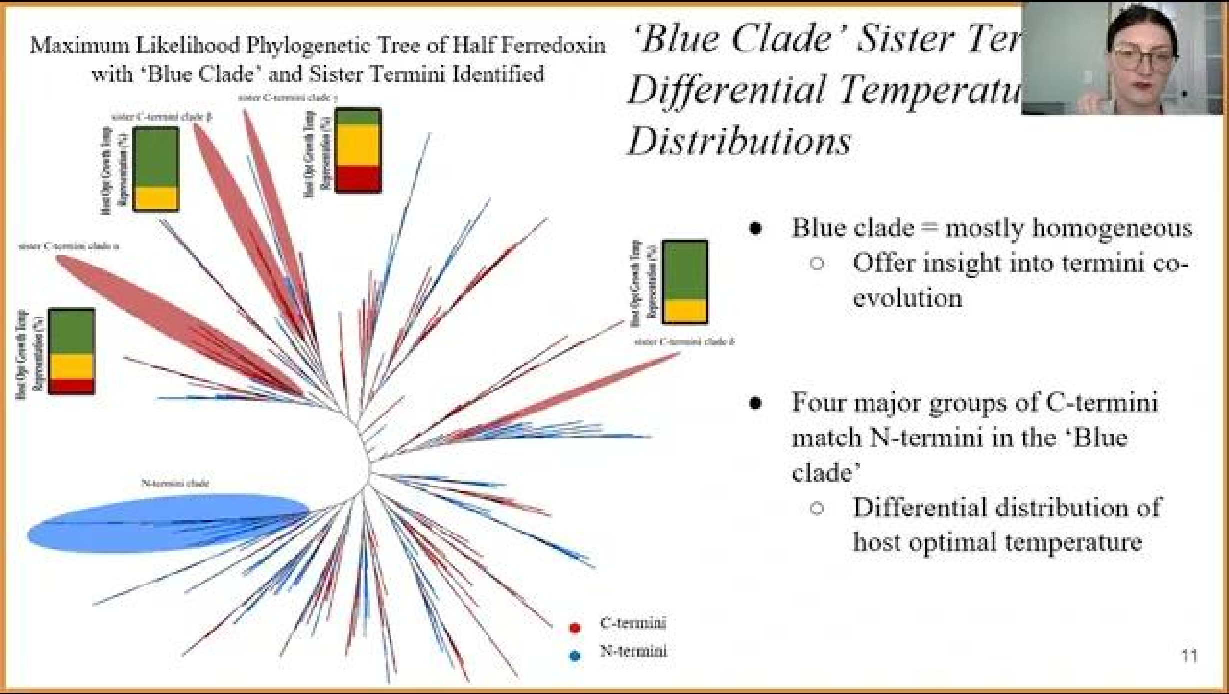

Maximum Likelihood Phylogenetic Tree of Half Ferredoxin with 'Blue Clade' and Sister Termini Identified

'Blue Clade' Sister Ter
Differential Temperatu
Distributions

- Blue clade = mostly homogeneous
  - Offer insight into termini co-evolution
- Four major groups of C-termini match N-termini in the 'Blue clade'
  - Differential distribution of host optimal temperature

1
00:00:05,670 --> 00:00:03,270
hello

2
00:00:06,389 --> 00:00:05,680
uh good whatever time it is wherever you

3
00:00:08,629 --> 00:00:06,399
are

4
00:00:09,910 --> 00:00:08,639
uh my name is julianna de giacomo i work

5
00:00:11,270 --> 00:00:09,920
with brock lewis university

6
00:00:13,110 --> 00:00:11,280
and i'm going to be discussing the

7
00:00:15,669 --> 00:00:13,120
environment as a driving force in the

8
00:00:17,670 --> 00:00:15,679
coevolution of pharadox and termini

9
00:00:19,830 --> 00:00:17,680
but first we are going to take a step

10
00:00:22,870 --> 00:00:19,840
backwards to really contextualize

11
00:00:25,029 --> 00:00:22,880
the significance of this claim

12
00:00:26,870 --> 00:00:25,039
biological redox reactions are catalyzed

13
00:00:29,189 --> 00:00:26,880

in nature by a set of enzymes called

14
00:00:31,509 --> 00:00:29,199
oxoduct cases this grouping can be

15
00:00:32,790 --> 00:00:31,519
separated into six major subsets which i

16
00:00:35,030 --> 00:00:32,800
have included here

17
00:00:36,229 --> 00:00:35,040
each with really diverse redox

18
00:00:38,470 --> 00:00:36,239
functionalities

19
00:00:39,990 --> 00:00:38,480
the origin of oxoroductase is likely

20
00:00:40,790 --> 00:00:40,000
parallels with the origins of life

21
00:00:43,110 --> 00:00:40,800
itself

22
00:00:43,990 --> 00:00:43,120
for the actual path of this evolution is

23
00:00:46,229 --> 00:00:44,000
largely unknown

24
00:00:47,830 --> 00:00:46,239
because of the immense structural and

25
00:00:49,029 --> 00:00:47,840
functional diversity that i've outlined

26
00:00:51,830 --> 00:00:49,039
here

27
00:00:53,510 --> 00:00:51,840
harold 2014 suggests 10 distinct origins

28
00:00:56,150 --> 00:00:53,520
for oxidative death cases where

29
00:00:57,430 --> 00:00:56,160
all iron-containing axoreductases share

30
00:00:59,670 --> 00:00:57,440
a single origin

31
00:01:01,830 --> 00:00:59,680
this is especially significant as this

32
00:01:02,470 --> 00:01:01,840
group constitutes nearly 50 percent of

33
00:01:05,509 --> 00:01:02,480
all

34
00:01:07,510 --> 00:01:05,519
metal containing oxidoreductases

35
00:01:09,109 --> 00:01:07,520
so now we are just focusing on

36
00:01:10,070 --> 00:01:09,119
understanding the origins of these

37
00:01:12,469 --> 00:01:10,080
iron-containing

38
00:01:14,230 --> 00:01:12,479

proteins there are numerous protein

39
00:01:15,030 --> 00:01:14,240
folds which contain these iron hosing

40
00:01:16,950 --> 00:01:15,040
domains

41
00:01:19,270 --> 00:01:16,960
and the most ancient of which likely

42
00:01:21,350 --> 00:01:19,280
looked like excellent ferridoxin

43
00:01:23,190 --> 00:01:21,360
ferrodoxins are a subset of electrons

44
00:01:24,149 --> 00:01:23,200
transfer proteins which contain iron

45
00:01:27,030 --> 00:01:24,159
sulfur clusters

46
00:01:27,830 --> 00:01:27,040
stabilized by assisting binding motif

47
00:01:30,310 --> 00:01:27,840
therefore

48
00:01:31,590 --> 00:01:30,320
the study of the origin of ferredoxin

49
00:01:37,190 --> 00:01:31,600
serves as a proxy

50
00:01:41,590 --> 00:01:39,910
paradoxines function using a critical

51
00:01:43,749 --> 00:01:41,600
iron sulfur cluster

52
00:01:44,630 --> 00:01:43,759
life originated on earth when the levels

53
00:01:46,710 --> 00:01:44,640
of oxygen

54
00:01:47,830 --> 00:01:46,720
was limited to completely absent which

55
00:01:50,310 --> 00:01:47,840
permitted metals

56
00:01:51,590 --> 00:01:50,320
including iron to exist as soluble

57
00:01:54,550 --> 00:01:51,600
compounds in the ocean

58
00:01:57,270 --> 00:01:54,560
such as these iron silver clusters which

59
00:01:58,870 --> 00:01:57,280
really could have formed spontaneously

60
00:02:00,550 --> 00:01:58,880
also supporting the antiquity of

61
00:02:02,469 --> 00:02:00,560
ferridoxin uh

62
00:02:03,590 --> 00:02:02,479
is that the feradoxan binding fold is

63
00:02:05,910 --> 00:02:03,600

highly evolvable

64
00:02:08,710 --> 00:02:05,920
derived from the low hydrogen bond

65
00:02:09,749 --> 00:02:08,720
energy per residue this along with the

66
00:02:13,030 --> 00:02:09,759
fact

67
00:02:15,670 --> 00:02:13,040
that uh paradoxes are also ubiquitous

68
00:02:17,430 --> 00:02:15,680
uh to all life supports that feradoxin

69
00:02:20,390 --> 00:02:17,440
is also very ancient

70
00:02:22,150 --> 00:02:20,400
in terms of how this feradoxin evolves

71
00:02:24,150 --> 00:02:22,160
it is accepted that at one point

72
00:02:26,150 --> 00:02:24,160
the gene of some ancient iron sulfur

73
00:02:28,710 --> 00:02:26,160
redox protein was duplicated

74
00:02:30,470 --> 00:02:28,720
leading to a symmetric protoferidoxin

75
00:02:33,750 --> 00:02:30,480
which i've outlined at the bottom

76
00:02:35,990 --> 00:02:33,760
of this slide however

77
00:02:37,350 --> 00:02:36,000
accent ferrodoxins are asymmetric iron

78
00:02:39,910 --> 00:02:37,360
sulfur proteins

79
00:02:42,229 --> 00:02:39,920
at this point each half could have

80
00:02:43,750 --> 00:02:42,239
acquired random mutations which led to

81
00:02:46,869 --> 00:02:43,760
modern feradoxin

82
00:02:47,830 --> 00:02:46,879
or the evolution of each half could have

83
00:02:50,790 --> 00:02:47,840
been driven by

84
00:02:52,309 --> 00:02:50,800
some external force previous

85
00:02:55,430 --> 00:02:52,319
phylogenetic analysis

86
00:02:55,990 --> 00:02:55,440
um with my lab of the amino acid

87
00:02:59,350 --> 00:02:56,000
sequence

88
00:03:01,509 --> 00:02:59,360

of each independent terminus of

89
00:03:04,630 --> 00:03:01,519
ferrodoxin displays significant sequence

90
00:03:07,670 --> 00:03:04,640
of conservation unique to each

91
00:03:09,110 --> 00:03:07,680
half this suggests that the two halves

92
00:03:11,030 --> 00:03:09,120
of ferrodoxin

93
00:03:13,030 --> 00:03:11,040
evolved independently under the

94
00:03:15,430 --> 00:03:13,040
influence of some external factors

95
00:03:16,869 --> 00:03:15,440
that specifically drove certain sequence

96
00:03:20,309 --> 00:03:16,879
changes to be conserved

97
00:03:21,910 --> 00:03:20,319
in each half thus the objectives of the

98
00:03:23,750 --> 00:03:21,920
study is to

99
00:03:25,509 --> 00:03:23,760
directly to understand the external

100
00:03:27,190 --> 00:03:25,519
forces which drove independent

101
00:03:30,470 --> 00:03:27,200
coevolution of ferradoxin

102
00:03:32,309 --> 00:03:30,480
and therefore oxo reductases but more

103
00:03:34,869 --> 00:03:32,319
broadly we hope to identify

104
00:03:35,350 --> 00:03:34,879
the basic factors of electron transfer

105
00:03:37,190 --> 00:03:35,360
and

106
00:03:38,390 --> 00:03:37,200
contribute to establishing the possible

107
00:03:41,430 --> 00:03:38,400
conditions for

108
00:03:43,110 --> 00:03:41,440
the origins of life

109
00:03:44,470 --> 00:03:43,120
of course central the study is the

110
00:03:46,550 --> 00:03:44,480
hypothesis

111
00:03:48,630 --> 00:03:46,560
in this study we suppose that following

112
00:03:50,470 --> 00:03:48,640
a gene duplication event environment

113
00:03:52,229 --> 00:03:50,480

environmental factors drove the

114
00:03:52,869 --> 00:03:52,239
independent co-evolution of each

115
00:03:55,350 --> 00:03:52,879
terminus

116
00:03:57,589 --> 00:03:55,360
of ancient paradoxin we specifically

117
00:03:59,110 --> 00:03:57,599
chose to start by looking at the impacts

118
00:04:00,550 --> 00:03:59,120
of temperature and ph

119
00:04:02,869 --> 00:04:00,560
because not only are these some of the

120
00:04:05,030 --> 00:04:02,879
most simple factors of environment

121
00:04:06,470 --> 00:04:05,040
but consistent with uh the hypothesis

122
00:04:09,110 --> 00:04:06,480
that life originated at

123
00:04:11,030 --> 00:04:09,120
a deep sea hydrothermal vent temperature

124
00:04:12,550 --> 00:04:11,040
and ph are two characteristics of the

125
00:04:14,390 --> 00:04:12,560
environment of early life

126
00:04:15,750 --> 00:04:14,400
which would have drastically changed as

127
00:04:19,749 --> 00:04:15,760
life diversified

128
00:04:21,509 --> 00:04:19,759
from that point and moving forward

129
00:04:23,510 --> 00:04:21,519
for the study it was most advantageous

130
00:04:24,950 --> 00:04:23,520
for us to focus on the most ancient

131
00:04:27,350 --> 00:04:24,960
group of paradoxes

132
00:04:29,990 --> 00:04:27,360
and we find this in four iron first

133
00:04:34,469 --> 00:04:30,000
sulfur cluster feradox and specifically

134
00:04:37,590 --> 00:04:34,479
in methanogenic archaeal host organisms

135
00:04:39,270 --> 00:04:37,600
so i'm moving on to methods and results

136
00:04:41,030 --> 00:04:39,280
in order to analyze independent

137
00:04:41,830 --> 00:04:41,040
coevolution we had to split each

138
00:04:44,469 --> 00:04:41,840

feradoxin

139
00:04:44,870 --> 00:04:44,479
into its respective termini we did this

140
00:04:47,990 --> 00:04:44,880
with

141
00:04:49,749 --> 00:04:48,000
a novel methodology to my group by

142
00:04:52,469 --> 00:04:49,759
locating the halfway point

143
00:04:53,590 --> 00:04:52,479
in protein three structures and

144
00:04:56,230 --> 00:04:53,600
comparing this

145
00:04:58,150 --> 00:04:56,240
to an alignment with all relevant

146
00:04:59,350 --> 00:04:58,160
ferridoxin sequences which were analyzed

147
00:05:01,749 --> 00:04:59,360
in this study

148
00:05:02,790 --> 00:05:01,759
this created an n-terminus and

149
00:05:05,749 --> 00:05:02,800
c-terminus for

150
00:05:06,390 --> 00:05:05,759
every protein finally we arranged these

151
00:05:08,310 --> 00:05:06,400
termini

152
00:05:09,749 --> 00:05:08,320
into maximum likelihood phylogenetic

153
00:05:11,990 --> 00:05:09,759
trees which

154
00:05:15,430 --> 00:05:12,000
allowed us to analyze the evolutionary

155
00:05:16,629 --> 00:05:15,440
relationships between these proteins

156
00:05:18,390 --> 00:05:16,639
here you can see the initial

157
00:05:19,670 --> 00:05:18,400
phylogenetic tree which was generated

158
00:05:22,950 --> 00:05:19,680
from the aligned

159
00:05:24,469 --> 00:05:22,960
paradox in termini the same termini

160
00:05:26,230 --> 00:05:24,479
are generally clustered together

161
00:05:28,390 --> 00:05:26,240
indicating that they are similar

162
00:05:29,749 --> 00:05:28,400
across different proteins as you can see

163
00:05:32,710 --> 00:05:29,759

in the bottom highlighted

164
00:05:33,350 --> 00:05:32,720
clade however you're um above this you

165
00:05:34,710 --> 00:05:33,360
can see

166
00:05:38,390 --> 00:05:34,720
that there are several clades that are

167
00:05:41,270 --> 00:05:38,400
mixed showing that the termini still

168
00:05:44,230 --> 00:05:41,280
display a level of sequence similarity

169
00:05:46,629 --> 00:05:44,240
within the same proteins

170
00:05:47,990 --> 00:05:46,639
when we consider host optimal growth

171
00:05:50,230 --> 00:05:48,000
temperature data

172
00:05:51,990 --> 00:05:50,240
it appears that in the instances where

173
00:05:54,230 --> 00:05:52,000
thermophile and hypothermophile

174
00:05:55,830 --> 00:05:54,240
sequences appear to be deeply branching

175
00:05:57,110 --> 00:05:55,840
in the clade which means they are closer

176
00:05:59,430 --> 00:05:57,120
to the origin

177
00:06:01,110 --> 00:05:59,440
the sequence is typically from the c

178
00:06:02,469 --> 00:06:01,120
terminus of ferridoxin which i've

179
00:06:05,350 --> 00:06:02,479
outlined here

180
00:06:07,430 --> 00:06:05,360
you can also see that the thermophile

181
00:06:08,629 --> 00:06:07,440
and hypothermophile hosts are generally

182
00:06:10,390 --> 00:06:08,639
clustered together

183
00:06:11,830 --> 00:06:10,400
whether they are at the closer to the

184
00:06:14,790 --> 00:06:11,840
origin or

185
00:06:16,469 --> 00:06:14,800
uh elsewhere in the tree unfortunately

186
00:06:19,990 --> 00:06:16,479
no such

187
00:06:25,990 --> 00:06:20,000
correlations existed when data ph data

188
00:06:31,029 --> 00:06:28,790

further investigation was focused on a

189
00:06:34,469 --> 00:06:31,039
large clade that was almost

190
00:06:37,189 --> 00:06:34,479
entirely composed of feradox and termini

191
00:06:37,590 --> 00:06:37,199
this is highlighted in blue on the slide

192
00:06:43,270 --> 00:06:37,600
the

193
00:06:47,430 --> 00:06:43,280
formed four distinct clusters

194
00:06:49,430 --> 00:06:47,440
the sister alpha beta gamma and delta

195
00:06:50,870 --> 00:06:49,440
while there are no differential

196
00:06:54,150 --> 00:06:50,880
distribution of ph

197
00:06:55,350 --> 00:06:54,160
among these sister clades we do see a

198
00:06:59,350 --> 00:06:55,360
pretty

199
00:06:59,990 --> 00:06:59,360
clear-cut distribution with temperature

200
00:07:02,390 --> 00:07:00,000
where

201
00:07:04,790 --> 00:07:02,400
especially sister-clad gamma is really

202
00:07:07,670 --> 00:07:04,800
dominated by hypothermophiles and

203
00:07:12,790 --> 00:07:10,870
to investigate further multiple sequence

204
00:07:15,189 --> 00:07:12,800
alignments were compared between the

205
00:07:16,790 --> 00:07:15,199
termini of each sister clade

206
00:07:18,309 --> 00:07:16,800
all alignments were visualized as

207
00:07:21,350 --> 00:07:18,319
sequence logos

208
00:07:23,510 --> 00:07:21,360
um visualized on the slide here each

209
00:07:23,830 --> 00:07:23,520
sister clade contained residues which

210
00:07:25,909 --> 00:07:23,840
were

211
00:07:26,870 --> 00:07:25,919
differentially present but the most

212
00:07:29,909 --> 00:07:26,880
notable

213
00:07:31,270 --> 00:07:29,919

was in sister clay gamma which is

214
00:07:33,350 --> 00:07:31,280
dominated by thermophiles and

215
00:07:35,510 --> 00:07:33,360
hypothermophiles i've mentioned before

216
00:07:36,870 --> 00:07:35,520
which contains a unique and really

217
00:07:40,390 --> 00:07:36,880
heavily conserved valine

218
00:07:44,309 --> 00:07:40,400
at position 5 within this termini

219
00:07:47,430 --> 00:07:44,319
and a tripeptide segment

220
00:07:49,029 --> 00:07:47,440
so now what does it all mean uh the real

221
00:07:50,230 --> 00:07:49,039
takeaway from the study is that these

222
00:07:52,790 --> 00:07:50,240
findings suggest that

223
00:07:53,749 --> 00:07:52,800
the paradox in termini did likely

224
00:07:55,670 --> 00:07:53,759
co-evolve

225
00:07:57,510 --> 00:07:55,680
independently where temperature was one

226
00:07:57,990 --> 00:07:57,520
of the driving factors for residue

227
00:08:00,790 --> 00:07:58,000
change

228
00:08:01,189 --> 00:08:00,800
in each terminus and ph as far as we

229
00:08:04,070 --> 00:08:01,199
know

230
00:08:05,110 --> 00:08:04,080
does not appear to be in such a direct

231
00:08:07,830 --> 00:08:05,120
way

232
00:08:09,029 --> 00:08:07,840
the c-termini sequence also appears to

233
00:08:11,909 --> 00:08:09,039
be more ancient

234
00:08:14,230 --> 00:08:11,919
than the n-termini and perhaps reflects

235
00:08:16,309 --> 00:08:14,240
the protoferidoxin which underwent a

236
00:08:19,110 --> 00:08:16,319
gene duplication originally

237
00:08:20,950 --> 00:08:19,120
within the c-termini a valine near the

238
00:08:23,189 --> 00:08:20,960

beginning of the sequence

239
00:08:27,270 --> 00:08:23,199
appears to be critical to thermophile

240
00:08:30,790 --> 00:08:27,280
and hypothermophile ferrodoxin

241
00:08:31,670 --> 00:08:30,800
um we also know that an ancient

242
00:08:34,389 --> 00:08:31,680
ferrodoxin

243
00:08:35,829 --> 00:08:34,399
likely evolved near the origins of life

244
00:08:37,909 --> 00:08:35,839
and is diversified

245
00:08:39,990 --> 00:08:37,919
into all iron-containing proteins

246
00:08:42,550 --> 00:08:40,000
including 50 percent of all

247
00:08:43,509 --> 00:08:42,560
metal containing oxidoreduct cases

248
00:08:45,750 --> 00:08:43,519
therefore

249
00:08:47,750 --> 00:08:45,760
this investigation also suggests that

250
00:08:48,230 --> 00:08:47,760
environmental temperature was a driving

251
00:08:51,269 --> 00:08:48,240
factor

252
00:08:53,590 --> 00:08:51,279
in the ancient evolution of life's

253
00:08:55,509 --> 00:08:53,600
really earliest oxorbectases and

254
00:08:56,630 --> 00:08:55,519
proteins that were capable of electron

255
00:08:58,870 --> 00:08:56,640
transfer

256
00:09:00,710 --> 00:08:58,880
as thermophile and hypothermia file

257
00:09:03,350 --> 00:09:00,720
sequences appear to be

258
00:09:05,350 --> 00:09:03,360
deeply branching the study also supports

259
00:09:06,470 --> 00:09:05,360
that life's earliest oxo reductases

260
00:09:08,949 --> 00:09:06,480
developed

261
00:09:10,389 --> 00:09:08,959
in organisms with extremely hot

262
00:09:13,030 --> 00:09:10,399
environments

263
00:09:14,710 --> 00:09:13,040

these and uh observations aligned with

264
00:09:18,389 --> 00:09:14,720
previous studies

265
00:09:20,389 --> 00:09:18,399
um on small subunit rna and of course

266
00:09:23,829 --> 00:09:20,399
the hypothesized origin of life at

267
00:09:26,710 --> 00:09:23,839
superheated hydrothermal vents

268
00:09:27,430 --> 00:09:26,720
and that is all i have for you today um

269
00:09:30,470 --> 00:09:27,440
i would like to

270
00:09:31,430 --> 00:09:30,480
acknowledge my uh head of my lab dr

271
00:09:33,509 --> 00:09:31,440
vikas nanda

272
00:09:35,110 --> 00:09:33,519
and everyone else in the nonda lab the

273
00:09:38,550 --> 00:09:35,120
nod to dream team

274
00:09:40,710 --> 00:09:38,560
and everyone else who

275
00:09:41,990 --> 00:09:40,720
helped with collecting source materials