AbSciCon 2019

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

2
00:00:12,090 --> 00:00:09,320
[Applause]

3
00:00:14,520 --> 00:00:12,100
hello my name is Anton Mohammed I'm from

4
00:00:16,529 --> 00:00:14,530
Oakland University in Michigan and I'm a

5
00:00:19,380 --> 00:00:16,539
master student and I'm working with dr.

6
00:00:21,960 --> 00:00:19,390
favia but the suit see on domain

7
00:00:26,009 --> 00:00:21,970
signatures of genome complexity and

8
00:00:29,099 --> 00:00:26,019
extremophiles so what are extremophiles

9
00:00:31,980 --> 00:00:29,109
extremophiles are those interesting and

10
00:00:34,260 --> 00:00:31,990
unique organisms that live in extreme

11
00:00:38,910 --> 00:00:34,270
environments and they live in hot and

12
00:00:41,040 --> 00:00:38,920
cold niches and they also live in acid

13
00:00:43,170 --> 00:00:41,050

solutions alkaline solutions salty

14
00:00:46,950 --> 00:00:43,180
environments and they include members of

15
00:00:49,259 --> 00:00:46,960
all they glued members of all domain all

16
00:00:53,189 --> 00:00:49,269
domains of life are kiai eukaryotes and

17
00:00:55,860 --> 00:00:53,199
bacteria and archaea is the major group

18
00:00:57,779 --> 00:00:55,870
that survive in extreme environments and

19
00:00:59,880 --> 00:00:57,789
although members of this group are less

20
00:01:02,040 --> 00:00:59,890
versatile than bacteria and eukaryotes

21
00:01:05,340 --> 00:01:02,050
they're quite skilled at adapting to

22
00:01:08,820 --> 00:01:05,350
extreme conditions and extremophiles

23
00:01:11,310 --> 00:01:08,830
exist in all of the archaeal phyla and

24
00:01:13,380 --> 00:01:11,320
in our data we only have the

25
00:01:15,690 --> 00:01:13,390
crenarchaeota the URI archaea and the

26
00:01:19,950 --> 00:01:15,700
thumb mark iota and these are some

27
00:01:22,830 --> 00:01:19,960
examples of the extremophiles we have

28
00:01:25,530 --> 00:01:22,840
the hyperthermophiles in red and then

29
00:01:27,660 --> 00:01:25,540
the thermo files in green and then we

30
00:01:30,510 --> 00:01:27,670
have the psycho files in blue and then

31
00:01:32,700 --> 00:01:30,520
piezo files also called viral files in

32
00:01:34,410 --> 00:01:32,710
black and as you may know that the

33
00:01:37,950 --> 00:01:34,420
extreme environments are unstable

34
00:01:39,780 --> 00:01:37,960
environments so that they're likely to

35
00:01:43,920 --> 00:01:39,790
lead to high variability at the genomic

36
00:01:45,990 --> 00:01:43,930
level so on any sort of adaptation or

37
00:01:48,719 --> 00:01:46,000
surviving strategy is likely to have

38
00:01:50,550 --> 00:01:48,729

genomic signatures and here we're

39
00:01:53,639 --> 00:01:50,560
looking at these genomic signatures and

40
00:01:55,560 --> 00:01:53,649
the organisms and part of these genomic

41
00:01:58,080 --> 00:01:55,570
signatures which are part of the bigger

42
00:02:00,660 --> 00:01:58,090
group are LCR's the low-complexity

43
00:02:02,819 --> 00:02:00,670
regions and LCR's have been known to be

44
00:02:04,920 --> 00:02:02,829
evolutionarily significant because they

45
00:02:07,709 --> 00:02:04,930
provide genomes and individual genes

46
00:02:12,240 --> 00:02:07,719
with adjustable turning knobs for

47
00:02:13,920 --> 00:02:12,250
efficient adaptation so previous

48
00:02:16,589 --> 00:02:13,930
research has been done by hearting

49
00:02:18,059 --> 00:02:16,599
Golding and showed that LCR's are not

50
00:02:20,730 --> 00:02:18,069
conserved regions and that they

51
00:02:23,400 --> 00:02:20,740
accumulate mutations so here

52
00:02:24,900 --> 00:02:23,410
number zero we have the LCR's and these

53
00:02:27,090 --> 00:02:24,910
are the number of mutations they

54
00:02:29,760 --> 00:02:27,100
accumulate the proportion of snips the

55
00:02:31,200 --> 00:02:29,770
single nucleotide polymorphisms and the

56
00:02:33,390 --> 00:02:31,210
different colors here and the graph

57
00:02:37,200 --> 00:02:33,400
shows the different conditions they

58
00:02:38,700 --> 00:02:37,210
tested the LCR Xen and genomic

59
00:02:41,010 --> 00:02:38,710
variability has been tested in

60
00:02:43,530 --> 00:02:41,020
eukaryotes and viruses and in the

61
00:02:45,810 --> 00:02:43,540
viruses it was suggested that LCR's are

62
00:02:47,910 --> 00:02:45,820
an important source of genomic

63
00:02:49,530 --> 00:02:47,920

variability for the lengthy virus and

64
00:02:51,960 --> 00:02:49,540
that they're an important source for

65
00:02:55,710 --> 00:02:51,970
antigenic variability in the HIV

66
00:02:58,050 --> 00:02:55,720
populations so what are the low

67
00:03:00,120 --> 00:02:58,060
complexity regions the LCR's are

68
00:03:03,360 --> 00:03:00,130
repetitive sequences that contain few

69
00:03:04,800 --> 00:03:03,370
nucleotides or amino acid types and some

70
00:03:06,840 --> 00:03:04,810
regions contain few different

71
00:03:09,930 --> 00:03:06,850
nucleotides or amino acids while others

72
00:03:12,660 --> 00:03:09,940
only contain one and LCR's have various

73
00:03:14,760 --> 00:03:12,670
configurations ranging from periodic to

74
00:03:16,800 --> 00:03:14,770
a periodic motives and they have

75
00:03:19,020 --> 00:03:16,810
different nomenclatures ranging from low

76
00:03:21,060 --> 00:03:19,030
complexity sequences tandem repeats

77
00:03:23,280 --> 00:03:21,070
simple sequence repeats amino acid

78
00:03:27,560 --> 00:03:23,290
repeats but they all basically share the

79
00:03:31,310 --> 00:03:27,570
same low complex or low diverse sequence

80
00:03:33,690 --> 00:03:31,320
so LCR's are ubiquitous in eukaryotes

81
00:03:35,760 --> 00:03:33,700
about still the function and

82
00:03:38,430 --> 00:03:35,770
evolutionary forces that act on these

83
00:03:40,500 --> 00:03:38,440
regions are unknown and are poorly

84
00:03:42,720 --> 00:03:40,510
understood and here on the right we have

85
00:03:45,270 --> 00:03:42,730
examples of LCR's and different

86
00:03:47,910 --> 00:03:45,280
eukaryotic species so for example in the

87
00:03:50,460 --> 00:03:47,920
homo sapiens the humans we have about

88
00:03:53,010 --> 00:03:50,470

18% and then in the Plasmodium

89
00:03:55,980 --> 00:03:53,020
falciparum the malaria agent we have

90
00:03:59,060 --> 00:03:55,990
about 50% so the proportion of genes

91
00:04:02,870 --> 00:03:59,070
that contain LCR's is highly variable

92
00:04:06,180 --> 00:04:02,880
across the different eukaryotic species

93
00:04:09,990 --> 00:04:06,190
so all the information that we have so

94
00:04:12,300 --> 00:04:10,000
far on LCR's is a result of research on

95
00:04:14,370 --> 00:04:12,310
eukaryotes and here you can see the

96
00:04:15,870 --> 00:04:14,380
eukaryotic tree and the parts that are

97
00:04:17,520 --> 00:04:15,880
marked with a star are the ones that

98
00:04:20,520 --> 00:04:17,530
have been studied for LCR's

99
00:04:22,530 --> 00:04:20,530
and even those form the minority in

100
00:04:25,290 --> 00:04:22,540
comparison to the other eukaryotic

```
101
00:04:27,840 --> 00:04:25,300
species and no one actually has looked

102
00:04:29,700 --> 00:04:27,850
at prokaryotes for LCR's and although

103
00:04:32,370 --> 00:04:29,710
the prokaryotes formed the majority of

104
00:04:34,680 --> 00:04:32,380
the fully sequenced genomes thus they

105
00:04:37,740 --> 00:04:34,690
provide us with a huge set of data

106
00:04:40,410 --> 00:04:37,750
to analyze and work with and prokaryotes

107
00:04:43,110 --> 00:04:40,420
also also form two out of the three

108
00:04:45,450 --> 00:04:43,120
three main domains of life and all the

109
00:04:49,470 --> 00:04:45,460
analysis that has been done so far was

110
00:04:51,930 --> 00:04:49,480
done on model eukaryotic species so in

111
00:04:54,060 --> 00:04:51,940
order to study a full domain we need to

112
00:04:56,820 --> 00:04:54,070
develop new testing methods and analysis

113
00:04:58,650 --> 00:04:56,830
```

in order to automate the process so we

114
00:05:01,860 --> 00:04:58,660
decided to focus on archaea in our

115
00:05:04,170 --> 00:05:01,870
research because archaea has about 200

116
00:05:07,470 --> 00:05:04,180
species while bacteria have 13,000

117
00:05:09,270 --> 00:05:07,480
species so when we study our km we will

118
00:05:11,400 --> 00:05:09,280
be able to develop these new testing

119
00:05:13,620 --> 00:05:11,410
methods and to automate the process in

120
00:05:15,750 --> 00:05:13,630
our future research so what we actually

121
00:05:18,530 --> 00:05:15,760
did we got the fully sequenced genomes

122
00:05:21,600 --> 00:05:18,540
of archaea from ncbi and it was about

123
00:05:24,240 --> 00:05:21,610
221 species and then we did basic

124
00:05:26,640 --> 00:05:24,250
statistics to see the LCR patterns and

125
00:05:28,950 --> 00:05:26,650
trends in the archaea and at that point

126
00:05:33,000 --> 00:05:28,960
we don't really know if LCR's even

127
00:05:34,920 --> 00:05:33,010
exists in archaea so now i'm going to be

128
00:05:38,460 --> 00:05:34,930
talking about the LCR patterns that we

129
00:05:40,770 --> 00:05:38,470
explored in archaea and the LCR

130
00:05:43,290 --> 00:05:40,780
frequency LCR length versus protein

131
00:05:45,810 --> 00:05:43,300
length the LCR frequency versus the GC

132
00:05:49,640 --> 00:05:45,820
content the amino acid composition and

133
00:05:53,070 --> 00:05:49,650
usage and the LCR location the proteins

134
00:05:55,890 --> 00:05:53,080
so to calculate the LCR frequency we

135
00:05:57,840 --> 00:05:55,900
just followed a simple equation a number

136
00:06:00,510 --> 00:05:57,850
of proteins with LCR's over the total

137
00:06:02,730 --> 00:06:00,520
number of proteins times 100 around this

138
00:06:04,800 --> 00:06:02,740

tree you can see the red are the

139
00:06:06,600 --> 00:06:04,810
temperature extremophiles and then the

140
00:06:08,909 --> 00:06:06,610
yellow are the temperature meter files

141
00:06:12,300 --> 00:06:08,919
and the orange bars are the LCR

142
00:06:13,980 --> 00:06:12,310
frequency so on average the healthy our

143
00:06:16,080 --> 00:06:13,990
frequency arranged between three and

144
00:06:18,270 --> 00:06:16,090
seven percent and when we looked

145
00:06:21,300 --> 00:06:18,280
specifically at certain groups for

146
00:06:23,760 --> 00:06:21,310
example the halobacteria the salt loving

147
00:06:25,860 --> 00:06:23,770
extremophiles we saw that they have the

148
00:06:29,670 --> 00:06:25,870
highest LCR frequency which was about

149
00:06:31,650 --> 00:06:29,680
14% and halobacteria else our salt

150
00:06:33,960 --> 00:06:31,660
loving extremophiles they live in salt

151
00:06:38,159 --> 00:06:33,970
extreme environments and the red water

152
00:06:39,900 --> 00:06:38,169
here shows the halobacteria and then

153
00:06:43,620 --> 00:06:39,910
when we looked at other groups such as

154
00:06:45,350 --> 00:06:43,630
the clan RKO de which are temperature

155
00:06:47,879 --> 00:06:45,360
extremophiles and to be exact

156
00:06:50,010 --> 00:06:47,889
hyperthermophiles and they all

157
00:06:52,320 --> 00:06:50,020
so have a high LCR frequency but it's

158
00:06:57,269 --> 00:06:52,330
lower than the halobacteria and there's

159
00:06:59,939 --> 00:06:57,279
range between 10 and 12 percent so we

160
00:07:01,649 --> 00:06:59,949
decide we are compared the LCR frequency

161
00:07:03,869 --> 00:07:01,659
between these groups to see if there is

162
00:07:05,369 --> 00:07:03,879
any significant difference and as you

163
00:07:07,529 --> 00:07:05,379

can see we saw that the halobacteria

164
00:07:10,080 --> 00:07:07,539
have a higher LCR frequency than the

165
00:07:12,570 --> 00:07:10,090
crenarchaeota and they have a higher LCR

166
00:07:14,939 --> 00:07:12,580
frequency than the URI RK Ora and the

167
00:07:17,879 --> 00:07:14,949
URI are Keota in this case includes both

168
00:07:19,980 --> 00:07:17,889
extremophiles and Misa files so the main

169
00:07:22,800 --> 00:07:19,990
question here is is this high LCR

170
00:07:24,689 --> 00:07:22,810
frequency a characteristic of salt

171
00:07:27,059 --> 00:07:24,699
extremophiles or temperature

172
00:07:30,659 --> 00:07:27,069
extremophiles or extremophiles in

173
00:07:33,689 --> 00:07:30,669
general so to answer this question we

174
00:07:35,640 --> 00:07:33,699
did a further analysis comparing the LCR

175
00:07:37,830 --> 00:07:35,650
frequency between halobacteria and

176
00:07:40,230 --> 00:07:37,840
archaea and as you can see we saw that

177
00:07:42,659 --> 00:07:40,240
halobacteria has a higher LCR frequency

178
00:07:44,490 --> 00:07:42,669
than archaea in general and even when we

179
00:07:46,619 --> 00:07:44,500
compared the needs of files and

180
00:07:48,239 --> 00:07:46,629
extremophiles in the URI our Keota we

181
00:07:51,149 --> 00:07:48,249
see that the Mesa files have a higher

182
00:07:53,129 --> 00:07:51,159
LCR frequency than the extremophiles so

183
00:07:55,050 --> 00:07:53,139
this suggests that the high LCR

184
00:07:57,600 --> 00:07:55,060
frequency is a characteristic and D

185
00:07:59,700 --> 00:07:57,610
salt-loving extremophiles and there are

186
00:08:01,860 --> 00:07:59,710
actually a few possible explanations to

187
00:08:04,740 --> 00:08:01,870
this trend that we just saw and one of

188
00:08:07,079 --> 00:08:04,750

them is that LCR's are an emergent

189
00:08:08,879 --> 00:08:07,089
property of proteins meaning longer

190
00:08:10,860 --> 00:08:08,889
proteins will have longer LCR's

191
00:08:13,589 --> 00:08:10,870
and when we compared them we didn't see

192
00:08:15,389 --> 00:08:13,599
this correlation between them so if the

193
00:08:17,670 --> 00:08:15,399
LCR s are not an emergent property of

194
00:08:19,589 --> 00:08:17,680
proteins the LCR's might be driven by

195
00:08:22,529 --> 00:08:19,599
the GC content of the genome whether

196
00:08:24,749 --> 00:08:22,539
it's high or low GC content so previous

197
00:08:27,749 --> 00:08:24,759
research done by TN showed that the LCR

198
00:08:31,019 --> 00:08:27,759
frequency is driven by the 80 content of

199
00:08:32,850 --> 00:08:31,029
the genome in eukaryotes and in our data

200
00:08:34,829 --> 00:08:32,860
as you can see the hello bacteria and

201
00:08:37,019 --> 00:08:34,839
the yellow temperature extremophiles in

202
00:08:40,259 --> 00:08:37,029
blue temperature mesa files in red they

203
00:08:42,089 --> 00:08:40,269
all follow the same trend and when we

204
00:08:43,680 --> 00:08:42,099
did a comparison for significant

205
00:08:46,560 --> 00:08:43,690
differences we saw that the halobacteria

206
00:08:49,290 --> 00:08:46,570
has a higher GC content than all of the

207
00:08:51,990 --> 00:08:49,300
other groups that Grande URI and then

208
00:08:54,060 --> 00:08:52,000
all of the archaea in general and even

209
00:08:57,240 --> 00:08:54,070
the mesa files had a higher GC content

210
00:08:59,460 --> 00:08:57,250
than the extremophiles so that means yes

211
00:09:01,410 --> 00:08:59,470
the patterns of GC and LC are

212
00:09:02,850 --> 00:09:01,420
frequencies are the same

213
00:09:05,870 --> 00:09:02,860

and that they'll see our frequency is

214
00:09:08,430 --> 00:09:05,880
driven by the GC content of the genome

215
00:09:10,740 --> 00:09:08,440
then we looked at the specific amino

216
00:09:13,050 --> 00:09:10,750
acid usage between extremophiles and

217
00:09:15,750 --> 00:09:13,060
mizore files and we saw that they both

218
00:09:18,360 --> 00:09:15,760
use our GC rich amino acids but

219
00:09:21,060 --> 00:09:18,370
extremophiles also use a T rich amino

220
00:09:23,940 --> 00:09:21,070
acids which is expected as we saw they

221
00:09:25,380 --> 00:09:23,950
have a lower GC content every one we

222
00:09:28,110 --> 00:09:25,390
looked specifically at the hello

223
00:09:30,210 --> 00:09:28,120
bacteria we saw that they use similar

224
00:09:33,120 --> 00:09:30,220
amino acids to the miso files and there

225
00:09:36,180 --> 00:09:33,130
are GC rich amino acids and then we

226
00:09:37,890 --> 00:09:36,190
compared in the amino acids usage

227
00:09:39,810 --> 00:09:37,900
between the halobacteria and the urea

228
00:09:41,370 --> 00:09:39,820
Ark iota and we found that they have

229
00:09:43,740 --> 00:09:41,380
different amino acids in their

230
00:09:46,650 --> 00:09:43,750
composition and actually previous

231
00:09:48,720 --> 00:09:46,660
research in eukaryotes showed that the

232
00:09:50,670 --> 00:09:48,730
amino acid composition is species

233
00:09:53,520 --> 00:09:50,680
specific and it's different between the

234
00:09:54,960 --> 00:09:53,530
groups and as you can see these are

235
00:09:57,210 --> 00:09:54,970
three different research done on

236
00:09:59,640 --> 00:09:57,220
vertebrates drosophila and Plasmodium

237
00:10:01,140 --> 00:09:59,650
falciparum and if you look closely all

238
00:10:02,790 --> 00:10:01,150

of them use different amino acid

239
00:10:04,740 --> 00:10:02,800
composition in their LCR's

240
00:10:07,730 --> 00:10:04,750
and the reason actually behind this

241
00:10:10,530 --> 00:10:07,740
difference is still unknown

242
00:10:13,260 --> 00:10:10,540
so up to this point just to summarize we

243
00:10:16,590 --> 00:10:13,270
saw that the LCR's are driven by the GC

244
00:10:18,990 --> 00:10:16,600
content of the genome and the amino acid

245
00:10:21,660 --> 00:10:19,000
composition confirmed this as we saw

246
00:10:23,490 --> 00:10:21,670
that they use GC rich amino acids so

247
00:10:25,200 --> 00:10:23,500
this suggests the neutrality of these

248
00:10:27,990 --> 00:10:25,210
regions and that there are functionless

249
00:10:30,060 --> 00:10:28,000
regions so we decided to look at the LCR

250
00:10:32,840 --> 00:10:30,070
locations in the proteins because this

251
00:10:35,490 --> 00:10:32,850
might be a sign of potential selection

252
00:10:37,860 --> 00:10:35,500
so we followed a previously developed

253
00:10:39,930 --> 00:10:37,870
methods where we divided the protein

254
00:10:42,240 --> 00:10:39,940
into three segments the n-terminal the

255
00:10:44,160 --> 00:10:42,250
middle and then the c-terminal and then

256
00:10:46,260 --> 00:10:44,170
we calculated the difference between the

257
00:10:48,870 --> 00:10:46,270
observed distribution of the middle

258
00:10:51,930 --> 00:10:48,880
point of the LCR's and the expected

259
00:10:54,300 --> 00:10:51,940
distribution to locate where HL c our

260
00:10:56,700 --> 00:10:54,310
sequence would fall and our null

261
00:10:59,220 --> 00:10:56,710
hypothesis suggests that the LCR's are

262
00:11:00,960 --> 00:10:59,230
randomly distributed and our alternative

263
00:11:03,300 --> 00:11:00,970

hypothesis suggests that they are not

264
00:11:05,520 --> 00:11:03,310
randomly distributed and that means that

265
00:11:08,640 --> 00:11:05,530
there is a potential biological function

266
00:11:11,010 --> 00:11:08,650
for them so the results of our

267
00:11:13,200 --> 00:11:11,020
chi-square shows that in extremophiles

268
00:11:15,660 --> 00:11:13,210
we have 82 percent and in the mesial

269
00:11:17,850 --> 00:11:15,670
files we have 75% significant

270
00:11:19,499 --> 00:11:17,860
which means that the LCR's do prefer

271
00:11:22,139 --> 00:11:19,509
certain locations and that they're not

272
00:11:24,689 --> 00:11:22,149
randomly distributed and when we look at

273
00:11:26,699 --> 00:11:24,699
the extremophiles in the Misa files we

274
00:11:29,280 --> 00:11:26,709
see that they both follow the same trend

275
00:11:32,009 --> 00:11:29,290
where the LCR's are located at the

276
00:11:34,049 --> 00:11:32,019
terminal of the proteins and when we

277
00:11:35,999 --> 00:11:34,059
look at the halobacteria specifically we

278
00:11:38,309 --> 00:11:36,009
also see that they follow the same trend

279
00:11:40,559 --> 00:11:38,319
and the LCR's are at the terminal of the

280
00:11:42,769 --> 00:11:40,569
proteins and this actually follows a

281
00:11:45,600 --> 00:11:42,779
previous trend that was found in

282
00:11:47,910 --> 00:11:45,610
research in eukaryotes and this one

283
00:11:50,009 --> 00:11:47,920
specifically in the drosophila where it

284
00:11:51,989 --> 00:11:50,019
shows that the LCR's prefer to be at the

285
00:11:54,780 --> 00:11:51,999
extremities of the protein rather than

286
00:11:57,329 --> 00:11:54,790
in the middle and this suggests that

287
00:11:58,559 --> 00:11:57,339
there might be a potential function for

288
00:12:02,999 --> 00:11:58,569

these regions and that they're not

289
00:12:05,449 --> 00:12:03,009
functionalist regions so do LCR's

290
00:12:08,730 --> 00:12:05,459
actually help extremophiles to survive

291
00:12:11,489 --> 00:12:08,740
well maybe but we do know the LCR's are

292
00:12:13,799 --> 00:12:11,499
a common feature in archaea and that

293
00:12:16,049 --> 00:12:13,809
they exist in salt-loving extremophiles

294
00:12:19,259 --> 00:12:16,059
and a higher frequency than the other

295
00:12:22,860 --> 00:12:19,269
extremophiles we also saw that the LCR's

296
00:12:24,869 --> 00:12:22,870
are are driven by the GC content of the

297
00:12:26,910 --> 00:12:24,879
genome and this was confirmed with the

298
00:12:27,929 --> 00:12:26,920
amino acid usage as we saw that they use

299
00:12:30,840 --> 00:12:27,939
gc-rich

300
00:12:33,360 --> 00:12:30,850
amino acids and we also saw that the

301
00:12:36,389 --> 00:12:33,370
amino acid usage is species specific and

302
00:12:38,429 --> 00:12:36,399
different between the groups so all of

303
00:12:41,340 --> 00:12:38,439
that with the non-random distribution of

304
00:12:44,249 --> 00:12:41,350
the LCR's we can say that the LCR's

305
00:12:45,600 --> 00:12:44,259
maybe creating genomic variability for

306
00:12:47,730 --> 00:12:45,610
the halobacteria the salt-loving

307
00:12:50,040 --> 00:12:47,740
extremophiles specifically so they can

308
00:12:52,590 --> 00:12:50,050
survive and there are extreme

309
00:12:54,030 --> 00:12:52,600
environment but we did not really see

310
00:12:57,329 --> 00:12:54,040
this trend in the temperature

311
00:12:59,249 --> 00:12:57,339
extremophiles so this suggests that they

312
00:13:01,379 --> 00:12:59,259
may be using different adaptive

313
00:13:04,470 --> 00:13:01,389

strategies to survive in their extreme

314
00:13:06,210 --> 00:13:04,480
environments so I'd like to thank

315
00:13:08,669 --> 00:13:06,220
Oakland University central for

316
00:13:11,189 --> 00:13:08,679
biomedical research NASA and IH for

317
00:13:13,410 --> 00:13:11,199
funding this program and I would like to

318
00:13:15,509 --> 00:13:13,420
thank dr. Sheila Sheila Randall Westrick

319
00:13:17,590 --> 00:13:15,519
and all of my lab members for helping me

320
00:13:18,340 --> 00:13:17,600
in this research thank you for attention

321
00:13:19,350 --> 00:13:18,350
[Applause]

322
00:13:21,010 --> 00:13:19,360
[Music]

323
00:13:23,540 --> 00:13:21,020
[Applause]

324
00:13:25,280 --> 00:13:23,550
all right Thank You Anton so we're

325
00:13:27,920 --> 00:13:25,290
running a little bit short on time but

326
00:13:33,200 --> 00:13:27,930
we have time I'm sorry we have time for

327
00:13:33,210 --> 00:13:40,970
thanks for the good talk

328
00:13:48,170 --> 00:13:44,080
my name is Tian from the University of I

329
00:13:50,780 --> 00:13:48,180
was wondering if you looked at the

330
00:13:59,510 --> 00:13:50,790
distribution of other hit archaea

331
00:14:04,700 --> 00:13:59,520
halophiles we looked specifically at the

332
00:14:06,230 --> 00:14:04,710
fully sequenced genomes so these were

333
00:14:09,970 --> 00:14:06,240
the only ones that were available on

334
00:14:12,980 --> 00:14:09,980
NCBI to look for their distribution

335
00:14:16,310 --> 00:14:12,990
they're like two or three other hello

336
00:14:20,000 --> 00:14:16,320
Felix lineages of like really weird GC

337
00:14:23,090 --> 00:14:20,010
content so well there was actually a

338
00:14:25,850 --> 00:14:23,100

paper back in 2008 that explained the

339
00:14:27,770 --> 00:14:25,860
high GC content and the hallow files and

340
00:14:31,640 --> 00:14:27,780
that convergent evolution have to do

341
00:14:34,370 --> 00:14:31,650
with this weird GC content in the hallow