DATA SUPPLEMENT

A machine learning strategy for gut microbiome-based diagnostic screening of cardiovascular disease

Sachin Aryal, Ahmad Alimadadi, Ishan Manandhar, Bina Joe* and Xi Cheng*

Bioinformatics & Artificial Intelligence Laboratory, Center for Hypertension and Precision Medicine, Program in Physiological Genomics, Department of Physiology and Pharmacology, University of Toledo College of Medicine and Life Sciences, Toledo, OH 43614, USA

Running title: AI and microbiome for CVD diagnosis

***Corresponding Authors**

Xi Cheng, Ph.D.

Bioinformatics & Artificial Intelligence Laboratory

Center for Hypertension and Precision Medicine

Program in Physiological Genomics

Department of Physiology and Pharmacology

University of Toledo College of Medicine and Life Sciences

Block Health Science Bldg. Rm 320, 3000 Arlington Ave., Toledo, OH 43614, USA

Phone: 419-383-4076 Email: Xi.Cheng@utoledo.edu

Bina Joe, Ph.D., FAHA, FCVS (APS)

Bioinformatics & Artificial Intelligence Laboratory

Center for Hypertension and Precision Medicine

Program in Physiological Genomics

Department of Physiology and Pharmacology

University of Toledo College of Medicine and Life Sciences

Block Health Science Bldg. Rm 237, 3000 Arlington Ave., Toledo, OH 43614, USA

Phone: 419-383-4144 Email: bina.joe@utoledo.edu

Supplementary Table S1. Thirty-nine differential bacterial taxa and their LDA scores.

Taxonomic Features	Enriched Group	LDA Score
k__Bacteria.p__Proteobacteria.c__Betaproteobacteria.o__Neisseriales.f__Neisseriaceae.g__Neisseria	non-cvd	2.22
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Veillonellaceae.g__Acidaminococcus	cvd	2.29
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria.o__RF32.f__	non-cvd	2.60
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Peptostreptococcaceae	non-cvd	2.58
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Ruminococcaceae.g__Ruminococcus	non-cvd	3.32
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Veillonellaceae.g__Veillonella	cvd	2.62
k__Bacteria.p__Proteobacteria.c__Betaproteobacteria.o__Neisseriales.f__Neisseriaceae	non-cvd	2.26
k__Bacteria.p__Cyanobacteria.c__4C0d_2.o__YS2	non-cvd	2.55
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria.o__RF32.f__.g__	non-cvd	2.61
k__Bacteria.p__Cyanobacteria.c__4C0d_2.o__YS2.f__.g__	non-cvd	2.56
k__Bacteria.p__Firmicutes.c__Bacilli.o__Bacillales.f__Listeriaceae	cvd	2.24
k__Bacteria.p__Proteobacteria.c__Gammaproteobacteria.o__Enterobacteriales.f__Enterobacteriaceae.g__Proteus	non-cvd	3.17
k__Bacteria.p__Firmicutes.c__Erysipelotrichi.o__Erysipelotrichales.f__Erysipelotrichaceae.g__	non-cvd	2.72
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Clostridiaceae.g__	non-cvd	2.23
k__Bacteria.p__Firmicutes.c__Erysipelotrichi.o__Erysipelotrichales.f__Erysipelotrichaceae.g___Eubacterium_	cvd	2.64
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria.o__RF32	non-cvd	2.62
k__Bacteria.p__Bacteroidetes.c__Bacteroidia.o__Bacteroidales.f___Paraprevotellaceae_	non-cvd	2.75
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Ruminococcaceae.g__Faecalibacterium	non-cvd	3.67
k__Bacteria.p__Proteobacteria.c__Betaproteobacteria.o__Neisseriales	non-cvd	2.25
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria.o__Caulobacterales.f__Caulobacteraceae.g__Brevundimonas	non-cvd	2.90
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria.o__Caulobacterales.f__Caulobacteraceae	non-cvd	2.87
k__Bacteria.p__Firmicutes.c__Bacilli.o__Bacillales.f__Listeriaceae.g__Listeria	cvd	2.25
k__Bacteria.p__Bacteroidetes.c__Bacteroidia.o__Bacteroidales.f__Bacteroidaceae	cvd	4.23

k__Bacteria.p__Bacteroidetes.c__Bacteroidia.o__Bacteroidales.f__Bacteroidaceae.g__Bacteroides	cvd	4.23
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Ruminococcaceae.g__Subdoligranulum	cvd	3.35
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria	non-cvd	3.19
k__Bacteria.p__Bacteroidetes.c__Bacteroidia.o__Bacteroidales.f__Rikenellaceae.g__Alistipes	non-cvd	2.86
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Veillonellaceae.g__Megasphaera	cvd	2.77
k__Bacteria.p__Cyanobacteria	non-cvd	2.55
k__Bacteria.p__Bacteroidetes.c__Bacteroidia.o__Bacteroidales.f__Rikenellaceae	non-cvd	3.12
k__Bacteria.p__Cyanobacteria.c__4C0d_2	non-cvd	2.56
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Lachnospiraceae.g__Lachnospira	non-cvd	2.94
k__Bacteria.p__Firmicutes.c__Clostridia.o__Clostridiales.f__Lachnospiraceae.g__Clostridium	cvd	2.99
k__Bacteria.p__Cyanobacteria.c__4C0d_2.o__YS2.f__	non-cvd	2.55
k__Bacteria.p__Lentisphaerae.c__.o__.f__	non-cvd	2.83
k__Bacteria.p__Lentisphaerae.c__.o__.f__.g__	non-cvd	2.65
k__Bacteria.p__Lentisphaerae.c__.o__	non-cvd	2.82
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria.o__Caulobacterales	non-cvd	2.85
k__Bacteria.p__Lentisphaerae.c__	non-cvd	3.08

Supplementary Table S2. Top 100 highly contributing OTU features and their importance scores.

OTU ID	Highly Contributing OTU	Importance Score
1	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGACGGGCAAGTCTGAAGTGAAAGGCAGGGGCT CAACCCCTGGACTGCTTTGGAAACTGTCCATCTAGAGTGCCGGAGAGGTAAGCGGAATTCCTAG	98.60
2	TACGTAGGAGGCAAGCGTTATCCGGAATGACTGGGCGTAAAGGGTGCGTAGGTGGTTCAGCAAGTTGGCAGCGTAACTCCGGGGCTC AACCTCGGCACTACTGCCAAAACTGTCGAACTTGAGTGCAGGAGGGGCAATCGGAATTCCTAG	64.05
3	TACGTAGGGAGCGAGCGTTATCCGGATTTATTGGGTGTAAAGGGTGCGTAGACGGGAAGTCAAGTTAGTTGTGAAATCCCTCGGCTTA ACTGAGGAACTGCAACTAAAACTGATTTTCTTGAGTACTGGAGAGGAAAGTGGAATTCCTAG	63.04
4	TACGTAGGGTGCAAGCGTTATCCGGAATTATTGGGCGTAAAGGGCTCGTAGGCGGTTCGTCGCGTCCGGTGTGAAAGTCCATCGCTTA ACGGTGGATCCGCGCCGGGTACGGGCGGGCTTGAGTGCGGTAGGGGAGACTGGAATTCCCGG	58.30
5	TACGGAGGATTCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGTTTGATAAGTTAGAGGTGAAATTTCGGGGCTCA ACCCTGAACGTGCCTCTAATACTGTTGAGCTAGAGAGTAGTTGCGGTAGGCGGAATGTATGG	58.27
6	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGATGGATGTTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGATATCTTGAGTGCAGTTGAGGCAGGCGGAATTCGTGG	58.12
7	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGACAGTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGCTGTCTTGAGTACAGTAGAGGTGGGCGGAATTCGTGG	56.95
8	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGATTGTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGAAACTGGCAGTCTTGAGTACAGTAGAGGTGGGCGGAATTCGTGG	55.83
9	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATGGACAAGTCTGATGTGAAAGGCTGGGGCTC AACCCCGGGACTGCATTGGAAACTGCCCGTCTTGAGTGCCGGAGAGGTAAGCGGAATTCCTAG	55.14
10	AACGTAGGTGGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGACCGGCAAGTTGGAAGTGAAATCTATGGGCTC AACCCATAAATTGCTTTCAAAACTGCTGGCCTTGAGTAGTGCAGAGGTAGGCGGAATTCCCGG	53.78
11	TACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTTTGTTAAGTCAGATGTGAAATCCCCGGGCTCA ACCTGGGAACTGCATCTGATACTGGCAAGCTTGAGTCTCGTAGAGGGGGGTAGAATTCCAGG	53.17
12	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGATTGTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGAAACTGGGAGTCTTGAGTACAGTAGAGGTGGGCGGAATTCGTGG	52.74
13	TACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGGGCGCGCAGGCGGCATCGCAAGTCGGTCTTAAAAGTGCGGGGCTT AACCCCGTGAGGGGACCGAAACTGTGAAGCTCGAGTGTCGGAGAGGAAAGCGGAATTCCTAGT	45.47
14	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGACGCTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGGTGTCTTGAGTACAGTAGAGGCAGGCGGAATTCGTGG	44.77
15	TACGTAGGGTGCAAGCGTTATCCGGAATTATTGGGCGTAAAGGGCTCGTAGGCGGTTCGTCGCGTCCGGTGTGAAAGTCCATCGCTTA ACGGTGGATCTGCGCCGGGTACGGGCGGGCTGGAGTGCGGTAGGGGAGACTGGAATTCCCGG	43.71
16	TACGTAGGGAGCAAGCGTTATCCGGATTTATTGGGTGTAAAGGGTGCGTAGACGGGACAACAAGTTAGTTGTGAAATCCCTCGGCTTA ACTGAGGAACTGCAACTAAAACTATTGTTCTTGAGTGTTGGAGAGGAAAGTGGAATTCCTAG	43.65

17	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGGGTGCGCAGGCGGTTGAGTAAGACAGATGTGAAATCCCCGAGCTT AACTCGGGAATGGCATATGTGACTGCTCGACTAGAGTGTGTCAGAGGGAGGTGGAATTCCACG	42.95
18	TACGTAGGGGGCTAGCGTTATCCGGATTTACTGGGCGTAAAGGGTGCGTAGGCGGTCTTTTAAGTCAGGAGTGAAAGGCTACGGCTCA ACCGTAGTAAGCTCTTGAAACTGGAGGACTTGAGTGCAGGAGAGGAGAGTGGAATTCCTAGT	40.96
19	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTGTGGCAAGTCTGATGTGAAAGGCATGGGCTC AACCTGTGGACTGCATTGGAAACTGTCATACTTGAGTGCCGGAGGGGTAAGCGGAATTCCTAG	40.27
20	AACGTAGGGTGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAAGACAAGTTGGAAGTGAAAACCATGGGCT CAACCCATGAATTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGATGGAATTCCCGG	38.22
21	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCACGGCAAGCCAGATGTGAAAGCCCGGGGCTC AACCCCGGGACTGCATTTGGAACTGCTGAGCTAGAGTGTCGGAGAGGCAAGTGGAATTCCTAG	37.70
22	TACGTAGGTGGCGAGCGTTATCCGGAATTATTGGGCGTAAAGAGGGAGCAGGCGGCACTAAGGGTCTGTGGTGAAAGATCGAAGCTT AACTTCGGTAAGCCATGGAAACCGTAGAGCTAGAGTGTGTGAGAGGATCGTGGAATTCCATGT	37.51
23	AACGTAGGTCACAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAAGACAAGTTGGAAGTGAAATCTATGGGCTC AACCCATAAACTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGGCGGAATTCCCGG	36.67
24	TACGTAGGGGGCAAGCGTTATCCGGAATTACTGGGTGTAAAGGGTGCGTAGGTGGTATGGCAAGTCAGAAGTGAAAACCCAGGGCTT AACTCTGGGACTGCTTTTGAAACTGTCAGACTGGAGTGCAGGAGAGGTAAGCGGAATTCCTAG	36.03
25	TACGTAGGGAGCGAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGATGGCAAGTCAGATGTGAAAACTATGGGCTC AACCCATAGACTGCATTTGAAACTGTTGTTCTTGAGTGAGGTAGAGGTAAGCGGAATTCCTGG	35.73
26	AACGTAGGTCACAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAGAACAAGTTGGAAGTGAAATCCATGGGCTC AACCCATGAACTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGGCGGAATTCCCGG	34.95
27	TACGGAGGATGCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGTGGTGATTTAAGTCAGCGGTGAAAGTTTGTGGCTCA ACCATAAAATTGCCGTTGAAACTGGGTTACTTGAGTGTGTTTGAGGTAGGCGGAATGCGTGG	34.91
28	AACGTAGGGTGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGACCGGCAAGTTGGAAGTGAAAACTATGGGCTC AACCCATAAATTGCTTTCAAAACTGCTGGCCTTGAGTAGTGCAGAGGTAGGTGGAATTCCCGG	34.72
29	TACGGAGGATGCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGCAGACGGGAGATTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGTTTCCTTGAGTGCAGTTGAGGCAGGCGGAATTCGTGG	34.51
30	TACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGTGCGGCAAGTCTGATGTGAAAGCCCGGGGCTC AACCCCGGTACTGCATTGGAAACTGTCGTACTAGAGTGTCGGAGGGGTAAGCGGAATTCCTAG	34.29
31	TACGTAGGGAGCGAGCGTTATCCGGAATTATTGGGTGTAAAGGGTGCGTAGACGGGAGAACAAGTTAGTTGTGAAATACCTCGGCTCA ACTGAGGAACTGCAACTAAAACTGTACTTCTTGAGTGCAGGAGAGGTAAGTGGAATTCCTAG	33.97
32	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCAAGGCAAGTCTGATGTGAAAACCCAGGGCTT AACCCTGGGACTGCATTGGAAACTGTCTGGCTCGAGTGCCGGAGAGGTAAGCGGAATTCCTAG	33.75
33	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGACTGGCAAGTCTGATGTGAAAGGCGGGGGCTC AACCCCTGGACTGCATTGGAAACTGTTAGTCTTGAGTGCCGGAGAGGTAAGCGGAATTCCTAG	33.68
34	TACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGATTGGTCAGTCTGTCTTAAAAGTTCGGGGCTTA ACCCCGTGATGGGATGGAAACTGCCAATCTAGAGTATCGGAGAGGAAAGTGGAATTCCTAGT	33.32

35	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGCCTTTTAAGTCAGCGGTGAAAGTCTGTGGCTCA ACCATAGAATTGCCGTTGAAACTGGGGGGCTTGAGTATGTTTGAGGCAGGCGGAATGCGTGG	33.28
36	TACGTAGGGGGCTAGCGTTATCCGGAATTACTGGGCGTAAAGGGTGCGTAGGTGGTTTCTTAAGTCAGAGGTGAAAGGCTACGGCTCA ACCGTAGTAAGCCTTTGAAACTGGGAAACTTGAGTGCAGGAGAGGAGAGTGGAATTCCTAGT	33.18
37	TACGGAGGGTGCAAGCGTTAATCGGAATCACTGGGCGTAAAGCGCACGTAGGCGGCTTGGTAAGTCAGGGGTGAAATCCCACAGCCC AACTGTGGAACTGCCTTTGATACTGCCAGGCTTGAGTACCGGAGAGGGTGGCGGAATTCCAGG	32.52
38	TACGGAGGATGCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGTGGTTAATTAAGTCAGCGGTGAAAGTTTGTGGCTCA ACCATAAAATTGCCGTTGAAACTGGTTGACTTGAGTATATTTGAGGTAGGCGGAATGCGTGG	32.52
39	TACGTAGGTCCCGAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCA ACCATAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGT	31.85
40	TACGTATGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGTGTAGGTGGCCATGCAAGTCAGAAGTGAAAATCCGGGGCTC AACCCCGGAACTGCTTTTGAAACTGTAAGGCTGGAGTGCAGGAGGGGTGAGTGGAATTCCTAG	31.38
41	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGCAGCAAGTCTGATGTGAAAGGCAGGGGCTT AACCCCTGGACTGCATTGGAAACTGCTGTGCTTGAGTGCCGGAGGGGTAAGCGGAATTCCTAG	31.20
42	TACAGAGGTCTCAAGCGTTGTTCGGAATCACTGGGCGTAAAGCGTGCGTAGGCTGTTTCGTAAGTCGTGTGTGAAAGGCGCGGGCTCA ACCCGCGGACGGCACATGATACTGCGAGACTAGAGTAATGGAGGGGGAACCGGAATTCTCGG	30.87
43	TACAGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGTGGTTTGTTAAGTTGAATGTGAAATCCCCGGGCTCA ACCTGGGAACTGCATCCAAAACTGGCAAGCTAGAGTATGGTAGAGGGTAGTGGAATTTCCTG	30.38
44	TACGTAGGGAGCAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGTGCGTAGGCGGCTTTGCAAGTCAGATGTGAAATCTATGGGCTCA ACCCATAAACTGCATTTGAAACTGTAGAGCTTGAGTGAAGTAGAGGCAGGCGGAATTCCCCG	29.68
45	TACGGAGGATTCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGTTTGATAAGTTAGAGGTGAAATCCCGGGGCTTA ACTCCGGAACTGCCTCTAATACTGTTAGACTAGAGAGTAGTTGCGGTAGGCGGAATGTATGG	29.62
46	TACGTATGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGTGTAGGTGGCCAGGCAAGTCAGAAGTGAAAGCCCGGGGCTC AACCCCGGGACTGCTTTTGAAACTGCAGGGCTAGAGTGCAGGAGGGGCAAGTGGAATTCCTAG	29.38
47	TACGTAGGGGGCAAGCGTTATCCGGAATTACTGGGTGTAAAGGGTGCGTAGGTGGTATGGCAAGTCAGAAGTGAAAACCCAGGGCTT AACTCTGGGACTGCTTTTGAAACTGTCAGACTAGAGTGCAGGAGAGGTAAGCGGAATTCCTAG	28.85
48	AACGTAGGTCACAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGGAAGACAAGTTGGAAGTGAAATCCATGGGCTC AACCCATGAACTGCTTTCAAAACTGTTTTTCTTGAGTAGTGCAGAGGTAGGCGGAATTCCCGG	28.00
49	TACGTAGGGGGCGAGCGTTATCCGGATTCATTGGGCGTAAAGCGCGCGTAGGCGGCCCGGCAGGCCGGGGGTCGAAGCGGGGGGCT CAACCCCCCGAAGCCCCCGGAACCTCCGCGGCTTGGGTCCGGTAGGGGAGGGTGGAACACCCGG	27.13
50	TACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGCAGCCGGGTCTGCAAGTCAGATGTGAAATCCATGGGCTCA ACCCATGAACTGCATTTGAAACTGTAGATCTTGAGTGTCGGAGGGGCAATCGGAATTCCTAG	27.03
51	TACGTATGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGTGCGTAGGTGGCAGTGCAAGTCAGATGTGAAAGGCCGGGGCTC AACCCCGGAGCTGCATTTGAAACTGCATAGCTAGAGTACAGGAGAGGCAGGCGGAATTCCTAG	26.99
52	TACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGCGGTCTGTCAAGTCGGATGTGAAATCCCCGGGCTC AACCTGGGAACTGCATTCGAAACTGGCAGGCTAGAGTCTTGTAGAGGGGGGTAGAATTCCAGG	26.83

53	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGAAGCAAGTCTGAAGTGAAAACCCAGGGCTC AACCCTGGGACTGCTTTGGAAACTGTTTTGCTAGAGTGTCGGAGAGGTAAGTGGAATTCCTAG	26.41
54	TACGTATGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGTGCGTAGGTGGCAGTGCAAGTCAGATGTGAAAGGCCGGGGCTC AACCCCGGAGCTGCATTTGAAACTGCTCGGCTAGAGTACAGGAGAGGCAGGCGGAATTCCTAG	26.40
55	TACGTAGGTGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGGATTGCAAGTCAGATGTGAAAACTGGGGGCTC AACCTCCAGCCTGCATTTGAAACTGTAGTTCTTGAGTGCTGGAGAGGCAATCGGAATTCCGTG	26.00
56	TACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGGTGGCAAGGCAAGCCAGAAGTGAAAACCCGGGGCTC AACCGCGGGATTGCTTTTGGAACTGTCATGCTAGAGTGCAGGAGGGGTGAGCGGAATTCCTAG	25.94
57	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGAATGGCAAGTCTGATGTGAAAGGCCGGGGCTC AACCCCGGGACTGCATTGGAAACTGTCAATCTAGAGTACCGGAGGGGTAAGTGGAATTCCTAG	25.88
58	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGTGGACTGGTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGTCAGTCTTGAGTACAGTAGAGGTGGGCGGAATTCGTGG	25.74
59	TACGTAGGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGCGCAGGCGGGCCGGCAAGTTGGAAGTGAAATCTATGGGCTT AACCCATAAACTGCTTTCAAAACTGCTGGTCTTGAGTGATGGAGAGGCAGGCGGAATTCCGTG	25.53
60	TACGTAGGGGGCGAGCGTTGTCCGGAATGATTGGGCGTAAAGGGCGCGTAGGCGGCCTGCTAAGTCTGGAGTGAAAGTCCTGCTTTC AAGGTGGGAATTGCTTTGGATACTGGTGGGCTGGAGTGCAGGAGAGGAAAGCGGAATTACCGG	25.36
61	TACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCTTCCCAAGTCCCTCTTAAAAGTGCGGGGCTTA ACCCCGTGATGGGAAGGAAACTGGGAAGCTGGAGTATCGGAGAGGAAAGTGGAATTCCTAGT	25.32
62	TACGGAGGGTGCAAGCGTTGTCCGGAATCATTGGGCGTAAAGAGTTCGTAGGTGGTTTGTTAAGTTTGGTGTTAAATGCAGAGGCTCA ACTTCTGTTCGGCATCGGATACTGGCAGACTAGAATGCGGTAGAGGTAAAGGGAATTCCTGG	24.99
63	TACGTATGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGTGTAGGTGGCCATGCAAGTCAGAAGTGAAAATCCGGGGCTC AACCCCGGAACTGCTTTTGAAACTGTAAGGCTAGAGTGCAGGAGGGGTGAGTGGAATTCCTAG	24.92
64	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGCATGATAAGTCTGATGTGAAAACCCAAGGCTC AACCATGGGACTGCATTGGAAACTGTCGTGCTGGAGTGTCGGAGAGGTAAGCGGAATTCCTAG	24.88
65	TACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCA ACCGGGGAGGGTCATTGGAAACTGGGAGACTTGAGTGCAGAAGAGGAGAGTGGAATTCCATG	24.83
66	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGGTTGTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGCGACCTTGAGTGCAACAGAGGTAGGCGGAATTCGTGG	24.65
67	TACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCTGTGTAAGTCTGAAGTGAAAGCCCGGGGCTCA ACCCCGGGACTGCTTTGGAAACTATGCAGCTAGAGTGTCGGAGAGGTAAGTGGAATTCCCAG	24.40
68	TACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGATGGGCAAGTCTGATGTGAAAACCCGGGGCTC AACCCCGGGACTGCATTGGAAACTGTTCATCTAGAGTGCTGGAGAGGTAAGTGGAATTCCTAG	24.36
69	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGATTGTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGCAGTCTTGAGTGCAGTAGAGGTGGGCGGAATTCGTGG	24.23
70	TACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGGCGGAGAAGCAAGTCAGAAGTGAAATCCATGGGCTT AACCCATGAACTGCTTTTGAAACTGTTTCCCTTGAGTATCGGAGAGGCAGGCGGAATTCCTAG	24.21

71	AACGTAGGGTGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGACCGGCAAGTTGGAAGTGAAATCCATGGGCTC AACCCGTGAATTGCTTTCAAAACTGCTGGCCTTGAGTAGTGCAGAGGTAGGTGGAATTCCCGG	24.09
72	TACGGAGGATCCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGTTTAGTAAGTCAGCGGTGAAATTTTGGTGCTTAA CACCAAACGTGCCGTTGATACTGCTGGGCTAGAGAGTAGTTGCGGTAGGCGGAATGTATGG	23.91
73	TACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGCAGGCGGAAGGCTAAGTCTGATGTGAAAGCCCGGGGCTC AACCCCGGTACTGCATTGGAAACTGGTCATCTAGAGTGTCGGAGGGGTAAGTGGAATTCCTAG	23.69
74	TACGGAGGATCCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGTTTGATAAGTTAGAGGTGAAATACCGGGGCTCA ACTCCGGAACTGCCTCTAATACTGTTGAACTAGAGAGTAGTTGCGGTAGGCGGAATGTATGG	23.48
75	TACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGCCGGGAAGGCAAGTCAGATGTGAAATCCACGGGCTC AACTCGTGAACTGCATTTGAAACTGTTTTTCTTGAGTATCGGAGAGGCAATCGGAATTCCTAG	23.31
76	TACGTAGGTGGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGTGTAGGCGGGATATCAAGTCAGAAGTGAAAATTACGGGCTC AACTCGTAACCTGCTTTTGAAACTGACATTCTTGAGTGAAGTAGAGGCAAGCGGAATTCCTAG	23.07
77	TACGTATGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGGCGGCATGGCAAGTCAGAAGTGAAAGCCTGGGGCTC AACCCCGGAATTGCTTTTGAAACTGTCAGGCTAGAGTGTCGGAGGGGTAAGCGGAATTCCTAG	23.05
78	TACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGTAGCCGGGAGGGCAAGTCAGATGTGAAATCCACGGGCTC AACTCGTGAACTGCATTTGAAACTACTCTTCTTGAGTATCGGAGAGGCAATCGGAATTCCTAG	23.02
79	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGATGGATGTTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGATGTCTTGAGTGCAGTTGAGGCAGGCGGAATTCGTGG	23.01
80	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGACGCAAGTCTGAAGTGAAATACCCGGGCTC AACCTGGGAACTGCTTTGGAAACTGTGTTGCTAGAGTGCTGGAGAGGTAAGCGGAATTCCTAG	23.01
81	AACGTAGGGTGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGCAGGCGGATTGGCAAGTTGGGAGTGAAATCTATGGGCTC AACCCATAAATTGCTTTCAAAACTGTCAGTCTTGAGTGGTGTAGAGGTAGGCGGAATTCCCGG	22.98
82	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCTGTGCAAGTCTGAAGTGAAAGGCATGGGCTC AACCTGTGGACTGCTTTGGAAACTGTGCAGCTAGAGTGTCGGAGAGGTAAGTGGAATTCCTAG	22.93
83	TACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGCAGCCGGGCATGCAAGTCAGATGTGAAATCTCAGGGCTTA ACCCTGAAACTGCATTTGAAACTGTATGTCTTGAGTGCCGGAGAGGTAATCGGAATTCCTTG	22.76
84	TACGTAGGTGGCGAGCGTTGTCCGGAATTATTGGGCGTAAAGAGCATGTAGGCGGCTTAATAAGTCGAGCGTGAAAATGCGGGGCTC AACCCCGTATGGCGCTGGAAACTGTTAGGCTTGAGTGCAGGAGAGGAAAGGGGAATTCCCAGT	22.75
85	TACGTAGGTGACAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGCGTAGGCGGACTGTCAAGTCAGTCGTGAAATACCGGGGCTTA ACCCCGGGGCTGCGATTGAAACTGACAGCCTTGAGTATCGGAGAGGAAAGCGGAATTCCTAG	22.63
86	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGGCGGACTATTAAGTCAGCTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAGTTGATACTGGTCGTCTTGAGTGCAGTAGAGGTAGGCGGAATTCGTGG	22.54
87	TACGTAGGTCCCGAGCGTTATCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTAGATAAGTCTGAAGTTAAAGGCTGTGGCTTA ACCATAGTACGCTTTGGAAACTGTTTAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGT	22.34
88	TACGTATGGTGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGTGCGTAGGTGGCAAGGCAAGTCTGAAGTGAAAATCCGGGGCTCA ACCCCGGAACTGCTTTGGAAACTGTTTAGCTGGAGTACAGGAGAGGTAAGTGGAATTCCTAG	22.26

89	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCATGGCAAGCCAGATGTGAAAGCCCGGGGCTC AACCCCGGGACTGCATTTGGAACTGTCAGGCTAGAGTGTCGGAGAGGAAAGCGGAATTCCTAG	22.23
90	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGAGCGTAGATGGGTTGTTAAGTCAGTTGTGAAAGTTTGCGGCTCA ACCGTAAAATTGCAATTGATACTGGCAGTCTTGAGTACAGTTGAGGTAGGCGGAATTCGTGG	22.14
91	TACGGAGGGTGCGAGCGTTAATCGGAATAACTGGGCGTAAAGGGCACGCAGGCGGTGACTTAAGTGAGGTGTGAAAGCCCCGGGCTT AACCTGGGAATTGCATTTCATACTGGGTCGCTAGAGTACTTTAGGGAGGGGTAGAATTCCACG	21.94
92	TACGTAGGTGGCAAGCGTTGTCCGGATTTACTGGGTGTAAAGGGCGTGCAGCCGGGCCGGCAAGTCAGATGTGAAATCTGGAGGCTT AACCTCCAAACTGCATTTGAAACTGTAGGTCTTGAGTACCGGAGAGGTTATCGGAATTCCTTG	21.93
93	TACGGAGGATGCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGCACGCCAAGTCAGCGGTGAAATTTTCGGGCTCA ACCCGGACTGTGCCGTTGAAACTGGCGAGCTAGAGTGCACAAGAGGCAGGCGGAATGCGTGG	21.87
94	TACGTATGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGTGTAGGTGGTATCACAAGTCAGAAGTGAAAGCCCGGGGCTC AACCCCGGGACTGCTTTTGAAACTGTGGAACTGGAGTGCAGGAGAGGTAAGTGGAATTCCTAG	21.71
95	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGCCTTTTAAGTCAGCGGTGAAAGTCTGTGGCTCA ACCATAGAATTGCCGTTGAAACTGGGAGGCTTGAGTATGTTTGAGGCAGGCGGAATGCGTGG	21.65
96	TACGTAGGGGGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGTTTTGCAAGTCTGAAGTGAAAGCCCGGGGCTT AACCCCGGGACTGCTTTGGAAACTGTAGGACTAGAGTGCAGGAGAGGTAAGTGGAATTCCTAG	21.60
97	TACGTAGGTGGCAAGCGTTGTCCGGAATTACTGGGTGTAAAGGGAGCGTAGGCGGGGAGGCAAGTTGAATGTCTAAACTATCGGCTC AACTGATAGTCGCGTTCAAAACTGCCACTCTTGAGTGCAGTAGAGGTAGGCGGAATTCCTAGT	21.20
98	TACGTAGGGGGCGAGCGTTATCCGGATTTACTGGGTGTAAAGGGAGCGTAGACGGCGTATCAAGTCTGATGTGAAAGGCAGGGGCTT AACCCCTGGACTGCATTGGAAACTGGTATGCTTGAGTGCCGGAGGGGTAAGCGGAATTCCTAG	20.85
99	TACGTAGGGAGCAAGCGTTATCCGGATTTACTGGGTGTAAAGGGCGCGCAGGCGGGCCGGTAAGTTGGAAGTGAAATCTATGGGCTT AACCCATAAACTGCTTTCAAAACTGCTGGTCTTGAGTGATGGAGAGGCAGGCGGAATTCCGTG	20.78
100	TACGGAGGATCCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGCGGTTTGATAAGTTAGAGGTGAAATACCGGTGCTTA ACACCGGAACTGCCTCTAATACTGTTGAACTAGAGAGTAGTTGCGGTAGGCGGAATGTATGG	20.75