
AbSciCon
2019

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

2
00:00:11,280 --> 00:00:09,140
[Applause]

3
00:00:13,350 --> 00:00:11,290
first I want to start off by thanking

4
00:00:16,859 --> 00:00:13,360
the conveners for inviting me to give

5
00:00:18,660 --> 00:00:16,869
this talk thank you Raghav and Nick were

6
00:00:20,580 --> 00:00:18,670
you on that a bunch of people out here

7
00:00:23,609 --> 00:00:20,590
so thank you

8
00:00:24,630 --> 00:00:23,619
so what I what I'm sort of interested in

9
00:00:26,730 --> 00:00:24,640
understanding a lot of a certain

10
00:00:29,069 --> 00:00:26,740
understanding is the early evolution of

11
00:00:30,689 --> 00:00:29,079
life and if you want to do that then

12
00:00:33,260 --> 00:00:30,699
you'd better consider the environment

13
00:00:36,330 --> 00:00:33,270

that that evolution is happening in and

14
00:00:39,120 --> 00:00:36,340
so to try to understand the influence of

15
00:00:40,920 --> 00:00:39,130
the environment on evolution what what

16
00:00:42,450 --> 00:00:40,930
our group has done is used in vitro

17
00:00:45,360 --> 00:00:42,460
evolution techniques to try to

18
00:00:49,319 --> 00:00:45,370
understand how RNA evolution is

19
00:00:50,819 --> 00:00:49,329
influenced by its environment so there

20
00:00:52,950 --> 00:00:50,829
are several sort of environments that we

21
00:00:55,200 --> 00:00:52,960
can consider so we can think about the

22
00:00:57,180 --> 00:00:55,210
early geochemical environment we can

23
00:00:58,319 --> 00:00:57,190
think about cellular environments and we

24
00:01:01,169 --> 00:00:58,329
can think about something I'm calling

25
00:01:03,299 --> 00:01:01,179
the genomic environment now at the last

26
00:01:05,009 --> 00:01:03,309
apps icon I talked about some of the

27
00:01:06,570 --> 00:01:05,019
experiments we did to understand the

28
00:01:07,950 --> 00:01:06,580
influence of the geochemical environment

29
00:01:10,080 --> 00:01:07,960
in the cellular environment we've

30
00:01:11,670 --> 00:01:10,090
written a couple of papers on that and

31
00:01:13,530 --> 00:01:11,680
so if you want to look at those or you

32
00:01:15,030 --> 00:01:13,540
want to know more about that work come

33
00:01:16,410 --> 00:01:15,040
find me afterwards today I'm actually

34
00:01:18,719 --> 00:01:16,420
going to focus on the genomic

35
00:01:19,950 --> 00:01:18,729
environment and if the genomic

36
00:01:21,870 --> 00:01:19,960
environment sounds like kind of an odd

37
00:01:23,999 --> 00:01:21,880
term to you what I really just mean by

38
00:01:26,160 --> 00:01:24,009

that is that you have RNA structures

39
00:01:28,440 --> 00:01:26,170
that are evolving they're you know

40
00:01:29,999 --> 00:01:28,450
residing inside of a genome and those

41
00:01:32,700 --> 00:01:30,009
are those RNA structures can move around

42
00:01:34,649 --> 00:01:32,710
in the genome sequences can can drop in

43
00:01:36,660 --> 00:01:34,659
next to them and so in a certain way

44
00:01:38,160 --> 00:01:36,670
from the perspective of an RNA structure

45
00:01:40,080 --> 00:01:38,170
the genome is also something of an

46
00:01:42,480 --> 00:01:40,090
environment and specifically what I'm

47
00:01:44,789 --> 00:01:42,490
going to talk about today is something

48
00:01:47,069 --> 00:01:44,799
the role of duplication in in RNA

49
00:01:49,200 --> 00:01:47,079
evolution and so you know we know that

50
00:01:50,459 --> 00:01:49,210
we can have gene duplication and that's

51
00:01:52,499 --> 00:01:50,469
plays an important role in evolution

52
00:01:54,660 --> 00:01:52,509
because once you have a gene duplicated

53
00:01:56,370 --> 00:01:54,670
you have a redundant copy and if you

54
00:01:57,779 --> 00:01:56,380
have a redundant copy then the first

55
00:01:59,249 --> 00:01:57,789
copy can keep doing whatever it was

56
00:02:01,020 --> 00:01:59,259
doing before and the second copy can

57
00:02:03,359 --> 00:02:01,030
potentially evolve to do something new

58
00:02:05,370 --> 00:02:03,369
but another type of sequence duplication

59
00:02:07,440 --> 00:02:05,380
you can have is duplication with energy

60
00:02:10,380 --> 00:02:07,450
if you have duplication within a gene

61
00:02:12,150 --> 00:02:10,390
this can potentially lead to larger and

62
00:02:15,030 --> 00:02:12,160
more complex structures this is

63
00:02:16,350 --> 00:02:15,040

something that has been proposed for a

64
00:02:19,199 --> 00:02:16,360
number of RNAs and there's some evidence

65
00:02:20,820 --> 00:02:19,209
that for example tRNA evolution may have

66
00:02:22,260 --> 00:02:20,830
proceeded through some duplication event

67
00:02:24,900 --> 00:02:22,270
followed by structural rearrangement

68
00:02:26,580 --> 00:02:24,910
some folks like 80 oneth and others have

69
00:02:27,930 --> 00:02:26,590
looked at the ribosome and seen some

70
00:02:30,150 --> 00:02:27,940
symmetry in the core of the ribosome

71
00:02:31,980 --> 00:02:30,160
suggested that maybe it was a dimer and

72
00:02:34,470 --> 00:02:31,990
that eventually fused and so that's sort

73
00:02:36,600 --> 00:02:34,480
of a form of duplication there's

74
00:02:39,090 --> 00:02:36,610
certainly lots of protein architectures

75
00:02:41,430 --> 00:02:39,100
that seem to have evidence of past

76
00:02:42,780 --> 00:02:41,440
duplication events and so sequence

77
00:02:46,860 --> 00:02:42,790
duplication something is very important

78
00:02:48,210 --> 00:02:46,870
in biopolymer evolution in general so

79
00:02:49,830 --> 00:02:48,220
you might ask you know what are some of

80
00:02:52,920 --> 00:02:49,840
the mechanisms that could be involved in

81
00:02:54,390 --> 00:02:52,930
RNA duplication and just sort of very

82
00:02:56,670 --> 00:02:54,400
similar to it what we're talking about

83
00:02:58,949 --> 00:02:56,680
the previous talk you could have two

84
00:03:00,810 --> 00:02:58,959
copies of an RNA so I'm showing one one

85
00:03:03,570 --> 00:03:00,820
copy as a redline and one copy as a blue

86
00:03:05,370 --> 00:03:03,580
line and those RNAs can from time to

87
00:03:06,930 --> 00:03:05,380
time spontaneously degrade and generate

88
00:03:08,730 --> 00:03:06,940

these two prime three prime cyclic

89
00:03:09,810 --> 00:03:08,740
phosphates we just heard about and as

90
00:03:12,420 --> 00:03:09,820
long as they're not lining up with

91
00:03:14,280 --> 00:03:12,430
Watson Crick pairing they can actually

92
00:03:17,730 --> 00:03:14,290
potentially react with each other and

93
00:03:19,680 --> 00:03:17,740
form an RNA that's now a duplicated

94
00:03:21,120 --> 00:03:19,690
version of the sequence another way to

95
00:03:23,250 --> 00:03:21,130
get duplication and it's a way that we

96
00:03:24,870 --> 00:03:23,260
continue to get it today is through

97
00:03:27,960 --> 00:03:24,880
rolling circle replication that's where

98
00:03:29,400 --> 00:03:27,970
you if you have a circular RNA and you

99
00:03:31,229 --> 00:03:29,410
have some sort of polymerase with it's a

100
00:03:33,300 --> 00:03:31,239
protein or a ribozyme or whatever it is

101
00:03:35,100 --> 00:03:33,310
as it goes around the circular RNA it

102
00:03:36,449 --> 00:03:35,110
makes one copy but because they are the

103
00:03:37,860 --> 00:03:36,459
template is circular it can just keep

104
00:03:40,620 --> 00:03:37,870
going you get multiple copies you can

105
00:03:42,240 --> 00:03:40,630
get to three for however many you want

106
00:03:44,970 --> 00:03:42,250
also you can imagine a polymerase is

107
00:03:46,979 --> 00:03:44,980
copying a template and then the the

108
00:03:48,509 --> 00:03:46,989
copied strands sort of slips moves

109
00:03:50,370 --> 00:03:48,519
dissociates three associates and it gets

110
00:03:51,960 --> 00:03:50,380
reacted the point is that there there

111
00:03:54,060 --> 00:03:51,970
are several different mechanisms that

112
00:03:55,530 --> 00:03:54,070
can lead to duplication when there's

113
00:03:57,150 --> 00:03:55,540

evidence that we've had duplication in

114
00:03:58,860 --> 00:03:57,160
the past so we kind of want to think

115
00:04:02,280 --> 00:03:58,870
about you know what are the consequences

116
00:04:03,960 --> 00:04:02,290
of duplication so for sake of example

117
00:04:06,660 --> 00:04:03,970
here's a little cartoon of a little

118
00:04:08,940 --> 00:04:06,670
imaginary RNA it's got an internal loop

119
00:04:10,530 --> 00:04:08,950
and it's got two stems and if we

120
00:04:12,750 --> 00:04:10,540
duplicate it what do we get well we get

121
00:04:14,970 --> 00:04:12,760
a spare copy of the functional loop

122
00:04:16,259 --> 00:04:14,980
right so similar to gene duplication so

123
00:04:18,150 --> 00:04:16,269
potentially that can evolve to do

124
00:04:19,860 --> 00:04:18,160
something else but another consequence

125
00:04:22,020 --> 00:04:19,870
that may not be immediately obvious is

126
00:04:23,610 --> 00:04:22,030
that once you duplicate a sequence you

127
00:04:25,530 --> 00:04:23,620
have this immediate folding

128
00:04:27,810 --> 00:04:25,540
heterogeneity problem because any base

129
00:04:29,670 --> 00:04:27,820
pair you can form within one copy of the

130
00:04:31,140 --> 00:04:29,680
sequence you can of course form between

131
00:04:32,700 --> 00:04:31,150
the two copies of the sequence and so

132
00:04:34,140 --> 00:04:32,710
you have to deal with that so

133
00:04:37,140 --> 00:04:34,150
these are a couple things that could be

134
00:04:38,760 --> 00:04:37,150
consequences of duplication it's all

135
00:04:40,770 --> 00:04:38,770
kind of you know theoretical ok let's

136
00:04:42,480 --> 00:04:40,780
let's actually do some experiments and

137
00:04:44,640 --> 00:04:42,490
so what we decided to do is we took a

138
00:04:46,890 --> 00:04:44,650

very well characterized functional RNA

139
00:04:49,379 --> 00:04:46,900
it's an ATP a primer that was evolved

140
00:04:50,850 --> 00:04:49,389
many years ago in the Shaw stack lab and

141
00:04:53,040 --> 00:04:50,860
it's since been shown to be present

142
00:04:55,409 --> 00:04:53,050
actually in in biological RNAs as well

143
00:04:57,570 --> 00:04:55,419
and we know a lot about it we know that

144
00:04:59,310 --> 00:04:57,580
ATP binds to this loop that's in the

145
00:05:01,590 --> 00:04:59,320
middle it's very simple similar to my

146
00:05:03,240 --> 00:05:01,600
cartoon on the previous slide we know

147
00:05:05,310 --> 00:05:03,250
what the 3d structure looks like thanks

148
00:05:07,589 --> 00:05:05,320
to NMR and so we decided we'd take the

149
00:05:10,080 --> 00:05:07,599
sequence we duplicate it we make two

150
00:05:11,999 --> 00:05:10,090
copies and then we would mutagenize it

151
00:05:13,740 --> 00:05:12,009
so introduce a bunch of mutations to

152
00:05:16,740 --> 00:05:13,750
this parent sequence that has two copies

153
00:05:18,719 --> 00:05:16,750
of the abdomen alright so now we have a

154
00:05:20,969 --> 00:05:18,729
library of mutants first thing we do is

155
00:05:24,060 --> 00:05:20,979
sequence it so we know what sequences

156
00:05:25,499 --> 00:05:24,070
we're dealing with and then we decided

157
00:05:27,300 --> 00:05:25,509
to to look at these mutants and screen

158
00:05:29,909 --> 00:05:27,310
them for find out which ones still have

159
00:05:32,670 --> 00:05:29,919
the ability to bind ATP after both

160
00:05:34,260 --> 00:05:32,680
duplication and mutation and so we run

161
00:05:36,060 --> 00:05:34,270
them over an ATP column we throw away

162
00:05:39,120 --> 00:05:36,070
everything every mutant that has lost

163
00:05:40,589 --> 00:05:39,130

its ability to bind ATP we add ATP to

164
00:05:43,260 --> 00:05:40,599
the column to recover everything that

165
00:05:45,689 --> 00:05:43,270
still binds ATP and we can then amplify

166
00:05:47,520 --> 00:05:45,699
those sequences and now we end up with a

167
00:05:49,649 --> 00:05:47,530
population that's enriched in mutants

168
00:05:51,540 --> 00:05:49,659
that tend to bind ATP and you can repeat

169
00:05:52,890 --> 00:05:51,550
that for several rounds and you

170
00:05:55,620 --> 00:05:52,900
basically have several generations of

171
00:05:57,000 --> 00:05:55,630
in-vitro revolution so that's one option

172
00:05:59,460 --> 00:05:57,010
and the other thing that we can do

173
00:06:02,100 --> 00:05:59,470
because this is this duplication event

174
00:06:04,230 --> 00:06:02,110
might allow us to have a second function

175
00:06:06,390 --> 00:06:04,240
evolve is instead of amplifying the ATP

176
00:06:09,029 --> 00:06:06,400
binders after we pull them off the ATP

177
00:06:10,890 --> 00:06:09,039
column we can just take those and add

178
00:06:12,870 --> 00:06:10,900
them directly to a second column a gtp

179
00:06:14,760 --> 00:06:12,880
column in this case and see if we can

180
00:06:17,430 --> 00:06:14,770
find RNAs that can simultaneously bind

181
00:06:20,430 --> 00:06:17,440
ATP and also bind to the gtp column so

182
00:06:22,140 --> 00:06:20,440
we recover the the mutants that that

183
00:06:23,760 --> 00:06:22,150
bind to both columns we amplify them we

184
00:06:25,499 --> 00:06:23,770
repeat this and again we get a

185
00:06:27,089 --> 00:06:25,509
presumably a population that's enriched

186
00:06:30,870 --> 00:06:27,099
on things that are retained on both

187
00:06:34,620 --> 00:06:30,880
columns so that's the set up what's the

188
00:06:36,629 --> 00:06:34,630

result well for the most part when we re

189
00:06:38,399 --> 00:06:36,639
selected for ATP binding and even when

190
00:06:41,310 --> 00:06:38,409
we did selections that included a gtp

191
00:06:44,159 --> 00:06:41,320
binding step pretty much every mutant

192
00:06:45,420 --> 00:06:44,169
that that was successful could fold into

193
00:06:45,869 --> 00:06:45,430
the structure shown here which is the

194
00:06:47,730 --> 00:06:45,879
same sir

195
00:06:49,290 --> 00:06:47,740
sure that the sort of wild type starting

196
00:06:51,929 --> 00:06:49,300
sequence had sort of a tandem

197
00:06:53,730 --> 00:06:51,939
arrangement of the abbe de mer so we had

198
00:06:56,249 --> 00:06:53,740
sequences so here's just an example so

199
00:06:58,889 --> 00:06:56,259
we had a mutant where this U is mutate

200
00:07:00,600 --> 00:06:58,899
to an a this a to a u this you to an a

201
00:07:02,699 --> 00:07:00,610
this a to a you it's just this is a

202
00:07:05,730 --> 00:07:02,709
4-way mutant that was very successful

203
00:07:08,159 --> 00:07:05,740
and you can see that this has the same

204
00:07:10,139 --> 00:07:08,169
secondary structure as the wild-type we

205
00:07:12,029 --> 00:07:10,149
don't disrupt any of the base pairs we

206
00:07:13,769 --> 00:07:12,039
just changed the sequence another

207
00:07:15,629 --> 00:07:13,779
consequence though this is does a little

208
00:07:16,889 --> 00:07:15,639
bit better than the wild-type because it

209
00:07:18,239 --> 00:07:16,899
helps to avoid some of that folding

210
00:07:19,999 --> 00:07:18,249
heterogeneity because we no longer have

211
00:07:21,570 --> 00:07:20,009
the same sequence two times in a row

212
00:07:23,309 --> 00:07:21,580
because that's what pretty much

213
00:07:24,329 --> 00:07:23,319

everything looks like there are some

214
00:07:27,059 --> 00:07:24,339
exceptions though there were some

215
00:07:29,339 --> 00:07:27,069
sequences that that did fairly well like

216
00:07:30,809 --> 00:07:29,349
this one here where I'm showing the

217
00:07:32,519 --> 00:07:30,819
different mutations mapped onto the

218
00:07:34,619 --> 00:07:32,529
structure and you can see that this

219
00:07:36,809 --> 00:07:34,629
three-way mutant all of these mutations

220
00:07:38,999 --> 00:07:36,819
are incompatible with the original

221
00:07:40,649 --> 00:07:39,009
structure the sequence did very well in

222
00:07:42,540 --> 00:07:40,659
our selections in all of our selections

223
00:07:45,989 --> 00:07:42,550
so did a number of other similar

224
00:07:47,790 --> 00:07:45,999
sequences but if we map these same

225
00:07:49,619 --> 00:07:47,800
mutants on to a different secondary

226
00:07:51,869 --> 00:07:49,629
structure this is what we're calling a

227
00:07:53,969 --> 00:07:51,879
nested secondary structure we see that

228
00:07:55,259 --> 00:07:53,979
that these mutations sort of make sense

229
00:07:56,850 --> 00:07:55,269
in the context of this structure and

230
00:07:58,949 --> 00:07:56,860
this is the predicted free energy

231
00:08:00,480 --> 00:07:58,959
structure for this sequence and what I

232
00:08:03,119 --> 00:08:00,490
want you to notice what this structure

233
00:08:04,859 --> 00:08:03,129
is that one this nested structure it

234
00:08:06,449 --> 00:08:04,869
doesn't have any base pairs in common

235
00:08:07,649 --> 00:08:06,459
with the tandem structure so the

236
00:08:10,499 --> 00:08:07,659
secondary structures couldn't be more

237
00:08:13,259 --> 00:08:10,509
different but we retain the two binding

238
00:08:14,369 --> 00:08:13,269

loops and so the actual

239
00:08:15,989 --> 00:08:14,379
three-dimensional structure that's

240
00:08:20,369 --> 00:08:15,999
responsible for binding hasn't changed

241
00:08:22,290 --> 00:08:20,379
at all so not surprisingly when we look

242
00:08:23,969 --> 00:08:22,300
at these two different structures and we

243
00:08:26,059 --> 00:08:23,979
we try to see how well they bind to the

244
00:08:29,399 --> 00:08:26,069
ATP column we see very similar

245
00:08:31,019 --> 00:08:29,409
affinities we do see a little bit more

246
00:08:33,059 --> 00:08:31,029
of this nested structure when our

247
00:08:36,629 --> 00:08:33,069
evolution experiments include the gtp

248
00:08:38,339 --> 00:08:36,639
binding step but it's not really because

249
00:08:41,579 --> 00:08:38,349
they're binding well the gtp column none

250
00:08:43,259 --> 00:08:41,589
either of these as I would expect binds

251
00:08:45,269 --> 00:08:43,269
particularly well to the GTP agarose

252
00:08:48,629 --> 00:08:45,279
column however it has some subtle effect

253
00:08:50,100 --> 00:08:48,639
on the balance of the two now one thing

254
00:08:51,629 --> 00:08:50,110
I think is kind of interesting about

255
00:08:54,120 --> 00:08:51,639
these duplication events is that it

256
00:08:55,860 --> 00:08:54,130
makes it very easy to evolve from one

257
00:08:58,110 --> 00:08:55,870
structure one functional structure to

258
00:08:59,530 --> 00:08:58,120
another so for example we had this this

259
00:09:01,780 --> 00:08:59,540
C mutated to at you

260
00:09:04,389 --> 00:09:01,790
poink mutant and it's predicted to form

261
00:09:06,850 --> 00:09:04,399
the the tandem secondary structure and

262
00:09:08,800 --> 00:09:06,860
then but then if if you pick up a second

263
00:09:10,990 --> 00:09:08,810

mutation in addition to that it's

264
00:09:12,220 --> 00:09:11,000
actually then favors the nested

265
00:09:13,509 --> 00:09:12,230
secondary structure so you get this

266
00:09:15,730 --> 00:09:13,519
shift in the minimum for energy

267
00:09:17,800 --> 00:09:15,740
structure in a single point mutation and

268
00:09:18,730 --> 00:09:17,810
you just go directly from one functional

269
00:09:19,960 --> 00:09:18,740
structure to the other which is

270
00:09:21,220 --> 00:09:19,970
something that's very difficult to do

271
00:09:24,340 --> 00:09:21,230
generally but in the case of duplication

272
00:09:27,639 --> 00:09:24,350
that seems to be pretty easy all right

273
00:09:30,759 --> 00:09:27,649
and so mmm that was all based on energy

274
00:09:32,470 --> 00:09:30,769
free energy structure calculations done

275
00:09:34,210 --> 00:09:32,480
on a computer but we wanted to do some

276
00:09:37,480 --> 00:09:34,220
experiments to verify that so we did

277
00:09:39,910 --> 00:09:37,490
some non denaturing gels and basically

278
00:09:42,220 --> 00:09:39,920
if we look at the wild-type sequence or

279
00:09:43,840 --> 00:09:42,230
we look at mutants that are expected to

280
00:09:46,199 --> 00:09:43,850
be in the tandem arrangement we get

281
00:09:48,519 --> 00:09:46,209
characteristic mobility on a native gel

282
00:09:50,410 --> 00:09:48,529
and if we look at sequences that are

283
00:09:52,840 --> 00:09:50,420
predicted to have the nested structure

284
00:09:54,639 --> 00:09:52,850
they have a different mobility if we go

285
00:09:56,949 --> 00:09:54,649
back to the point point mutant I had on

286
00:09:58,600 --> 00:09:56,959
the previous slide it actually is

287
00:10:00,040 --> 00:09:58,610
predicted the free energy difference

288
00:10:01,360 --> 00:10:00,050

between the two structures is predicted

289
00:10:03,550 --> 00:10:01,370
to be very small and so we kind of

290
00:10:05,110 --> 00:10:03,560
expect to have a combination of both

291
00:10:06,519 --> 00:10:05,120
tandem and nested within the same

292
00:10:08,710 --> 00:10:06,529
sequence and the native gel seems to

293
00:10:09,579 --> 00:10:08,720
support that because our our peak is

294
00:10:12,280 --> 00:10:09,589
shifted a little bit

295
00:10:14,530 --> 00:10:12,290
towards the towards the migration

296
00:10:15,759 --> 00:10:14,540
pattern of the the nested structure and

297
00:10:17,439 --> 00:10:15,769
there's even a little bit of a shoulder

298
00:10:19,689 --> 00:10:17,449
there and we make that second mutation

299
00:10:21,460 --> 00:10:19,699
then it's pretty strongly favors the

300
00:10:25,230 --> 00:10:21,470
nested conformation and then and and our

301
00:10:27,790 --> 00:10:25,240
native gel sort of pair that out so

302
00:10:29,650 --> 00:10:27,800
before and and so just the one last

303
00:10:32,079 --> 00:10:29,660
result that I wanted to make sure I

304
00:10:35,500 --> 00:10:32,089
covered is that there are some mutants

305
00:10:37,660 --> 00:10:35,510
at high at a distance that are totally

306
00:10:40,000 --> 00:10:37,670
incompatible with what we know about ATP

307
00:10:42,550 --> 00:10:40,010
binding so the like for example here's a

308
00:10:45,309 --> 00:10:42,560
mutant that disrupts one copy of the ATP

309
00:10:49,569 --> 00:10:45,319
binding motif and in fact it doesn't

310
00:10:51,550 --> 00:10:49,579
bind to the column nearly as well and we

311
00:10:53,650 --> 00:10:51,560
only see this sequence and many like it

312
00:10:55,509 --> 00:10:53,660
when the GTP binding step is included so

313
00:10:57,309 --> 00:10:55,519

you might think okay maybe we've found

314
00:10:59,350 --> 00:10:57,319
an effective GTP binder something that

315
00:11:02,439 --> 00:10:59,360
can simultaneously bind ATP at one site

316
00:11:03,699 --> 00:11:02,449
GTP at the other if it does bind GTP or

317
00:11:04,809 --> 00:11:03,709
the GTP column it doesn't do it very

318
00:11:06,550 --> 00:11:04,819
well

319
00:11:09,490 --> 00:11:06,560
if very little of the material is

320
00:11:11,470 --> 00:11:09,500
retained on the GTP column though make

321
00:11:13,590 --> 00:11:11,480
of it what you will it's a very very

322
00:11:15,510 --> 00:11:13,600
tiny but reproduce

323
00:11:17,910 --> 00:11:15,520
a higher amount that comes off than the

324
00:11:21,270 --> 00:11:17,920
other sequences but like I said it's not

325
00:11:23,490 --> 00:11:21,280
very much but anyway the main sort of

326
00:11:25,550 --> 00:11:23,500
takeaway from from our experiments is

327
00:11:28,170 --> 00:11:25,560
that you know we've shown an example

328
00:11:30,720 --> 00:11:28,180
where you can have a very large change

329
00:11:32,640 --> 00:11:30,730
in secondary structure without any real

330
00:11:33,720 --> 00:11:32,650
change in the the actual functional

331
00:11:37,110 --> 00:11:33,730
three-dimensional structure that's

332
00:11:38,430 --> 00:11:37,120
responsible for function and so we have

333
00:11:40,260 --> 00:11:38,440
these two completely different secondary

334
00:11:43,620 --> 00:11:40,270
structures we have the same motif two

335
00:11:44,780 --> 00:11:43,630
times and so I was thinking about you

336
00:11:47,100 --> 00:11:44,790
know what the implications are for

337
00:11:48,090 --> 00:11:47,110
looking at naturally occurring RNAs and

338
00:11:49,710 --> 00:11:48,100

what does this tell us about how we

339
00:11:53,180 --> 00:11:49,720
should look at them and one thing it

340
00:11:56,040 --> 00:11:53,190
tells us is that you yes you can have

341
00:11:57,570 --> 00:11:56,050
duplication followed by structural

342
00:11:59,040 --> 00:11:57,580
rearrangement with a very small number

343
00:12:00,330 --> 00:11:59,050
of mutations and so that would be

344
00:12:02,400 --> 00:12:00,340
consistent with some of the models for

345
00:12:04,470 --> 00:12:02,410
tRNA evolution that to propose the same

346
00:12:06,210 --> 00:12:04,480
thing the other thing that I think is

347
00:12:07,590 --> 00:12:06,220
important to think about is that if

348
00:12:09,000 --> 00:12:07,600
there were duplication events in the

349
00:12:10,740 --> 00:12:09,010
past or to be very careful when looking

350
00:12:12,570 --> 00:12:10,750
at modern rnas and trying to determine

351
00:12:14,730 --> 00:12:12,580
what the ancestral form looked like

352
00:12:17,280 --> 00:12:14,740
because there is this opportunity and

353
00:12:19,110 --> 00:12:17,290
least in this particular mechanism to

354
00:12:21,090 --> 00:12:19,120
have a change in the secondary structure

355
00:12:23,490 --> 00:12:21,100
and overall what I kind of take away

356
00:12:25,140 --> 00:12:23,500
from this is that it paints a consistent

357
00:12:26,400 --> 00:12:25,150
picture with a lot of other work which

358
00:12:28,680 --> 00:12:26,410
is that there's sort of a hierarchy of

359
00:12:29,970 --> 00:12:28,690
conservation and biomolecules which is

360
00:12:31,710 --> 00:12:29,980
that you know if you really want to

361
00:12:33,180 --> 00:12:31,720
identify what's going to be conserved

362
00:12:34,230 --> 00:12:33,190
over the longest period of time and it's

363
00:12:36,600 --> 00:12:34,240

going to be the best thing to look at

364
00:12:39,420 --> 00:12:36,610
it's a three-dimensional structure and

365
00:12:41,730 --> 00:12:39,430
then maybe a little bit less reliable

366
00:12:42,870 --> 00:12:41,740
secondary structure and then and then

367
00:12:43,890 --> 00:12:42,880
sort of the least conserved thing is

368
00:12:46,500 --> 00:12:43,900
sequences and so that I think that's

369
00:12:47,610 --> 00:12:46,510
that sort of hierarchy of confidence is

370
00:12:49,860 --> 00:12:47,620
something we should keep in mind when

371
00:12:50,910 --> 00:12:49,870
doing ancestral reconstructions and it

372
00:12:52,380 --> 00:12:50,920
highlights the importance of

373
00:12:54,410 --> 00:12:52,390
understanding the three-dimensional

374
00:12:55,950 --> 00:12:54,420
structure and with that I'll just

375
00:12:56,960 --> 00:12:55,960
acknowledge the people have been

376
00:12:59,310 --> 00:12:56,970
involved in some of this work

377
00:13:01,080 --> 00:12:59,320
particularly want to acknowledge and

378
00:13:03,300 --> 00:13:01,090
Rupa Banach who is an intern who

379
00:13:05,280 --> 00:13:03,310
basically designed these experiments did

380
00:13:07,560 --> 00:13:05,290
the selections was heavily involved the

381
00:13:10,290 --> 00:13:07,570
analysis and and did a lot more than

382
00:13:12,540 --> 00:13:10,300
then I would have any right to expect

383
00:13:14,550 --> 00:13:12,550
from an intern so that was really

384
00:13:16,170 --> 00:13:14,560
fantastic he's off at grad school at UC

385
00:13:24,500 --> 00:13:16,180
Berkeley now and if we have time

386
00:13:39,569 --> 00:13:33,660
yesterday overtime let's start here why

387
00:13:42,210 --> 00:13:39,579
is that so well because I would say if

388
00:13:44,579 --> 00:13:42,220

you're trying to conserve a function I

389
00:13:45,930 --> 00:13:44,589
mean it is it's the active site it's

390
00:13:47,940 --> 00:13:45,940
it's the three-dimensional structure

391
00:13:51,060 --> 00:13:47,950
arrangement of atoms that matters right

392
00:13:53,430 --> 00:13:51,070
so you can you can retain like it's like

393
00:13:55,230 --> 00:13:53,440
I've shown the simpler case the

394
00:13:57,060 --> 00:13:55,240
secondary structure right it's you know

395
00:13:59,069 --> 00:13:57,070
sequences you can serve well because you

396
00:14:00,480 --> 00:13:59,079
can you can do you can flip a base pair

397
00:14:02,100 --> 00:14:00,490
and it's a different sequence but it's

398
00:14:04,199 --> 00:14:02,110
the exact same structure well is the

399
00:14:07,970 --> 00:14:04,209
same thing here you can rearrange your

400
00:14:19,250 --> 00:14:07,980
secondary structures but keep the same

401
00:14:24,240 --> 00:14:22,769
so I guess what you have to recognize is

402
00:14:28,439 --> 00:14:24,250
that there that there are you know you

403
00:14:30,750 --> 00:14:28,449
have to look at a motif right and so you

404
00:14:31,380 --> 00:14:30,760
know it's somewhat modular right so you

405
00:14:33,150 --> 00:14:31,390
can't

406
00:14:34,380 --> 00:14:33,160
so yes the entire three-dimensional

407
00:14:37,170 --> 00:14:34,390
structure of the whole molecule is

408
00:14:39,150 --> 00:14:37,180
different obviously but there are units

409
00:15:08,700 --> 00:14:39,160
and modules within it and so that I

410
00:15:15,700 --> 00:15:12,280
okay there's a lot there but yeah so I

411
00:15:17,950 --> 00:15:15,710
mean we haven't sort of independently

412
00:15:19,540 --> 00:15:17,960
measured the KDS I mean you know you'd

413
00:15:22,330 --> 00:15:19,550

expect them for the individual binding

414
00:15:23,890 --> 00:15:22,340
sites to be pretty much the same but you

415
00:15:25,840 --> 00:15:23,900
know we do know that if you disrupt one

416
00:15:27,790 --> 00:15:25,850
copy then they don't bind nearly as well

417
00:15:31,900 --> 00:15:27,800
is so obviously the the presence of two

418
00:15:34,540 --> 00:15:31,910
copies they're doing better and as far

419
00:15:37,090 --> 00:15:34,550
as cooperativity and it's not clear yeah

420
00:15:39,010 --> 00:15:37,100
you know I was it's possible right if

421
00:15:45,850 --> 00:15:39,020
there were some new structure that form

422
00:15:47,470 --> 00:15:45,860
that was really tight oh I see what

423
00:15:50,380 --> 00:15:47,480
you're saying yeah yeah I mean you know

424
00:15:52,740 --> 00:15:50,390
they seem to elute reason well yeah

425
00:16:00,569 --> 00:15:52,750
that's always a possibility for sure

426
00:16:04,289 --> 00:16:02,910
four letters which is why you get these

427
00:16:10,320 --> 00:16:04,299
misfortunes you get them as you

428
00:16:13,970 --> 00:16:12,090
of course obviously one solutions with

429
00:16:15,930 --> 00:16:13,980
one expansion

430
00:16:17,390 --> 00:16:15,940
there's another way that's to go to

431
00:16:32,590 --> 00:16:17,400
higher temperatures without

432
00:16:41,300 --> 00:16:39,170
well we haven't done that I would know

433
00:16:42,290 --> 00:16:41,310
we haven't done that I mean it would be

434
00:16:44,720 --> 00:16:42,300
it would be interesting to see what