
and small peptides small pieces of RNA

990
00:41:33,470 --> 00:41:31,560
and that grew in that way and we think

991
00:41:36,140 --> 00:41:33,480
there was a diverse population in fact

992
00:41:38,930 --> 00:41:36,150
that and we're maybe not talking about

993
00:41:40,700 --> 00:41:38,940
this the the origins of the ribosome we

994
00:41:42,410 --> 00:41:40,710
believe could have been before RNA the

995
00:41:44,300 --> 00:41:42,420
ancestor of the ribosome was not

996
00:41:48,200 --> 00:41:44,310
necessarily made of RNA it could have

997
00:41:50,240 --> 00:41:48,210
been an ancestral molecule in the and

998
00:41:52,099 --> 00:41:50,250
that product of the ribosome was not

999
00:41:54,620 --> 00:41:52,109
necessarily protein in fact we're pretty

1000
00:41:57,170 --> 00:41:54,630
sure it was not broken it was it was

1001
00:42:00,710 --> 00:41:57,180
peptide I mean it was esters and things

1002
00:42:02,359 --> 00:42:00,720
like that and I'm not answering your

1003
00:42:04,130 --> 00:42:02,369
question though so what tell me your

1004
00:42:06,349 --> 00:42:04,140
question again it's just something that

1005
00:42:08,450 --> 00:42:06,359
I wonder about when I look at all these

1006
00:42:10,190 --> 00:42:08,460
different ribosomal subunits you can

1007
00:42:11,390 --> 00:42:10,200
draw a Venn diagram and find that

1008
00:42:14,210 --> 00:42:11,400
there's all these different ones in

1009
00:42:16,280 --> 00:42:14,220
archaea and bacteria okay ribosomal

1010
00:42:17,599 --> 00:42:16,290
subunits in but too bright but this is a

1011
00:42:19,440 --> 00:42:17,609
different hierarchical level but it's

1012
00:42:21,090 --> 00:42:19,450
really in parallel to

1013
00:42:23,160 --> 00:42:21,100

you presented in that way and that table

1014
00:42:25,260 --> 00:42:23,170
of the green yellow and okay but the

1015
00:42:26,790 --> 00:42:25,270
ribosomal subunits are the same in

1016
00:42:28,860 --> 00:42:26,800
bacteria and archaea that's what I'm

1017
00:42:30,510 --> 00:42:28,870
getting confused by your crab but some

1018
00:42:32,520 --> 00:42:30,520
of them are different like bacteria have

1019
00:42:37,640 --> 00:42:32,530
different ones than archaea ones

1020
00:42:41,760 --> 00:42:37,650
different from eukaryote right no again

1021
00:42:43,140 --> 00:42:41,770
the ribosome has accreted yeah I'm not

1022
00:42:45,480 --> 00:42:43,150
I'm not understanding your question

1023
00:42:48,600 --> 00:42:45,490
there's no there's no different subunits

1024
00:42:51,360 --> 00:42:48,610
in archaea and vs. bacteria on the

1025
00:42:53,310 --> 00:42:51,370
periphery there is okay yes there's not

1026
00:42:54,570 --> 00:42:53,320
like how many ribosomal yeah okay

1027
00:42:56,040 --> 00:42:54,580
there's things added oh yeah we don't

1028
00:42:56,820 --> 00:42:56,050
call those subunits so some confused

1029
00:42:59,070 --> 00:42:56,830
yeah okay

1030
00:43:00,270 --> 00:42:59,080
there's expansion segments that's what

1031
00:43:01,890 --> 00:43:00,280
that's what you're talking about that

1032
00:43:03,450 --> 00:43:01,900
that are added onto the surface of the

1033
00:43:05,550 --> 00:43:03,460
ribosome we'll have to draw this out

1034
00:43:08,580 --> 00:43:05,560
later but if there's time I'll squeaking

1035
00:43:14,280 --> 00:43:08,590
another question okay so if I take that

1036
00:43:16,440 --> 00:43:14,290
common gene set and I and I throw that

1037
00:43:19,290 --> 00:43:16,450
in a tube what does it do and what can

1038
00:43:21,360 --> 00:43:19,300

that tell us about there Oh first life

1039
00:43:23,820 --> 00:43:21,370
if you take that common genes out you

1040
00:43:26,900 --> 00:43:23,830
can probably and a little bit other

1041
00:43:30,540 --> 00:43:26,910
stuff you could do in vitro translation

1042
00:43:34,430 --> 00:43:30,550
so that tells you I don't know you can

1043
00:43:37,800 --> 00:43:34,440
interpret that how you want but the way

1044
00:43:39,450 --> 00:43:37,810
but I think you can't ignore it right if

1045
00:43:43,050 --> 00:43:39,460
you have a model of the origin of life

1046
00:43:45,570 --> 00:43:43,060
that does not account for the extreme

1047
00:43:48,750 --> 00:43:45,580
conservation of the translation system

1048
00:43:51,060 --> 00:43:48,760
and the fact that that Luca had a

1049
00:43:53,460 --> 00:43:51,070
translation system that is untouchable

1050
00:43:54,930 --> 00:43:53,470
essentially by evolution and if you want

1051
00:43:56,580 --> 00:43:54,940
to take your next step back and say

1052
00:43:58,350 --> 00:43:56,590
there was no translation system and

1053
00:44:00,000 --> 00:43:58,360
there was a different biology then

1054
00:44:02,850 --> 00:44:00,010
that's what I find very difficult

1055
00:44:05,460 --> 00:44:02,860
because post luca the translation system

1056
00:44:07,800 --> 00:44:05,470
is the most permanent system in the

1057
00:44:10,350 --> 00:44:07,810
known universe and if you you know if

1058
00:44:12,390 --> 00:44:10,360
your evolutionary model goes a half a

1059
00:44:15,660 --> 00:44:12,400
step backwards and abolish as it totally

1060
00:44:18,510 --> 00:44:15,670
i have a problem with that as in the RNA

1061
00:44:20,250 --> 00:44:18,520
world for example right so the RNA world

1062
00:44:23,820 --> 00:44:20,260
says there was no translation system

1063
00:44:26,580 --> 00:44:23,830

biology matured and then a translation

1064
00:44:28,380 --> 00:44:26,590
system was invented that suddenly became

1065
00:44:31,560 --> 00:44:28,390
so important it could not it's

1066
00:44:32,650 --> 00:44:31,570
untouchable and that just seems kind of

1067
00:44:34,900 --> 00:44:32,660
incredible to me

1068
00:44:36,579 --> 00:44:34,910
I think this this this permanence and

1069
00:44:38,529 --> 00:44:36,589
the resilience and the robustness of the

1070
00:44:40,630 --> 00:44:38,539
translation system you've got to take it

1071
00:44:43,720 --> 00:44:40,640
back to the beginning you've got to say

1072
00:44:46,960 --> 00:44:43,730
that was built in but that's sort of to

1073
00:44:48,250 --> 00:44:46,970
me that's the simplest model well I

1074
00:44:50,289 --> 00:44:48,260
think you'll agree that this was a

1075
00:44:51,430 --> 00:44:50,299
fascinating talk so let's thank Laurie

1076
00:44:53,800 --> 00:44:51,440
thank you guys again

1077
00:45:06,480 --> 00:44:53,810
[Applause]