Find me during the conference!

Gathertown: Paula (ELSI)

(mostly Japanese afternoon session)

Discord: Paula Prondzinsky

(message me at any time!)

1
00:00:04,309 --> 00:00:02,710
hello everyone i'm paula and i'm a

2
00:00:06,389 --> 00:00:04,319
second year phd student at the earth

3
00:00:08,230 --> 00:00:06,399
life science institute at tokyo tech i

4
00:00:09,589 --> 00:00:08,240
work in the group of sean mcglenn at lc

5
00:00:11,669 --> 00:00:09,599
and my research interests are in the

6
00:00:13,749 --> 00:00:11,679
fields of extremified microbiology and

7
00:00:15,589 --> 00:00:13,759
computational biology

8
00:00:17,349 --> 00:00:15,599
the title of my talk today is single

9
00:00:18,790 --> 00:00:17,359
cell and genomically resolved analytical

10
00:00:20,630 --> 00:00:18,800
techniques elucidate metabolic

11
00:00:22,550 --> 00:00:20,640
adaptations of methanogenic ikea to

12
00:00:25,029 --> 00:00:22,560
varying temperatures i will walk you

13
00:00:26,870 --> 00:00:25,039

through this step by step

14
00:00:28,390 --> 00:00:26,880
looking at my title the first thing you

15
00:00:30,630 --> 00:00:28,400
might be wondering is why should we even

16
00:00:32,229 --> 00:00:30,640
be interested in methanogens well

17
00:00:34,069 --> 00:00:32,239
methanogens may have been some of the

18
00:00:35,590 --> 00:00:34,079
earliest life forms on earth

19
00:00:37,350 --> 00:00:35,600
looking at some of the major events that

20
00:00:39,430 --> 00:00:37,360
led to life as we know it today starting

21
00:00:41,990 --> 00:00:39,440
with the moon-forming impact 4.5 billion

22
00:00:44,150 --> 00:00:42,000
years ago to ocean formation and the

23
00:00:46,389 --> 00:00:44,160
first isotopic signatures of life 4.1

24
00:00:48,150 --> 00:00:46,399
billion years ago methanogenesis and

25
00:00:49,750 --> 00:00:48,160
sulphate reduction are the earliest

26
00:00:51,990 --> 00:00:49,760
metabolism for which we have direct

27
00:00:53,830 --> 00:00:52,000
isotopic evidence at 3.5 billion years

28
00:00:55,430 --> 00:00:53,840
ago before we can find signs of

29
00:00:58,310 --> 00:00:55,440
photosynthesis and later on the great

30
00:01:00,150 --> 00:00:58,320
oxidation event

31
00:01:02,069 --> 00:01:00,160
looking at the evolution of biological

32
00:01:04,469 --> 00:01:02,079
processes and atmospheric oxygen and

33
00:01:05,910 --> 00:01:04,479
methane on an early earth without oxygen

34
00:01:08,230 --> 00:01:05,920
and an environment that is hostile to

35
00:01:09,750 --> 00:01:08,240
most modern life forms methanogenesis

36
00:01:11,750 --> 00:01:09,760
enabled other non-photosynthetic

37
00:01:13,990 --> 00:01:11,760
prokaryotic metabolisms such as

38
00:01:15,670 --> 00:01:14,000

methanotrophy to thrive which then in

39
00:01:17,350 --> 00:01:15,680
turn pave the way for photosynthetic

40
00:01:20,469 --> 00:01:17,360
primary producers and later on

41
00:01:24,310 --> 00:01:22,870
but what exactly is methanogenesis in

42
00:01:25,670 --> 00:01:24,320
simple terms it is the biological

43
00:01:27,590 --> 00:01:25,680
production of methane through one of

44
00:01:28,950 --> 00:01:27,600
three major pathways

45
00:01:31,109 --> 00:01:28,960
first we have hydrogenotropic

46
00:01:32,710 --> 00:01:31,119
methanogenesis which uses carbon dioxide

47
00:01:34,149 --> 00:01:32,720
as substrate and hydrogenous electron

48
00:01:35,510 --> 00:01:34,159
donor resulting in methane and water

49
00:01:37,510 --> 00:01:35,520
formation

50
00:01:39,350 --> 00:01:37,520
second acetyl-clastic methanogenesis

51
00:01:41,350 --> 00:01:39,360
where acetate is split by oxidizing the

52
00:01:43,590 --> 00:01:41,360
carboxyl group to carbon dioxide and

53
00:01:45,670 --> 00:01:43,600
reducing the methyl group to methane and

54
00:01:47,590 --> 00:01:45,680
third methylotrophic methanogenesis with

55
00:01:48,870 --> 00:01:47,600
methanol regular means as substrates

56
00:01:50,310 --> 00:01:48,880
where electrons are gained through the

57
00:01:52,630 --> 00:01:50,320
oxidation of an additional methyl group

58
00:01:53,910 --> 00:01:52,640
to carbon dioxide so with four methyl

59
00:01:56,230 --> 00:01:53,920
groups as input you will end up with

60
00:01:57,910 --> 00:01:56,240
three molecules of methane

61
00:01:59,590 --> 00:01:57,920
today methanogens are found in various

62
00:02:01,429 --> 00:01:59,600
environments such as animal rumens

63
00:02:03,830 --> 00:02:01,439

c-flow sediments subglacial lakes or

64
00:02:05,910 --> 00:02:03,840
industrial plants and in recent years in

65
00:02:07,830 --> 00:02:05,920
addition to the substrates mentioned

66
00:02:09,990 --> 00:02:07,840
here more and more additional substrates

67
00:02:12,150 --> 00:02:10,000
such as items have been identified for

68
00:02:13,910 --> 00:02:12,160
use by methanogens but that is a

69
00:02:15,350 --> 00:02:13,920
different story

70
00:02:16,710 --> 00:02:15,360
in addition to the white substrate

71
00:02:18,630 --> 00:02:16,720
utilization shown in part on the

72
00:02:20,630 --> 00:02:18,640
previous slide methanogens also have

73
00:02:22,309 --> 00:02:20,640
wide growth temperature ranges

74
00:02:24,229 --> 00:02:22,319
for my project i have classified them

75
00:02:26,070 --> 00:02:24,239
into three temperature groups

```
76
00:02:28,070 --> 00:02:26,080
we have psycho-tolerant organisms with a

77
00:02:29,990 --> 00:02:28,080
minimum growth temperature of 15 degrees

78
00:02:31,750 --> 00:02:30,000
celsius or less which are shown to the

79
00:02:34,150 --> 00:02:31,760
left of or touching the blue line in

80
00:02:35,750 --> 00:02:34,160
this plot thermo-tolerant organisms with

81
00:02:37,990 --> 00:02:35,760
a maximum growth temperature equal to

82
00:02:39,990 --> 00:02:38,000
above 45 degrees celsius which are

83
00:02:41,990 --> 00:02:40,000
either to the right of the red line or

84
00:02:43,670 --> 00:02:42,000
touching the red line and then finally

85
00:02:45,350 --> 00:02:43,680
mesospheric organisms which are all

86
00:02:47,430 --> 00:02:45,360
organisms in between the two lines and

87
00:02:48,949 --> 00:02:47,440
everything touching both lines

88
00:02:51,270 --> 00:02:48,959
```

the total temperature range is from

89
00:02:54,710 --> 00:02:51,280
minus 2.5 degrees celsius observed in an

90
00:02:56,470 --> 00:02:54,720
antarctic lake organism to 120 degrees

91
00:02:58,470 --> 00:02:56,480
celsius observed in an

92
00:02:59,430 --> 00:02:58,480
organism isolated from a black smoker

93
00:03:01,030 --> 00:02:59,440
wall

94
00:03:03,589 --> 00:03:01,040
it is worth mentioning here that there

95
00:03:05,270 --> 00:03:03,599
are no universal definitions for

96
00:03:07,190 --> 00:03:05,280
like tolerant thermotolerant or

97
00:03:08,710 --> 00:03:07,200
thermophilic temperature cutoffs and the

98
00:03:10,470 --> 00:03:08,720
thresholds here were set based on the

99
00:03:12,470 --> 00:03:10,480
overall distributions within this group

100
00:03:13,990 --> 00:03:12,480
of organisms

101
00:03:16,390 --> 00:03:14,000
now let's have a look at the phylogeny

102
00:03:18,630 --> 00:03:16,400
of the organisms i define methanogens as

103
00:03:21,110 --> 00:03:18,640
every archaem containing methyl coenzyme

104
00:03:23,830 --> 00:03:21,120
and reductase short mcr alpha beta and

105
00:03:26,229 --> 00:03:23,840
gamma subunits which are a total of 295

106
00:03:27,589 --> 00:03:26,239
organisms this here is a subset of those

107
00:03:29,030 --> 00:03:27,599
that have temperature data available

108
00:03:30,869 --> 00:03:29,040
which is a set i'm showing in the

109
00:03:33,509 --> 00:03:30,879
following analyses as well

110
00:03:34,869 --> 00:03:33,519
there are a total of 86 organisms the

111
00:03:36,550 --> 00:03:34,879
shapes in the tree indicate the

112
00:03:39,110 --> 00:03:36,560
different substrates where the collard

113
00:03:40,390 --> 00:03:39,120

boxes represent the temperature classes

114
00:03:42,390 --> 00:03:40,400
we can see that temperatures and

115
00:03:44,390 --> 00:03:42,400
substrates are distributed throughout

116
00:03:45,990 --> 00:03:44,400
the phylogenetic tree and that the only

117
00:03:47,190 --> 00:03:46,000
real trend that we can observe is that

118
00:03:49,910 --> 00:03:47,200
there are no acidic clastic

119
00:03:51,350 --> 00:03:49,920
psycho-tolerant organisms

120
00:03:53,190 --> 00:03:51,360
when seeing these wide temperature

121
00:03:55,030 --> 00:03:53,200
ranges you might be wondering what

122
00:03:56,949 --> 00:03:55,040
mechanisms allow methanogens to inhabit

123
00:03:59,110 --> 00:03:56,959
such a wide range of environments and

124
00:04:00,789 --> 00:03:59,120
how can we find those mechanisms in my

125
00:04:01,910 --> 00:04:00,799
thesis work i'm looking at methanogens

126
00:04:03,990 --> 00:04:01,920
genomes with respect to their

127
00:04:05,509 --> 00:04:04,000
composition function and structure

128
00:04:07,270 --> 00:04:05,519
this is what i'm going to talk about for

129
00:04:08,949 --> 00:04:07,280
the rest of this presentation

130
00:04:10,789 --> 00:04:08,959
i'm also interested in the physiological

131
00:04:12,710 --> 00:04:10,799
adaptations and i'm doing single cell

132
00:04:14,710 --> 00:04:12,720
analysis of organisms grown at different

133
00:04:16,150 --> 00:04:14,720
temperatures to investigate to what

134
00:04:18,710 --> 00:04:16,160
extent individual cells in the

135
00:04:20,469 --> 00:04:18,720
population differ from one another

136
00:04:21,430 --> 00:04:20,479
so let's have a look at the genome data

137
00:04:24,070 --> 00:04:21,440
and if you're interested in the

138
00:04:26,230 --> 00:04:24,080

physiological physiology part please

139
00:04:27,909 --> 00:04:26,240
contact me later

140
00:04:30,150 --> 00:04:27,919
i started by splitting the genome of

141
00:04:31,749 --> 00:04:30,160
each of the 86 species into their core

142
00:04:33,590 --> 00:04:31,759
and pan genomes

143
00:04:35,350 --> 00:04:33,600
the core is shared by all organisms in

144
00:04:37,270 --> 00:04:35,360
this analysis and appears in blue in the

145
00:04:38,710 --> 00:04:37,280
third while the pen genome consists of

146
00:04:41,110 --> 00:04:38,720
genes that are not present in every

147
00:04:42,870 --> 00:04:41,120
single organism but in some of them

148
00:04:45,110 --> 00:04:42,880
the pendulum can further be divided into

149
00:04:46,629 --> 00:04:45,120
unique genes that only appear in exactly

150
00:04:48,870 --> 00:04:46,639
one species and are in green here in

151
00:04:50,230 --> 00:04:48,880
this plot and shared genes which appear

152
00:04:51,909 --> 00:04:50,240
in at least two but not all of the

153
00:04:53,110 --> 00:04:51,919
species and which are the orange

154
00:04:54,950 --> 00:04:53,120
fraction

155
00:04:57,110 --> 00:04:54,960
in this plot each bar represents one

156
00:04:59,909 --> 00:04:57,120
organism and the core of all of the

157
00:05:02,550 --> 00:04:59,919
organisms consists of 225 genes

158
00:05:03,990 --> 00:05:02,560
which in size of the whole genome is

159
00:05:06,230 --> 00:05:04,000
between 5

160
00:05:08,870 --> 00:05:06,240
and 15

161
00:05:10,710 --> 00:05:08,880
that is actually relatively small

162
00:05:12,310 --> 00:05:10,720
so where are the adaptive mechanisms

163
00:05:14,070 --> 00:05:12,320

that we are looking for are they in the

164
00:05:15,590 --> 00:05:14,080
core and the pen genome

165
00:05:18,629 --> 00:05:15,600
we next have a look at the composition

166
00:05:21,029 --> 00:05:18,639
and structure of proteins of each core

167
00:05:23,189 --> 00:05:21,039
shared and unique gene groups to find

168
00:05:25,350 --> 00:05:23,199
out more since we cannot see a clear

169
00:05:27,350 --> 00:05:25,360
trend in the phylogeny from the previous

170
00:05:29,029 --> 00:05:27,360
slide or the genome composition on this

171
00:05:30,550 --> 00:05:29,039
slide

172
00:05:32,870 --> 00:05:30,560
i will first show you amino acid

173
00:05:34,710 --> 00:05:32,880
compositions of each of the fractions we

174
00:05:36,390 --> 00:05:34,720
are looking at amino acid substitutions

175
00:05:37,990 --> 00:05:36,400
because they cause structural changes

176
00:05:40,070 --> 00:05:38,000
the proteins that can cause an increase

177
00:05:41,749 --> 00:05:40,080
of decrease in flexibility and stability

178
00:05:43,590 --> 00:05:41,759
of the protein which is important in

179
00:05:45,749 --> 00:05:43,600
different temperature environments

180
00:05:47,430 --> 00:05:45,759
so for example cyclophilic proteins have

181
00:05:49,430 --> 00:05:47,440
a higher flexibility and lower thermal

182
00:05:50,870 --> 00:05:49,440
stability compared to their mesophilic

183
00:05:52,629 --> 00:05:50,880
counterparts

184
00:05:54,230 --> 00:05:52,639
we found that the observed substitutions

185
00:05:56,309 --> 00:05:54,240
follow the expectations that we get from

186
00:05:58,309 --> 00:05:56,319
the literature with the conserved core

187
00:05:59,830 --> 00:05:58,319
and variable unique proteins

188
00:06:01,430 --> 00:05:59,840

so let me walk you through each of the

189
00:06:03,830 --> 00:06:01,440
factions one by one starting with the

190
00:06:05,350 --> 00:06:03,840
overall genome composition

191
00:06:06,870 --> 00:06:05,360
this plot shows the differences in

192
00:06:09,189 --> 00:06:06,880
thermo-tolerant and psycho-tolerant

193
00:06:11,029 --> 00:06:09,199
species with respect to individual amino

194
00:06:12,710 --> 00:06:11,039
acids calculated from the overall mean

195
00:06:14,469 --> 00:06:12,720
appearance of an amino acid through all

196
00:06:15,990 --> 00:06:14,479
thermo-tolerant cyclotron species

197
00:06:17,909 --> 00:06:16,000
respectively

198
00:06:20,309 --> 00:06:17,919
a positive value means that the amino

199
00:06:21,830 --> 00:06:20,319
acid is more than cycle tolerant forms

200
00:06:23,350 --> 00:06:21,840
while a negative value means that the

201
00:06:25,590 --> 00:06:23,360
amino acid is more abundant in

202
00:06:28,150 --> 00:06:25,600
thermo-tolerant organisms

203
00:06:30,469 --> 00:06:28,160
looking at all genes together the core

204
00:06:31,590 --> 00:06:30,479
plus pen genomes so this is the black

205
00:06:33,830 --> 00:06:31,600
bar here

206
00:06:36,150 --> 00:06:33,840
we see some differences which are not

207
00:06:37,830 --> 00:06:36,160
too pronounced the main differences are

208
00:06:41,110 --> 00:06:37,840
higher glutamic acid and lysine and

209
00:06:43,270 --> 00:06:41,120
thermals and higher serine and uranium

210
00:06:45,110 --> 00:06:43,280
and psychotron species

211
00:06:46,830 --> 00:06:45,120
but those differences are mostly within

212
00:06:49,830 --> 00:06:46,840
the one percent

213
00:06:52,390 --> 00:06:49,840

range next let's look at the individual

214
00:06:54,150 --> 00:06:52,400
genome components starting with the core

215
00:06:55,430 --> 00:06:54,160
the differences are similar to that of

216
00:06:57,430 --> 00:06:55,440
the overall

217
00:06:58,790 --> 00:06:57,440
genome even a bit less pronounced than

218
00:07:00,390 --> 00:06:58,800
in the whole genome

219
00:07:02,629 --> 00:07:00,400
the trend for the shared fraction of the

220
00:07:06,309 --> 00:07:02,639
genome is very similar as well

221
00:07:08,790 --> 00:07:06,319
so finally let's look at the uniques and

222
00:07:10,950 --> 00:07:08,800
here we can see that the differences are

223
00:07:12,710 --> 00:07:10,960
more pronounced than the other fractions

224
00:07:15,029 --> 00:07:12,720
that we looked at before

225
00:07:16,870 --> 00:07:15,039
when we compare all of this data side by

226
00:07:18,150 --> 00:07:16,880
side we can see that indeed the unique

227
00:07:20,390 --> 00:07:18,160
fraction has the most pronounced

228
00:07:21,990 --> 00:07:20,400
differences in amino acid composition

229
00:07:25,830 --> 00:07:22,000
sometimes even a completely opposite

230
00:07:27,350 --> 00:07:25,840
trend as the other genome fractions

231
00:07:29,830 --> 00:07:27,360
let's see if we can observe a similar

232
00:07:31,430 --> 00:07:29,840
trend in protein domain architectures

233
00:07:33,189 --> 00:07:31,440
the architectures will tell us about the

234
00:07:35,350 --> 00:07:33,199
orientation of secondary structures of

235
00:07:37,670 --> 00:07:35,360
protein faults the core here is very

236
00:07:39,909 --> 00:07:37,680
very conserved which you can see in the

237
00:07:41,510 --> 00:07:39,919
absence of the blue bars

238
00:07:43,110 --> 00:07:41,520

there are some differences in the shared

239
00:07:44,710 --> 00:07:43,120
fraction but the most pronounced

240
00:07:45,990 --> 00:07:44,720
differences again and the unique

241
00:07:47,830 --> 00:07:46,000
fraction

242
00:07:49,670 --> 00:07:47,840
and in addition to this plot this is not

243
00:07:52,710 --> 00:07:49,680
displayed here but

244
00:07:55,350 --> 00:07:52,720
one important thing to note is that when

245
00:07:56,230 --> 00:07:55,360
it comes to the mapping coverage of the

246
00:08:00,469 --> 00:07:56,240
genes

247
00:08:02,309 --> 00:08:00,479
mapped only a small fraction of them had

248
00:08:03,749 --> 00:08:02,319
an architecture mapping onto them which

249
00:08:05,430 --> 00:08:03,759
means that they need more attention if

250
00:08:07,830 --> 00:08:05,440
we want to find out more about their

251
00:08:10,550 --> 00:08:07,840
structure and function

252
00:08:12,070 --> 00:08:10,560
so finally what does this data tell us

253
00:08:13,749 --> 00:08:12,080
basically it tells us that the core is

254
00:08:15,670 --> 00:08:13,759
very conserved when it comes to amino

255
00:08:17,670 --> 00:08:15,680
acid composition meaning flexibility and

256
00:08:19,029 --> 00:08:17,680
rigidity and protein domain

257
00:08:20,390 --> 00:08:19,039
architectures meaning secondary

258
00:08:21,909 --> 00:08:20,400
structures

259
00:08:23,430 --> 00:08:21,919
we know very little about the unique

260
00:08:24,869 --> 00:08:23,440
genes even though some of them might

261
00:08:27,430 --> 00:08:24,879
hold keys to the thermal adaptation

262
00:08:29,350 --> 00:08:27,440
mechanisms of methanogens

263
00:08:31,749 --> 00:08:29,360

so if you want to find out more about

264
00:08:33,509 --> 00:08:31,759
those we will need molecular biology

265
00:08:35,750 --> 00:08:33,519
studies

266
00:08:38,389 --> 00:08:35,760
and then finally what is next for me

267
00:08:40,949 --> 00:08:38,399
i have a bit more data to go through in

268
00:08:42,310 --> 00:08:40,959
this project such as gene phylogenies

269
00:08:43,909 --> 00:08:42,320
for the core genes from which we can

270
00:08:46,389 --> 00:08:43,919
relate the connection between

271
00:08:48,150 --> 00:08:46,399
temperatures and evolutionary rates and

272
00:08:50,150 --> 00:08:48,160
i'm also interested in the hypothesis of

273
00:08:52,230 --> 00:08:50,160
thermal reduction which was first

274
00:08:53,190 --> 00:08:52,240
introduced by forte more than 20 years

275
00:08:54,949 --> 00:08:53,200
ago

276
00:08:56,870 --> 00:08:54,959
as i mentioned in the beginning i'm also

277
00:08:58,710 --> 00:08:56,880
looking at the physiology of individual

278
00:09:00,710 --> 00:08:58,720
cells through laboratory experiments and

279
00:09:03,030 --> 00:09:00,720
ultimately my plan is to link the genome

280
00:09:04,389 --> 00:09:03,040
potential with the observed phenotypic

281
00:09:06,070 --> 00:09:04,399
variations

282
00:09:07,829 --> 00:09:06,080
if you're interested in any of this i'd

283
00:09:10,550 --> 00:09:07,839
be happy to meet up virtually during or

284
00:09:12,389 --> 00:09:10,560
after the conference as well

285
00:09:14,710 --> 00:09:12,399
on this slide i've just listed some of

286
00:09:16,630 --> 00:09:14,720
the references i used for my figures if

287
00:09:18,389 --> 00:09:16,640
you're interested you can read further

288
00:09:19,990 --> 00:09:18,399

here

289
00:09:22,070 --> 00:09:20,000
thank you for watching my presentation

290
00:09:23,829 --> 00:09:22,080
until the end you can find me during the

291
00:09:25,829 --> 00:09:23,839
conference either on gather town during

292
00:09:28,070 --> 00:09:25,839
the japanese afternoon session or

293
00:09:29,590 --> 00:09:28,080
message me through discord at any time