Bacterial ID by MALDI using DNA databases

SimulTOF Systems
Marlborough, MA

Ken Parker
Kenneth.parker@simultof.com

History

- Mass spec has been used for bacterial typing for decades
  - starting in the **1970**s, originally via pyrolysis of lipids
- **1996**: MALDI was used to type bacteria in protein mz range
  - ribosomal proteins correspond to many of the strongest masses
- **1997**: Pineda et al. -> matching organisms to ribosomal protein mass lists
- Since then, database look-up has largely been ignored
  - Except for selected clades (see next slide)
- Instead, FDA approved methods to identify bacteria via spectral matching
  - requires careful curation of library spectra
- In this talk, MALDI -> bacterial ID to DNA databases described
  - Want to extend MALDI market beyond pathogen ID
- Can easily search **10, 000** strains / species
  - no limit yet.

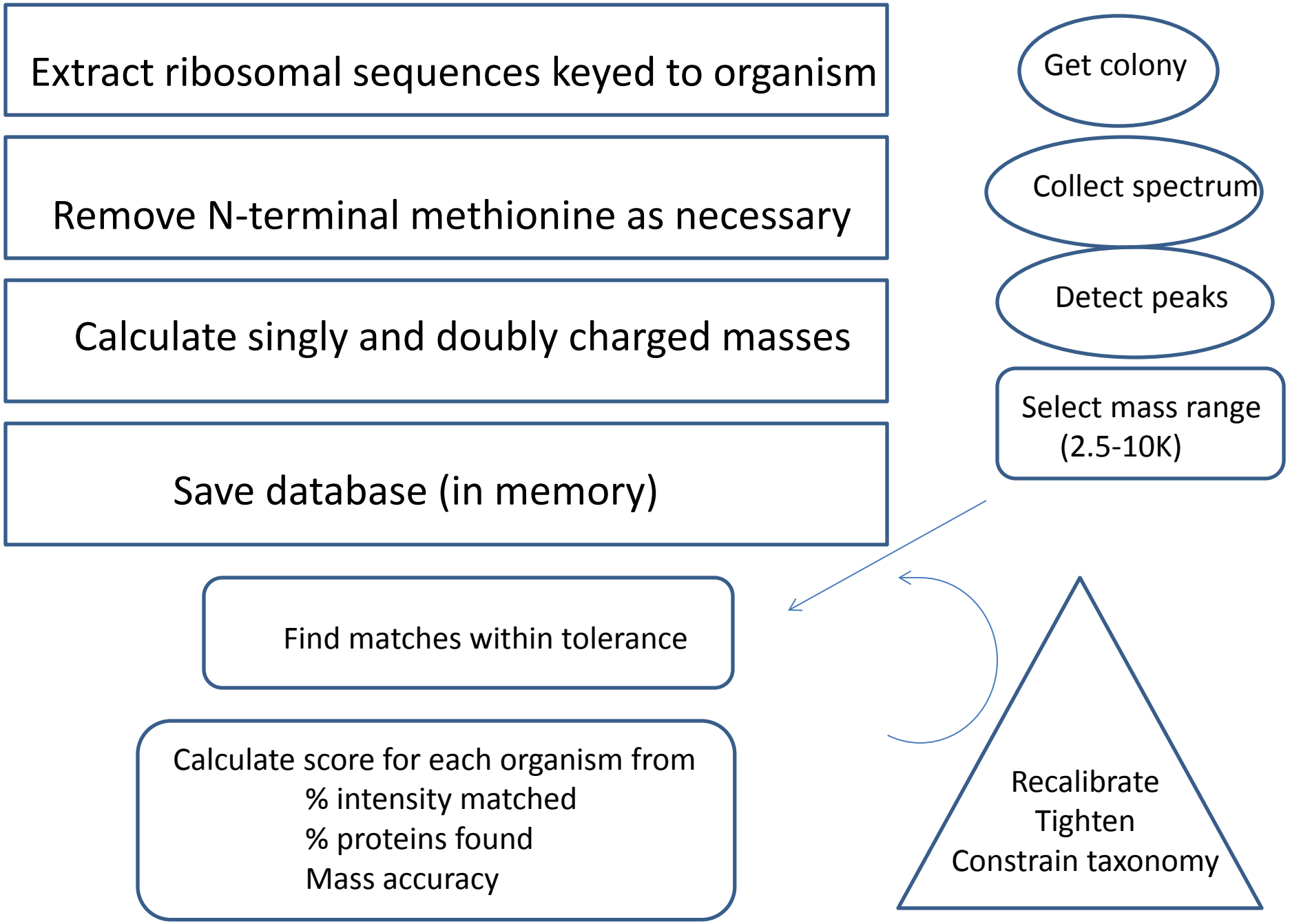
Discrimination of *Escherichia col*i O157, O26 and O111 from Other Serovars by MALDI-TOF MS Based on the S10-GERMS Method

Teruyo Ojima-Kato, Naomi Yamamoto, Mayumi Suzuki,
Tomohiro Fukunaga Hiroto Tamura
PLOS ONE | www.plosone.org 1 November 2014 | Volume 9 | Issue 11 | e113458

Identifies bacteria using the ribosomal proteins encoded together.
Called S10 GERM for S10-spc-alpha operon gene-encoded ribosomal protein mass spectrum

This uses about half of the ribosomal subunits

Our method uses all of the subunits in the desired mass range

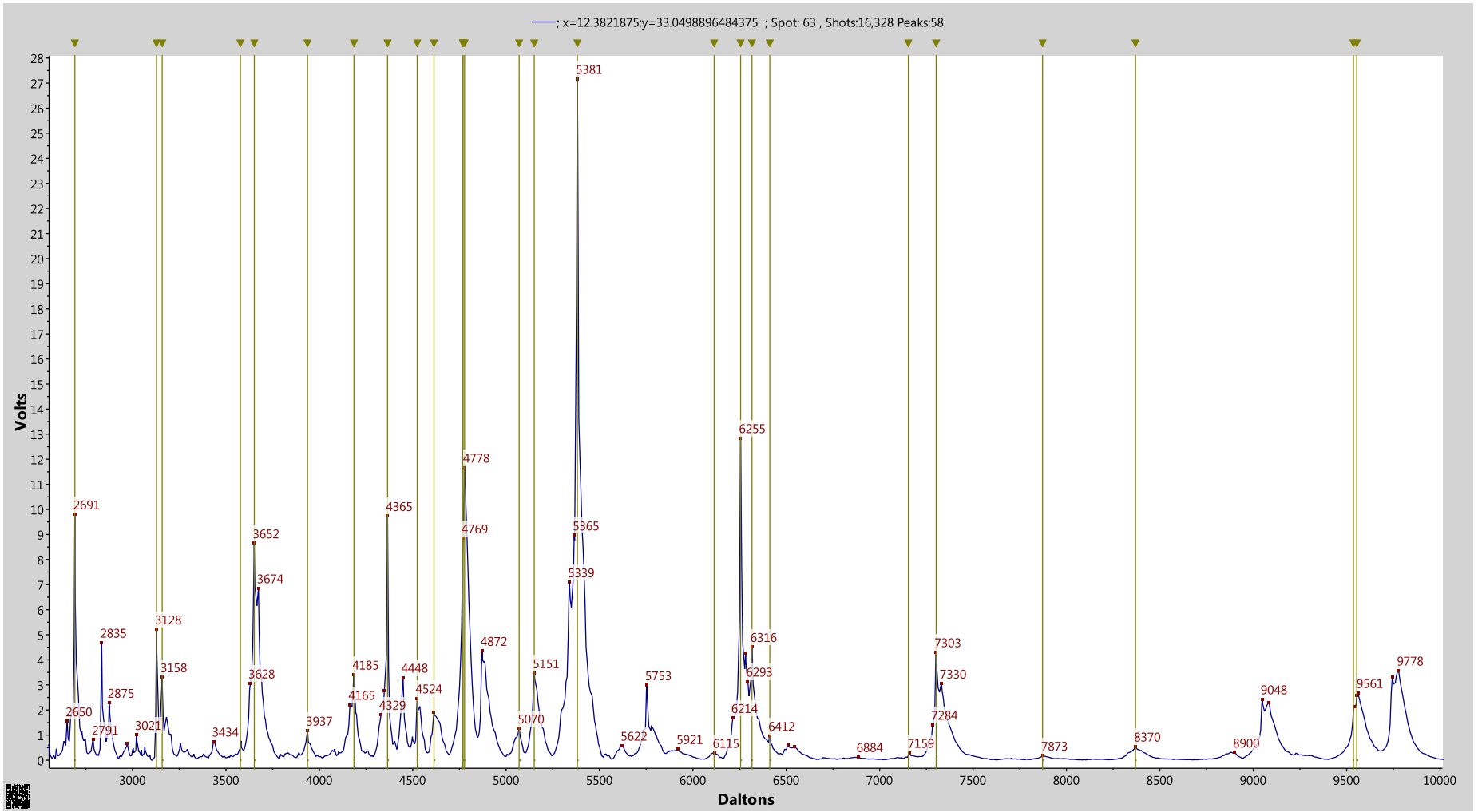

Extract ribosomal sequences keyed to organism
Remove N-terminal methionine as necessary
Calculate singly and doubly charged masses
Save database (in memory)
Get colony
Collect spectrum
Detect peaks
Select mass range
(2.5-10K)
Find matches within tolerance
Calculate score for each organism from
% intensity matched
% proteins found
Mass accuracy
Recalibrate
Tighten
Constrain taxonomy

When to search DNA databases?

- If you want to know why MALDI works (biochemistry)
- If you want to know what you are distinguishing
- When you are not happy with library searches
  - Returns scores of all organisms in the database
    - Useful for exploring relationships of closely related organisms
  - Returns sequences and names for all matches
- Can easily adjust mass tolerance or mass range to determine robustness of the answer
- Can restrict search to taxon id (saves time)
- Should work better for environmental organisms
- Come find me with any bacterial ID problem!

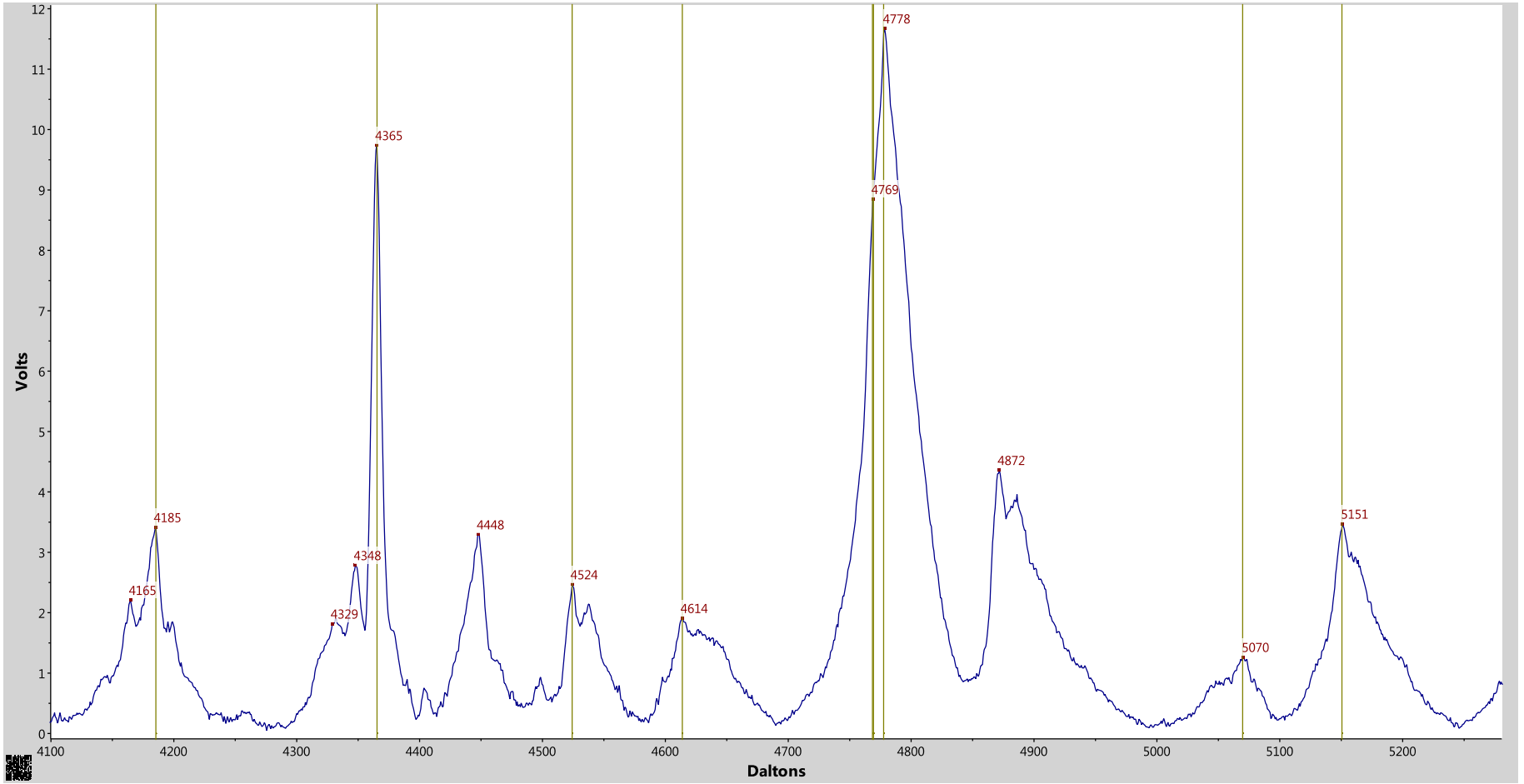
Matches of ribosomal proteins to *S. dysenteriae* within 500 ppm

Plate csm8 spot 63



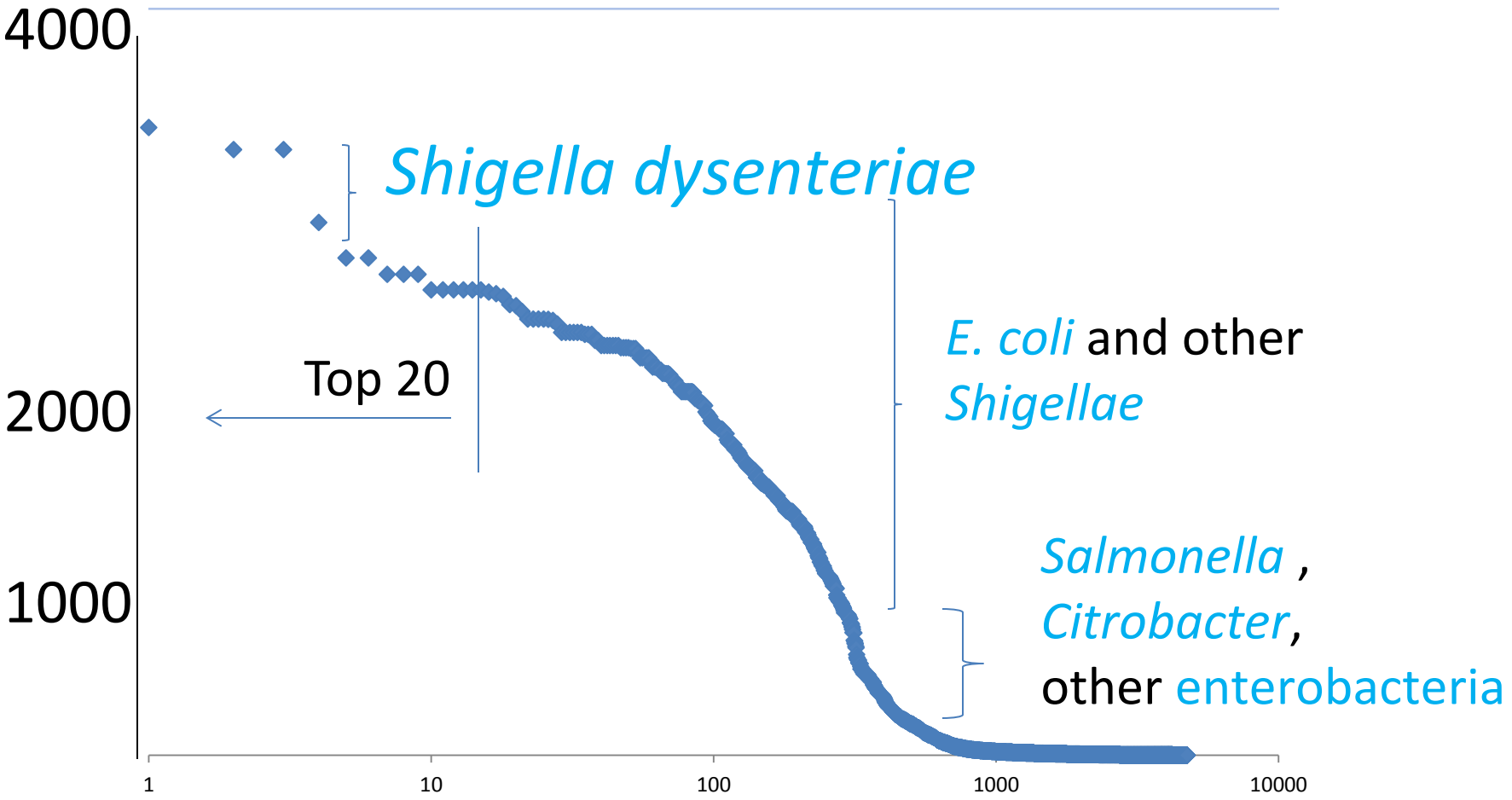
mz 2500-10000

Same spectrum
7 peaks matched in this region



mz 4100-5300

Score Vs. Organism



7641 strains

Top 20 scoring organisms; all related to *E. coli*

I	OrgID	Name	Proteins	Matches	Score	%rib	% int	ppm
1	1514	Shigella dysenteriae WRSd3	61	22	3365	36.1	64.7	80.8
2	70406	Shigella dysenteriae WRSd3	65	23	3245	35.4	64.8	80.7
3	70403	Shigella dysenteriae sd197	65	23	3245	35.4	64.8	80.7
4	6911	Escherichia coli KTE182	57	21	2856	36.8	52.6	88.7
5	8951	Escherichia coli 908573	59	21	2665	35.6	52.6	88.7
6	985	Escherichia coli KTE64	59	21	2665	35.6	52.6	88.7
7	4443	Escherichia coli MS 187-1	60	21	2577	35.0	52.6	88.7
8	1470	Escherichia coli KTE36	60	21	2577	35.0	52.6	88.7
9	1466	Escherichia coli KTE51	60	21	2577	35.0	52.6	88.7
10	2012	Escherichia coli UMEA 3240-1	61	21	2493	34.4	52.6	88.7
11	2088	Escherichia coli UMEA 3304-1	61	21	2493	34.4	52.6	88.7
12	3037	Escherichia coli 907357	61	21	2493	34.4	52.6	88.7
13	699	Escherichia coli KTE76	61	21	2493	34.4	52.6	88.7
14	1663	Escherichia coli KTE234	61	21	2493	34.4	52.6	88.7
15	1980	Escherichia coli KTE100	61	21	2493	34.4	52.6	88.7
16	70398	Shigella boydii CDC 3083-94	65	22	2482	33.8	54.2	86.3
17	15052	Escherichia coli TA280	46	17	2473	37.0	45.3	97.5
18	9460	Escherichia coli MS 79-10	58	20	2459	34.5	51.7	88.4
19	7783	Escherichia coli MS 182-1	58	20	2415	34.5	50.8	90.7
20	10135	Escherichia coli H299	59	20	2409	33.9	52.4	88.7

Matched peaks: 9 out of top 10 explained by ribosomal proteins

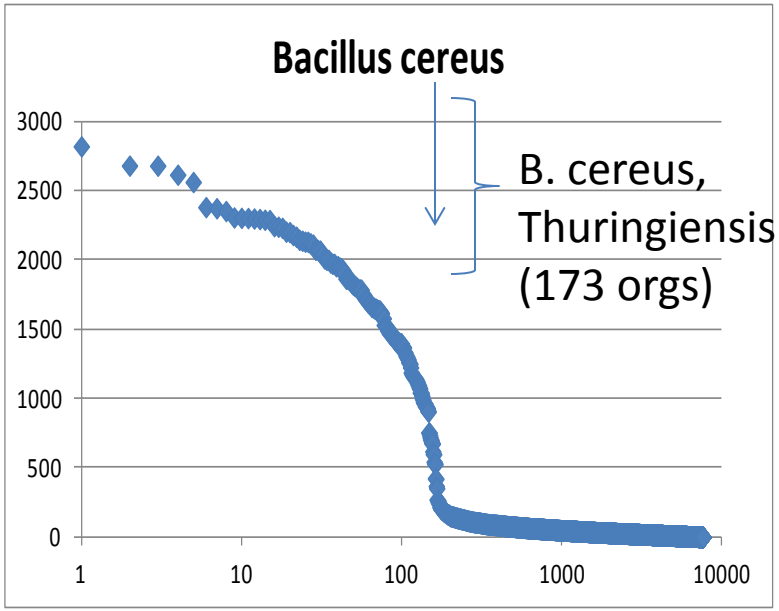
N	I	theo	exper	ppm	Acc	Protein	Seq
1	1	5381.4	5381.2	-45.5	A0A090NMR8	50S ribosomal protein L34	MKRTFQ
2	2	4777.6	4778.4	158.3	A0A090NGX7	30S ribosomal protein S20(2)	ANIKSAK
3	3	3651.8	3651.5	-71.6	A0A090NXZ7	50S ribosomal protein L29(2)	MKAKEL
4	4	6255.4	6255.0	-71.4	A0A090NL18	50S ribosomal protein L33Me	AKGIREK
5	5	2691.2	2690.8	-139.6	A0A090NMR8	50S ribosomal protein L34(2)	MKRTFQ
6	6	4365.4	4365.2	-31.0	A0A090NJL4	50S ribosomal protein L36	MKVRAS
7	8	7302.5	7302.5	0.4	A0A090NXZ7	50S ribosomal protein L29	MKAKEL
8	9	5150.6	5151.0	95.2	A0A090NJM9	30S ribosomal protein S19(2)	PRSLKKG
9	10	4185.4	4185.0	-82.9	A0A090NI81	30S ribosomal protein S21(2)	PVIKVRE
10	14	3128.2	3128.3	27.5	A0A090NL18	50S ribosomal protein L33Me(2)	AKGIREK
11	19	4613.8	4613.6	-33.5	A0A090NF12	DNA-binding protein HU beta(2)	MNKSQL
12	22	6316.2	6315.7	-74.0	A0A090NP57	50S ribosomal protein L32	AVQQN
13	23	3158.6	3157.9	-210.0	A0A090NP57	50S ribosomal protein L32(2)	AVQQN
14	24	3936.6	3937.0	119.2	A0A090NC24	50S ribosomal protein L31(2)	MKKDIH
15	25	5069.8	5070.3	108.9	A0A090NNF1	30S ribosomal protein S15(2)	SLSTEAT
16	29	8369.8	8370.0	29.2	A0A090NI81	30S ribosomal protein S21	PVIKVRE
17	40	4768.5	4768.9	79.0	A0A090NIK4	DNA-binding protein HU alpha(2)	MNKTQ
18	45	9536.0	9539.3	345.3	A0A090NIK4	DNA-binding protein HU alpha	MNKTQ
19	48	6114.2	6115.1	149.8	A0A090NJ47	50S ribosomal protein L22(2)	METIAK
20	49	6411.6	6412.2	87.5	A0A090NXZ9	50S ribosomal protein L30	AKTIKIT
21	50	7872.1	7873.3	147.9	A0A090NC24	50S ribosomal protein L31	MKKDIH
22	56	9554.2	9554.7	50.0	A0A090NGX7	30S ribosomal protein S20	ANIKSAK

Masses for the top 20 organism hits

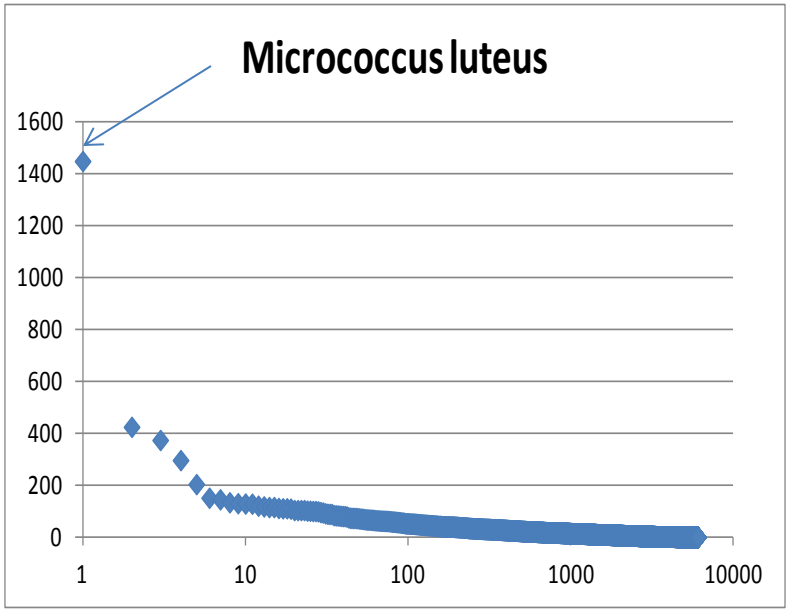
Green: same mass as top hit
Red: subunit missing
Orange: subunit has extension
Blue: subunit has substitution (V->A)

N	I	Protein	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20
1	1	50S ribosomal protein L34	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381	5381
2	2	30S ribosomal protein S20(2)	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778	4778
3	3	50S ribosomal protein L29(2)	3652	3652	3652	3638		3638	3638	3638		3638		3638	3638	3638	3638	3638	3638	3638	3638	3638
4	4	50S ribosomal protein L33Me	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255	6255
5	5	50S ribosomal protein L34(2)	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691	2691
6	6	50S ribosomal protein L36	4365	4365	4365	4365	4365	4365	4365	4365	4365	4365	4365	4365	4365	4365	4365	4365	5415	4365	4365	4365
7	8	50S ribosomal protein L29	7303	7303	7303	7274		7274	7274	7274		7274		7274	7274	7274	7274	7274	7274	7274	7274	7274
8	9	30S ribosomal protein S19(2)	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151	5151
9	10	30S ribosomal protein S21(2)	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185	4185
10	14	50S ribosomal protein L33Me(2)	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128	3128
11	19	DNA-binding protein HU beta(2)	4614	4614	4614	4614	4614	4614	4614	4614	4614	4614	4614	4614	4614	4614	4614	4614		4614		4614
12	22	50S ribosomal protein L32	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316	6316
13	23	50S ribosomal protein L32(2)	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158	3158
14	24	50S ribosomal protein L31(2)	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937	3937
15	25	30S ribosomal protein S15(2)	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070	5070		5070	5070
16	29	30S ribosomal protein S21	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370	8370
17	40	DNA-binding protein HU alpha(2)	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769	4769
18	45	DNA-binding protein HU alpha	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539	9539
19	48	50S ribosomal protein L22(2)	6115	6115	6115	6115	6115	6115	6115	6115	6115	6115	6115	6115	6115	6115	6115	6115		6115	6115	6115
20	49	50S ribosomal protein L30	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	6412	
21	50	50S ribosomal protein L31	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873	7873
22	56	30S ribosomal protein S20	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555	9555
23	20	50S ribosomal protein L25(2)	7162	5348	5348			5348	5348	5347	5348	5348	5347	5348	5347	5348	5348	5339	5347	5348	5348	5347
24	53	50S ribosomal protein L35	7891	7159	7159	7159	7159	7159	7159	7159	7159	7159	7159	7159	7159	7159	7159	7159		7159	7159	7159

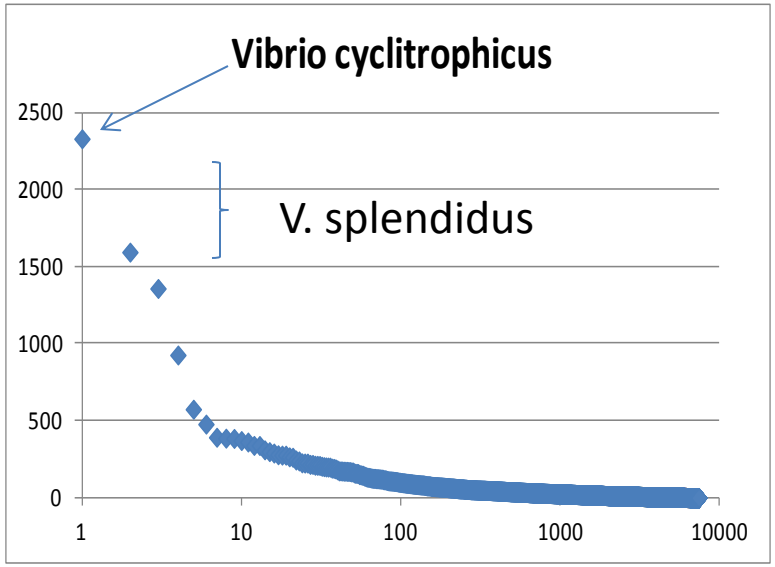
Overgrown LB medium



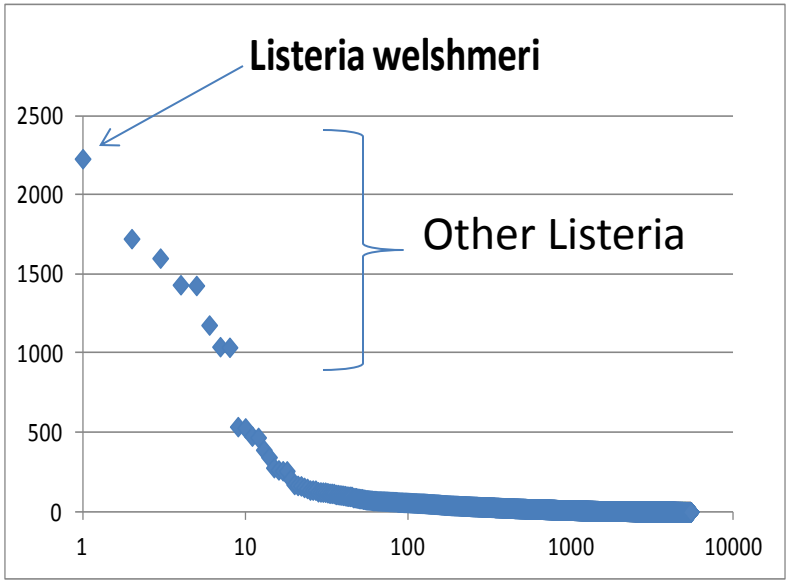
Colorado School of Mines



Martin Polz' lab MIT



Colorado School of Mines



What is a bacterial species?

- Originally:
  - disease
  - growth requirements / metabolism
  - ecological niche
- New consensus: genomic similarity -> proteomic similarity
  - Vertical vs Horizontal evolution
- Simplification: Substantial Ribosomal protein sequence identity
  - Highly conserved throughout prokaryotic evolution, transmitted vertically
  - 54 subunits typically
  - Complex nearly 1:1
- MALDI recognizes small, abundant, basic proteins like ribosomal subunits
- Therefore, MALDI can distinguish species
  - If organisms cannot be distinguished by MALDI, then they do not belong in different species
  - Limitations of species delineation by MALDI due to errors in species definition

Distinguishing *Shigella* from *E. coli*

- Presumably this is clinically important
- There are 4 defined *Shigella* species
  - *Boydii* **(35)**, *dysenteriae* **(14)**, *flexnerii* **(120)**, *sonnei* **(583)** vs *E. coli* **(3947)**
  - Known for 20 years *Shigella* not separable from *E. coli.*
  - *Shigella* 'species' not separable from one another.
  - Epidemiological differences between *Shigella* species
    - According to my primary source, Wikipedia.
  - *Shigella*'s pathogenicity due to a large plasmid that promotes invasion
    - A family of ~ 8 E3 ubiquitin ligases
    - Type III secretion system
- Carefully check ribosomal protein profile for 4 examples of *E. coli* and each *Shigella* species

PATRIC cladogram

https://

www.patricbrc.org/

Pathosystyems Resource Integration Center

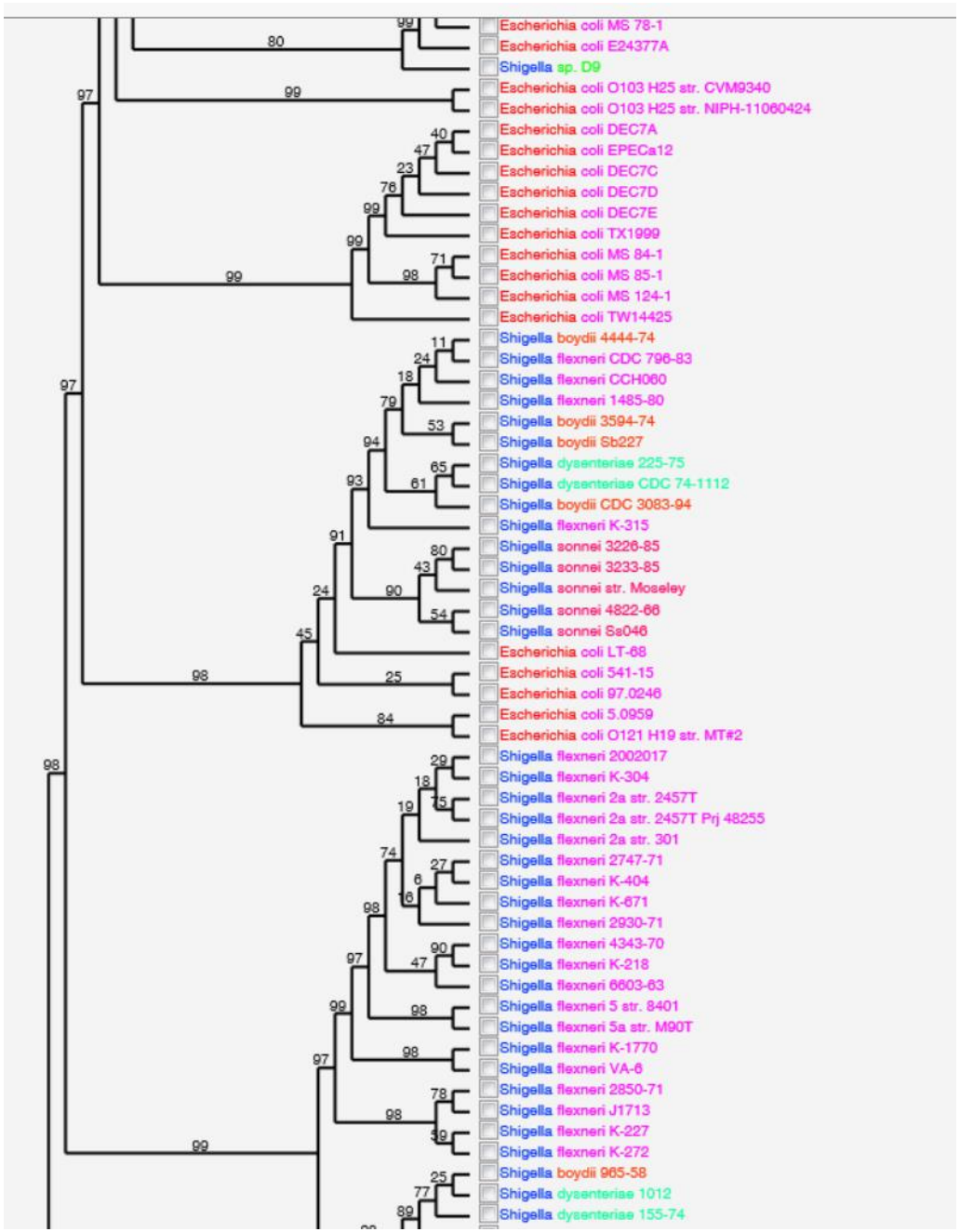
NIH/NIAID-funded project of The University of Chicago with subcontract to the Biocomplexity Institute of Virginia Tech

Escherichia **red**

Shigella blue

dysenteriae **green**

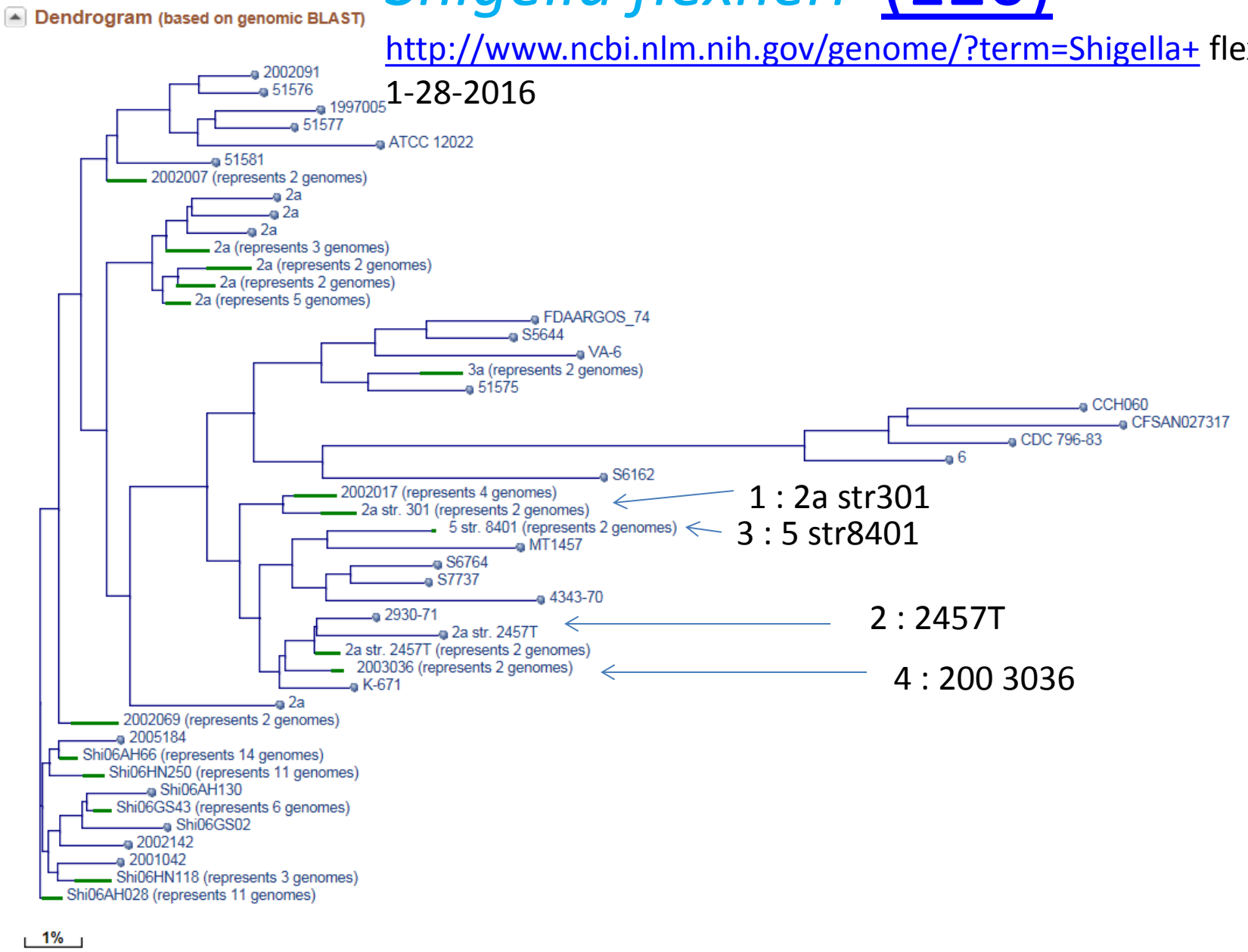
Lots of ***E. coli*** strains above and below, mixed with a few ***Shigellae***



Shigella flexneri (120)

http://www.ncbi.nlm.nih.gov/genome/?term=Shigella+ flexneri

1-28-2016



Clustering of selected Enterobacteria by ribosomal proteins + DNA-binding protein HU

I	Organism	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
1	Serratia marcescens Db	50	41	42	3	3	3	3	3	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	Serratia marcescens 66	41	49	42	3	3	3	3	3	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	Serratia marcescens 78	42	42	48	3	3	3	3	3	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	Citrobacter freundii CF	3	3	3	57	41	39	38	38	6	6	18	2	3	4	3	4	6	6	5	5	5	5	5	7	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	Citrobacter freundii CA	3	3	3	41	56	46	45	45	9	9	16	3	4	5	4	5	6	6	5	5	5	4	4	6	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	Citrobacter KTE151 2-9	3	3	3	39	46	56	54	54	8	8	14	4	4	6	5	6	7	7	6	6	6	5	5	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	Citrobacter freundii 98	3	3	3	38	45	54	56	56	8	8	14	4	4	6	5	6	7	7	6	6	6	5	5	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
8	Citrobacter koseri 989_	3	3	3	38	45	54	56	56	8	8	14	4	4	6	5	6	7	7	6	6	6	5	5	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
9	Klebsiella pneumoniae 1243_	2	2	2	6	9	8	8	8	58	57	10	7	6	9	8	8	8	8	7	7	7	8	8	11	10	10	10	10	10	10	10	10	10	8	9	9	9	10	8
10	Klebsiella pneumoniae KPNIH	2	2	2	6	9	8	8	8	57	59	10	7	6	9	8	8	8	8	7	7	7	8	8	11	10	10	10	10	10	10	10	10	10	8	9	9	9	10	8
11	Salmonella enterica serovar T	3	3	3	18	16	14	14	14	10	10	56	12	12	14	13	14	16	15	14	13	14	16	16	18	17	17	16	17	16	16	16	17	17	14	15	15	15	16	15
12	Shigella flexneri 200 30	1	1	1	2	3	4	4	4	7	7	12	49	38	38	40	42	29	27	26	29	30	30	29	30	34	34	34	34	34	34	34	36	36	33	32	31	32	34	30
13	Shigella flexneri 2a str3	1	1	1	3	4	4	4	4	6	6	12	38	56	44	46	48	33	31	30	34	34	36	35	36	39	39	39	39	39	39	39	41	41	38	37	36	37	39	36
14	Shigella flexneri 5 str 8	1	1	1	4	5	6	6	6	9	9	14	38	44	56	48	51	38	36	37	40	41	44	43	44	48	48	48	48	49	49	48	50	50	46	45	45	45	48	43
15	Shigella flexneri 20020	1	1	1	3	4	5	5	5	8	8	13	40	46	48	54	52	37	35	34	36	37	39	38	40	44	44	44	45	43	43	43	45	45	42	40	40	41	44	39
16	Shigella flexneri 2a str2	1	1	1	4	5	6	6	6	8	8	14	42	48	51	52	56	40	38	37	40	41	43	42	43	47	47	47	47	47	47	47	49	49	45	44	43	44	47	42
17	Escherichia alberti ECO6-170	1	1	1	6	6	7	7	7	8	8	16	29	33	38	37	40	56	52	33	35	36	39	38	40	42	42	41	42	42	42	42	42	42	39	39	38	39	41	37
18	Escherichia alberti KF1	1	1	1	6	6	7	7	7	8	8	15	27	31	36	35	38	52	56	30	32	33	37	36	37	41	41	39	40	40	40	39	40	40	37	36	36	37	39	35
19	Shigella dysenteriae dy	1	1	1	5	5	6	6	6	7	7	14	26	30	37	34	37	33	30	56	49	50	36	36	38	40	40	40	40	40	40	39	40	40	38	37	37	37	40	36
20	Shigella dysenteriae WI	1	1	1	5	5	6	6	6	7	7	13	29	34	40	36	40	35	32	49	56	55	39	38	40	42	42	43	42	43	43	42	43	43	41	41	40	40	42	37
21	Shigella dysenteriae sd	1	1	1	5	5	6	6	6	7	7	14	30	34	41	37	41	36	33	50	55	56	40	39	41	43	43	44	43	44	44	43	44	44	41	41	40	40	43	38
22	Shigella sonnei 536 pla	1	1	1	5	4	5	5	5	8	8	16	30	36	44	39	43	39	37	36	39	40	56	54	45	47	47	47	47	47	47	48	50	50	46	47	46	47	49	44
23	Shigella sonnei Ss046 p	1	1	1	5	4	5	5	5	8	8	16	29	35	43	38	42	38	36	36	38	39	54	56	44	46	46	46	46	46	46	47	49	49	45	46	45	46	48	43
24	Escherichia fergusonii ATCC3	2	2	2	7	6	7	7	7	11	11	18	30	36	44	40	43	40	37	38	40	41	45	44	56	49	49	48	49	48	48	48	49	49	45	46	45	46	48	43
25	Escherichia coli O157 H	1	1	1	5	5	6	6	6	10	10	17	34	39	48	44	47	42	41	40	42	43	47	46	49	56	56	52	53	52	52	51	53	53	49	48	48	49	52	47
26	Escherichia coli O157:H	1	1	1	5	5	6	6	6	10	10	17	34	39	48	44	47	42	41	40	42	43	47	46	49	56	57	52	53	52	52	51	53	53	49	48	48	49	52	47
27	Escherichia coli BL21 D	1	1	1	5	5	6	6	6	10	10	16	34	39	48	44	47	41	39	40	43	44	47	46	48	52	52	56	53	52	52	51	53	53	48	48	47	48	51	46
28	Shigella boydii 775-SBC	1	1	1	5	5	6	6	6	10	10	17	34	39	48	45	47	42	40	40	42	43	47	46	49	53	53	53	56	52	52	51	53	53	49	48	48	49	52	47
29	Escherichia coli K12 MC	1	1	1	5	5	6	6	6	10	10	16	34	39	49	43	47	42	40	40	43	44	47	46	48	52	52	52	52	56	56	53	53	53	48	48	48	48	51	46
30	Escherichia coli K12 W3	1	1	1	5	5	6	6	6	10	10	16	34	39	49	43	47	42	40	40	43	44	47	46	48	52	52	52	52	56	56	53	53	53	48	48	48	48	51	46
31	Shigella dysenteriae SD	1	1	1	5	5	6	6	6	10	10	16	34	39	48	43	47	42	39	39	42	43	48	47	48	51	51	51	51	53	53	56	54	54	49	50	48	49	52	47
32	Escherichia coli O104:H	1	1	1	5	5	6	6	6	10	10	17	36	41	50	45	49	42	40	40	43	44	50	49	49	53	53	53	53	53	53	53	56	56	51	51	50	51	54	49
33	Escherichia coli O111:H	1	1	1	5	5	6	6	6	10	10	17	36	41	50	45	49	42	40	40	43	44	50	49	49	53	53	53	53	53	53	53	56	56	51	51	50	51	54	49
34	Shigella boydii Sb227 p	1	1	1	5	5	6	6	6	8	8	14	33	38	46	42	45	39	37	37	41	41	46	45	45	49	49	48	49	48	48	49	51	51	56	52	53	54	52	47
35	Shigella flexneri CDC 79	1	1	1	5	5	6	6	6	9	9	15	32	37	45	40	44	39	36	37	41	41	47	46	46	48	48	48	48	48	48	50	51	51	52	56	53	54	51	46
36	Shigella boydii 4444-74	1	1	1	5	5	6	6	6	9	9	15	31	36	45	40	43	38	36	37	40	40	46	45	45	48	48	47	48	48	48	48	50	50	53	53	56	55	51	46
37	Shigella boydii ATCC 92	1	1	1	5	5	6	6	6	9	9	15	32	37	45	41	44	39	37	37	40	40	47	46	46	49	49	48	49	48	48	49	51	51	54	54	55	56	52	47
38	Shigella boydii CDC 308	1	1	1	5	5	6	6	6	10	10	16	34	39	48	44	47	41	39	40	42	43	49	48	48	52	52	51	52	51	51	52	54	54	52	51	51	52	56	50
39	Shigella dysenteriae 22	1	1	1	5	5	6	6	6	8	8	15	30	36	43	39	42	37	35	36	37	38	44	43	43	47	47	46	47	46	46	47	49	49	47	46	46	47	50	54

Clustering of selected Enterobacteria by sequence identity of all proteins (nearly the same pattern)

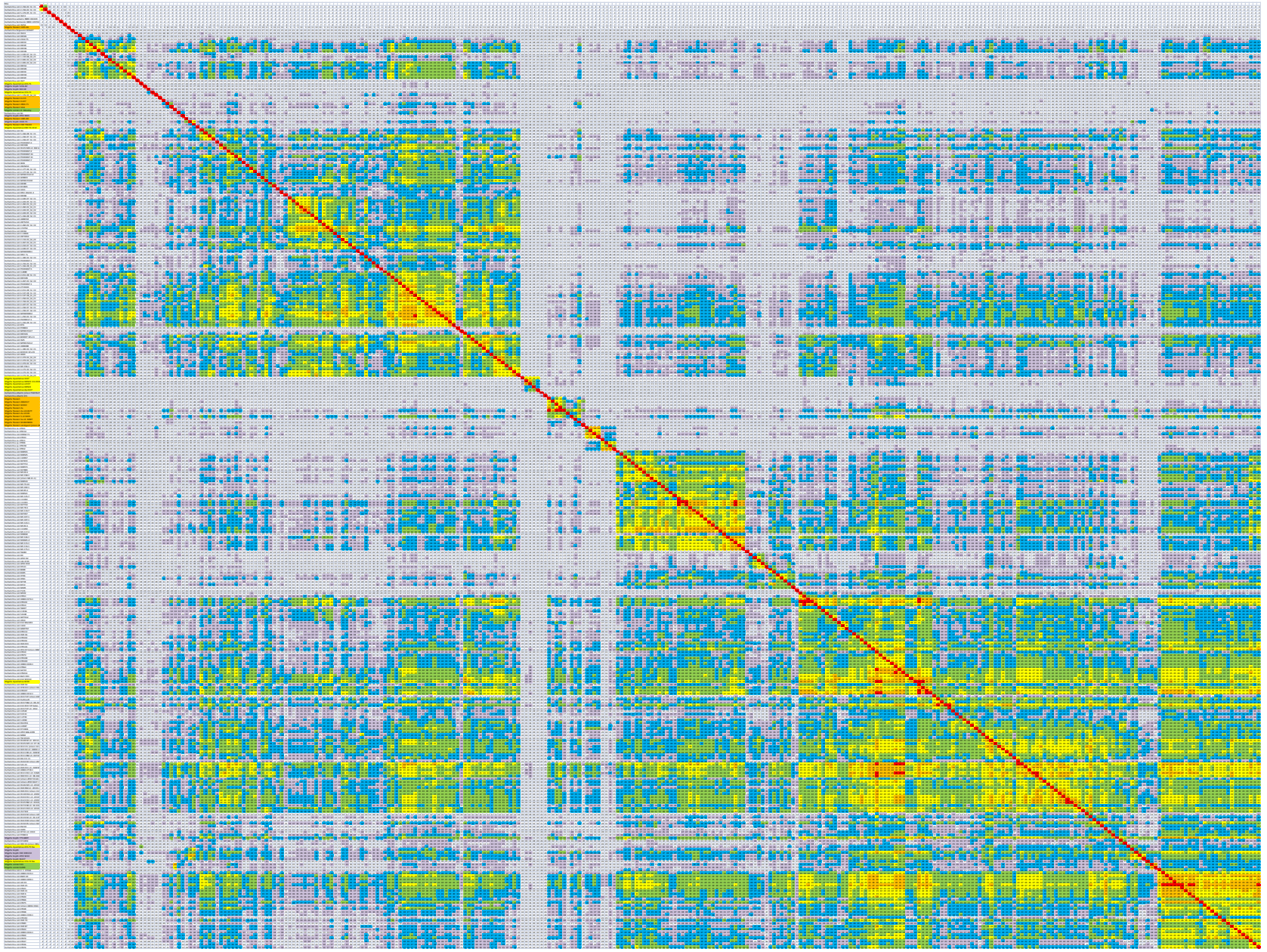
I	Organism	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
1	Serratia marcescens Db1	4600	549	512	4	4	4	5	5	5	5	5	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	Serratia marcescens 666_	549	4916	580	5	13	23	19	8	8	10	10	2	2	2	2	2	3	3	2	3	3	3	2	2	2	3	2	2	3	3	4	2	2	5	4	4	4	7	3
3	Serratia marcescens 786_	512	580	5067	5	21	10	16	9	32	12	12	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	5	2	2	2	6	3
4	Salmonella enterica serovar Typ	4	5	5	4418	23	35	53	43	33	33	34	38	30	25	29	29	42	34	35	31	38	32	33	34	38	37	34	33	35	34	47	36	37	35	33	35	35	51	174
5	Klebsiella pneumoniae 1243_CK	4	13	21	23	5054	2244	26	30	30	26	26	16	16	15	17	17	26	19	16	14	17	17	20	20	19	20	19	17	20	18	21	18	18	33	19	20	20	26	21
6	Klebsiella pneumoniae KPNIH10	4	23	10	35	2244	5643	41	40	31	34	34	19	16	15	17	17	22	19	21	15	23	17	20	20	21	26	21	19	20	20	31	18	18	25	20	20	20	37	32
7	Citrobacter freundii CFNI	5	19	16	53	26	41	4929	331	295	297	296	19	17	14	16	15	18	15	20	13	23	15	15	16	16	20	16	17	17	19	26	17	17	27	21	16	16	34	23
8	Citrobacter freundii CAV1	5	8	9	43	30	40	331	5038	425	426	423	17	18	14	20	16	23	18	17	12	17	20	16	17	15	18	15	15	16	16	23	16	16	40	16	18	18	42	18
9	Citrobacter KTE151 2-9-2	5	8	32	33	30	31	295	425	4766	1895	1875	17	17	14	17	16	19	17	15	11	14	16	16	17	15	15	15	15	16	16	18	17	17	21	15	21	21	21	16
10	Citrobacter freundii 980-	5	10	12	33	26	34	297	426	1895	4577	4477	17	18	14	17	16	19	19	16	13	16	18	16	17	16	16	15	15	17	16	19	18	18	23	16	24	24	25	19
11	Citrobacter koseri 989_C	5	10	12	34	26	34	296	423	1875	4477	4566	17	18	14	17	16	19	19	16	13	16	18	16	17	16	16	15	15	17	16	19	18	18	23	16	24	24	25	19
12	Escherichia alberti ECO6-170	2	2	2	38	16	19	19	17	17	17	17	4289	1561	112	158	154	170	167	147	143	170	174	205	216	190	185	191	190	180	176	224	191	185	209	200	203	210	213	215
13	Escherichia alberti KF1	2	2	2	30	16	16	17	18	17	18	18	1561	4185	102	146	144	154	160	140	138	166	168	199	224	176	174	179	180	177	166	214	191	184	198	202	199	205	206	219
14	Shigella dysenteriae dys1	2	2	2	25	15	15	14	14	14	14	14	113	103	5520	1807	1917	163	278	263	294	288	280	416	412	300	300	301	307	316	300	373	366	357	385	378	394	401	405	356
15	Shigella dysenteriae WRS	2	2	2	29	17	17	16	20	17	17	17	158	146	1800	3589	2451	255	455	383	383	489	491	605	711	522	546	511	516	538	490	640	631	616	657	629	692	625	611
16	Shigella dysenteriae sd19	2	2	2	29	17	17	15	16	16	16	16	154	144	1911	2451	3531	229	401	363	409	422	409	629	603	449	438	444	449	465	426	559	549	535	571	560	603	602	598	536
17	Escherichia fergusonii ATCC354	3	3	3	42	26	22	18	23	19	19	19	170	154	163	255	229	4269	299	265	244	304	309	368	419	332	341	341	328	328	317	470	378	373	442	419	405	431	428	412
18	Shigella flexneri 5 str 840	2	3	2	34	19	19	15	18	17	19	19	167	160	275	455	401	299	4028	1891	1813	2347	2400	640	736	715	730	731	777	706	666	737	869	842	853	851	810	891	856	872
19	Shigella flexneri 200 303	2	2	2	35	16	21	20	17	15	16	16	147	140	261	383	363	265	1891	4565	2441	3053	2924	603	648	586	588	595	647	576	557	631	728	702	737	732	752	771	783	752
20	Shigella flexneri 2a str301	2	3	2	31	14	15	13	12	11	13	13	143	138	291	383	409	244	1813	2441	3957	2762	2703	594	589	572	564	565	635	569	533	587	709	685	694	682	714	720	739	686
21	Shigella flexneri 2002017	2	3	2	38	17	23	23	17	14	16	16	170	166	286	489	422	304	2347	3053	2762	4343	3654	645	743	750	726	718	806	743	697	735	913	902	862	859	818	900	888	877
22	Shigella flexneri 2a str245	2	3	2	32	17	17	15	20	16	18	18	174	168	278	491	409	309	2400	2924	2703	3654	4079	657	758	726	736	744	805	716	688	747	889	861	884	888	842	929	899	893
23	Escherichia coli O157 H7	2	2	2	33	20	20	15	16	16	16	16	205	199	413	605	629	368	640	603	594	645	657	5162	3922	735	760	767	739	728	715	1004	901	862	1034	1057	1116	1107	1151	1209
24	Escherichia coli O157:H7	2	2	2	34	20	20	16	17	17	17	17	216	224	409	711	603	419	736	648	589	743	758	3922	5413	831	859	867	835	835	820	1135	1039	990	1179	1192	1162	1260	1209	1388
25	Shigella flexneri CDC 796-	2	2	2	38	19	21	16	15	15	16	16	190	176	299	522	449	332	715	586	572	750	726	735	831	4143	3028	3126	2954	1818	1677	814	1112	1078	988	978	923	1022	1064	1101
26	Shigella boydii 4444-74	2	3	2	37	20	26	20	18	15	16	16	185	174	299	546	438	341	730	588	564	726	736	760	859	3028	4322	3509	3256	1765	1706	830	1112	1077	1025	1016	956	1062	1099	1125
27	Shigella boydii ATCC 9210	2	2	2	34	19	21	16	15	15	15	15	191	179	300	511	444	341	731	595	565	718	744	767	867	3126	3509	3979	3405	1794	1714	831	1117	1083	1028	1021	956	1057	1104	1134
28	Shigella boydii Sb227 plas	2	2	2	33	17	19	17	15	15	15	15	190	180	306	516	449	328	777	647	635	806	805	739	835	2954	3256	3405	4038	1719	1634	803	1091	1059	992	988	925	1020	1059	1081
29	Shigella boydii CDC 3083-	2	3	2	35	20	20	17	16	16	17	17	180	177	315	538	465	328	706	576	569	743	716	728	835	1818	1765	1794	1719	4161	1727	819	1114	1076	978	982	920	1015	1064	1086
30	Shigella dysenteriae 225-	2	3	2	34	18	20	19	16	16	16	16	176	166	299	490	426	317	666	557	533	697	688	715	820	1677	1706	1714	1634	1727	4065	782	1036	989	949	950	885	982	1008	1033
31	Shigella boydii 775-SBOY	2	4	3	47	21	31	26	23	18	19	19	224	214	372	640	559	470	737	631	587	735	747	1004	1135	814	830	831	803	819	782	5000	1008	965	1164	1243	1203	1303	1194	1178
32	Shigella sonnei 536 plasm	2	2	2	36	18	18	17	16	17	18	18	191	191	363	631	549	378	869	728	709	913	889	901	1039	1112	1112	1117	1091	1114	1036	1008	4576	3656	1228	1241	1187	1299	1366	1367
33	Shigella sonnei Ss046 pla	2	2	2	37	18	18	17	16	17	18	18	185	184	355	616	535	373	842	702	685	902	861	862	990	1078	1077	1083	1059	1076	989	965	3656	4354	1186	1190	1151	1257	1317	1309
34	Shigella dysenteriae SD1	2	5	5	35	33	25	27	40	21	23	23	209	198	381	657	571	442	853	737	694	862	884	1034	1179	988	1025	1028	992	978	949	1164	1228	1186	4321	1865	1719	1874	1497	1421
35	Escherichia coli BL21 DE3	2	4	2	33	19	20	21	16	15	16	16	200	202	375	647	560	419	851	732	682	859	888	1057	1192	978	1016	1021	988	982	950	1243	1241	1190	1865	4244	2055	2276	1512	1494
36	Escherichia coli K12 MG1	2	4	2	35	20	20	16	18	21	24	24	203	199	392	629	603	405	810	752	714	818	842	1116	1162	923	956	956	925	920	885	1203	1187	1151	1719	2055	4066	3858	1539	1433
37	Escherichia coli K12 W31	2	4	2	35	20	20	16	18	21	24	24	210	205	399	692	602	431	891	771	720	900	929	1107	1260	1022	1062	1057	1020	1015	982	1303	1299	1257	1874	2276	3858	4347	1570	1554
38	Escherichia coli O104:H4	2	7	6	51	26	37	34	42	21	25	25	213	206	403	625	598	428	856	783	739	888	899	1151	1209	1064	1099	1104	1059	1064	1008	1194	1366	1317	1497	1512	1539	1570	5083	2022
39	Escherichia coli O111:H s	2	3	3	174	21	32	23	18	16	19	19	215	219	354	611	536	412	872	752	686	877	893	1209	1388	1101	1125	1134	1081	1086	1033	1178	1367	1309	1421	1494	1433	1554	2022	5483

Clustering of selected Enterobacteria by sequence identity of C-terminal hexapeptide

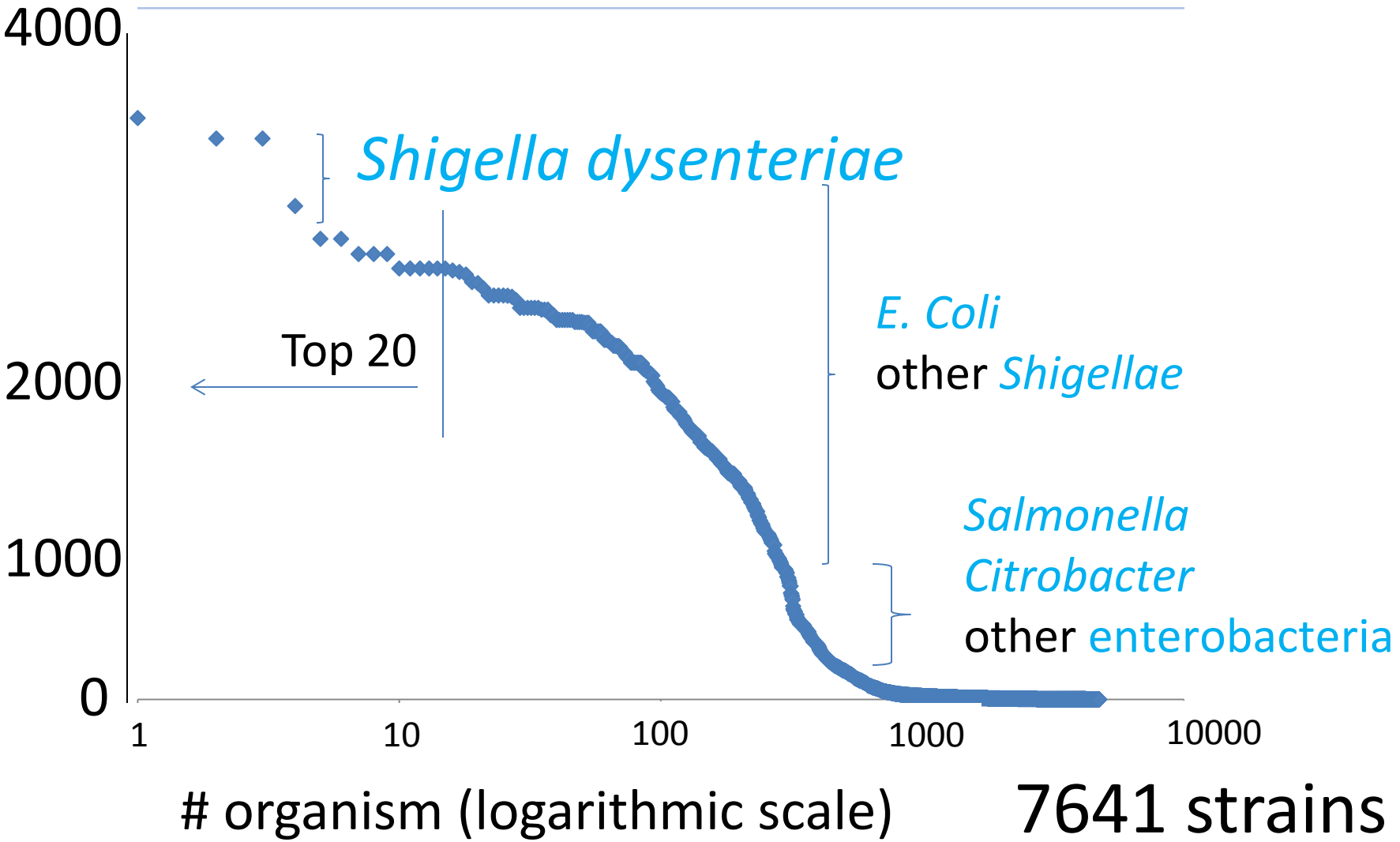
Serratia marcescens 666_SMA	4928	3079	3019	278	350	285	291	289	293	322	151	229	265	234	271	268	270	268	281	259	263	261	252	248	229	236	254	254	254
Serratia marcescens 786_SSo	3077	5083	3104	301	305	349	302	300	300	302	152	256	260	236	277	273	274	273	281	266	268	263	259	254	233	236	259	256	258
Serratia marcescens Db11	3021	3106	4617	266	278	282	282	280	280	281	151	228	268	236	272	271	273	271	264	259	262	253	258	253	232	240	256	255	257
Klebsiella pneumoniae 1243_(	285	307	272	5065	4126	836	832	826	852	842	402	606	690	645	744	721	734	728	763	714	713	674	677	666	610	588	667	662	667
Klebsiella pneumoniae KPNIH1	352	308	279	4126	5762	923	916	909	941	945	423	655	726	672	799	763	780	781	790	784	797	792	758	749	716	668	750	745	748
Citrobacter KTE151 2-9-2016	293	357	289	831	916	4794	3836	3780	2784	2608	608	886	1086	948	1182	1144	1170	1130	1146	1123	1125	1066	1048	1031	986	927	1068	1048	1064
Citrobacter freundii 980-SSon	296	306	287	825	908	3837	4585	4484	2746	2616	622	905	1095	956	1182	1136	1161	1121	1147	1102	1092	1027	1027	1012	941	913	1040	1029	1041
Citrobacter koseri 989_CKOS	294	304	285	819	901	3781	4484	4574	2710	2581	616	896	1083	944	1167	1124	1148	1109	1135	1094	1082	1014	1016	1001	933	906	1029	1019	1031
Citrobacter freundii CAV1321	296	305	285	849	933	2783	2746	2709	5108	2753	612	888	1105	950	1198	1147	1173	1123	1171	1109	1107	1036	1035	1017	960	920	1060	1050	1064
Citrobacter freundii CFNIH1 2-	326	306	286	835	934	2610	2617	2582	2754	4991	608	877	1091	929	1176	1125	1151	1139	1125	1123	1135	1086	1059	1040	1017	931	1074	1061	1066
Shigella dysenteriae dys1617	152	154	154	400	418	609	623	617	614	612	5978	2664	1553	1843	1450	1465	1470	1450	1451	1497	1481	1488	1401	1413	1299	1510	1382	1416	1380
Shigella dysenteriae sd197 pla	232	259	231	601	648	889	909	900	890	880	2664	7473	2355	2648	2144	2237	2198	2138	2180	2494	2362	2534	2229	2347	1944	2607	2384	2505	2345
Escherichia coli O157 H7 Saka	266	261	269	684	719	1088	1097	1086	1105	1095	1555	2357	5545	2228	3038	3012	2985	2926	2927	2805	2739	2598	2537	2521	2323	2477	2614	2600	2621
Shigella dysenteriae WRSd3	238	239	239	643	666	952	961	949	955	934	1843	2650	2229	3597	2435	2275	2446	2407	2420	2471	2461	2344	2271	2287	1909	1917	2233	2333	2275
Shigella boydii 775-SBOY	274	280	275	739	792	1185	1186	1172	1201	1180	1454	2150	3038	2438	5032	3188	3427	3348	3303	3096	3023	2758	2810	2770	2416	2331	2839	2824	2841
Escherichia coli K12 MG1655	270	274	272	715	756	1148	1141	1130	1152	1130	1465	2241	3015	2276	3187	4106	3897	3425	3247	3006	2961	2699	2739	2697	2505	2463	2800	2767	2803
Escherichia coli K12 W3110 Ci	272	275	274	729	774	1173	1165	1153	1178	1155	1469	2202	2986	2445	3426	3896	4387	3740	3530	3245	3197	2913	2956	2904	2561	2452	2999	2971	3017
Escherichia coli BL21 DE3	269	273	271	721	773	1130	1123	1112	1125	1142	1452	2145	2926	2408	3347	3422	3739	4294	3486	3249	3232	2959	2989	2934	2558	2422	2973	2947	2986
Shigella dysenteriae SD1D	284	283	266	758	784	1148	1151	1140	1177	1130	1449	2184	2926	2419	3299	3247	3531	3483	4351	3113	3089	2831	2891	2845	2428	2359	2870	2874	2903
Shigella sonnei 536 plasmid	263	268	261	707	778	1125	1105	1097	1112	1128	1501	2500	2804	2472	3096	3006	3248	3250	3116	4962	4516	3464	3224	3257	2709	2769	3222	3321	3181
Shigella sonnei Ss046 plasmid	265	268	262	705	788	1126	1094	1084	1108	1139	1485	2368	2739	2462	3024	2962	3201	3234	3093	4516	4864	3540	3258	3264	2754	2731	3266	3357	3208
Shigella boydii CDC 3083-94 p	264	264	255	668	785	1065	1027	1014	1037	1087	1491	2539	2596	2346	2757	2698	2914	2961	2835	3465	3540	5410	3870	3953	2842	2967	3367	3448	3351
Shigella boydii ATCC 9210	255	260	259	672	752	1049	1027	1016	1035	1060	1403	2235	2537	2273	2810	2741	2959	2993	2894	3226	3259	3869	4637	4447	2600	2796	3206	3245	3273
Shigella boydii Sb227 plasmid	251	255	254	663	744	1030	1012	1001	1017	1041	1415	2353	2521	2289	2768	2697	2905	2936	2847	3258	3264	3951	4446	4800	2591	2922	3237	3365	3325
Shigella flexneri 200 3036	230	235	234	603	707	990	946	938	963	1024	1300	1950	2331	1912	2415	2505	2562	2563	2429	2712	2755	2840	2600	2593	5169	3202	3280	3527	3483
Shigella flexneri 2a str301 plas	233	233	237	583	663	929	916	909	921	936	1511	2613	2482	1920	2329	2460	2451	2424	2361	2772	2733	2965	2795	2922	3198	5215	3360	3737	3604
Shigella flexneri 5 str 8401	253	258	255	662	743	1069	1042	1031	1062	1078	1383	2387	2620	2234	2840	2799	3000	2974	2875	3222	3265	3366	3207	3240	3278	3359	4800	4317	4294
Shigella flexneri 2002017 5 pl	254	256	255	659	739	1049	1031	1021	1052	1065	1416	2509	2603	2334	2821	2764	2969	2947	2876	3320	3356	3445	3244	3366	3526	3738	4316	5369	4761
Shigella flexneri 2a str2457T	253	257	256	664	743	1065	1042	1032	1066	1069	1380	2349	2626	2276	2839	2801	3017	2987	2905	3181	3207	3348	3271	3326	3479	3603	4293	4761	4901

Clustering of ***E. coli*** / ***Shigella*** strains by ribosomal proteins < 200 aa long

Shigella strains are colored by ‘species’ in the left-most column (flexneri orange)



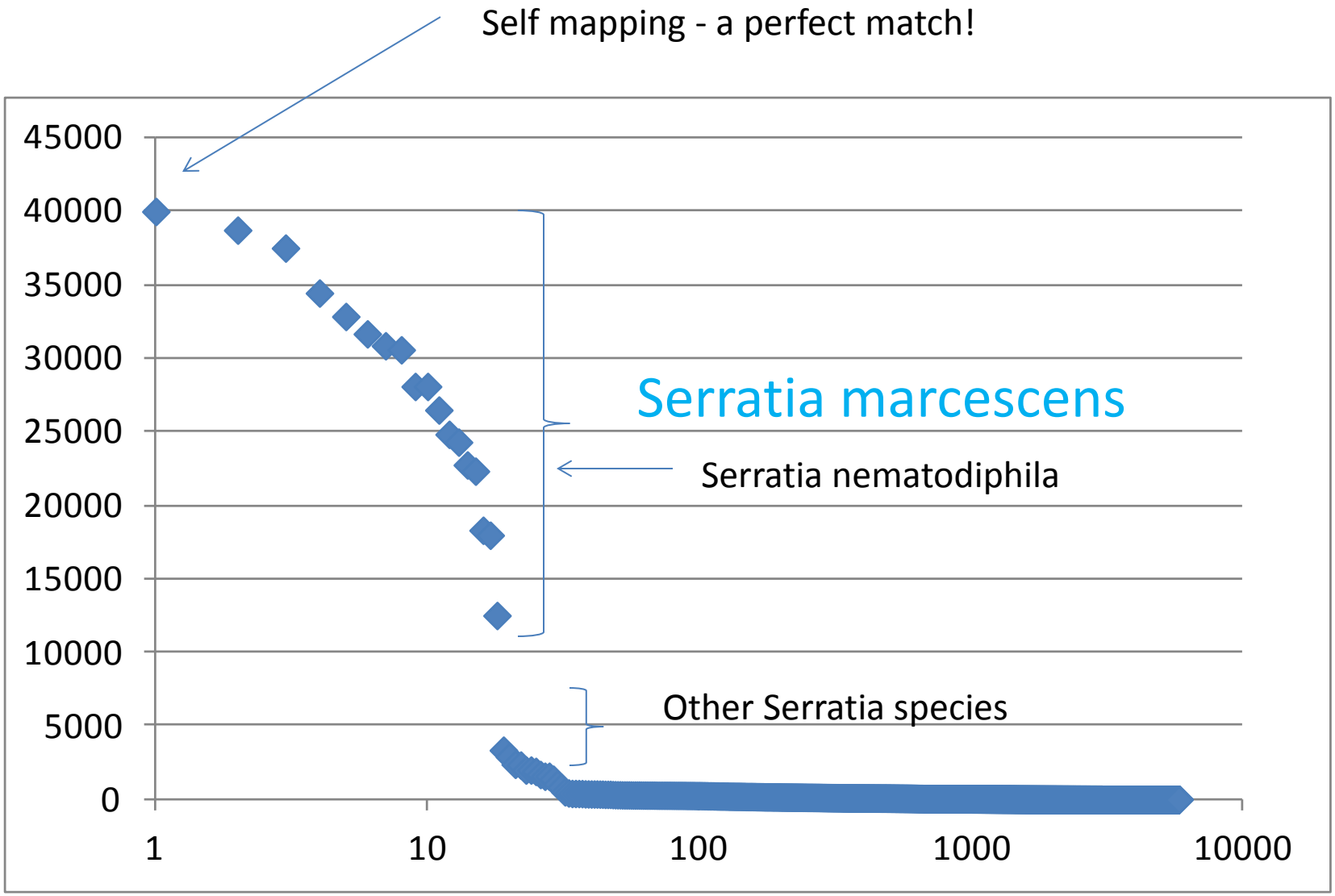
Score Vs. Organism



In silico mapping at 500 ppm

Ribosomal protein theoretical masses

786_SSON, thought to be *Shigella sonnei*



Improving public repositories

- Solution -get NCBI? to develop a ribosomal protein database
  - This was originally envisioned by Pineda et al. in 2003. (Analytical Chem. 75:3817)
  - Many more genomes available (>25000)
  - Can reduce search space by pooling genomes with identical ribosome sequences
    - Lots of identity in
      - *Staphylococcus aureus* **(5584)**
      - *Escherichia coli* **(3947)**
      - *Mycobacterium tuberculosis* **(3612)**

Proving the methodology

- I have succeeded in every case library methods succeed (> 100 species)
- Both library methods and genome matching methods will fail if the organism is not in the database
- Genomic matching can be tested using any peak list (ideally with intensity)
  - Calibration is important
- Genomic matching can also be tested in silico
  - Generate a ribosomal mass list
  - Can now reduce mass tolerance to nearly zero.
- Challenge me with a bacterial ID problem!

Biological confusing factors

- Some bacteria are harder to lyse so as to release ribosomal proteins
- Not all ribosomal proteins are detected
- Probably some differential degradation
- Probably at lag phase ribosomes are sometimes less prominent in the proteome

Informatic Confounding Factors

- Some genomes have mistakes in ribosomal proteins
  - Some subunits entirely missing in some proteomes
    - In E coli and B. subtilis, only ~ ½ of ribosomal proteins are necessary for growth
  - N-terminal extensions or deletions in the databases
  - Some sequences are mis-annotated as ribosomal subunits
- Some ribosomal proteins are present in multiple forms, at least for some clades
  - L31 has a special zinc-deficient form
- Some proteins are modified
  - L33 -> methylated in gammaproteobacteria
- Everything works better without these problems

Conclusions

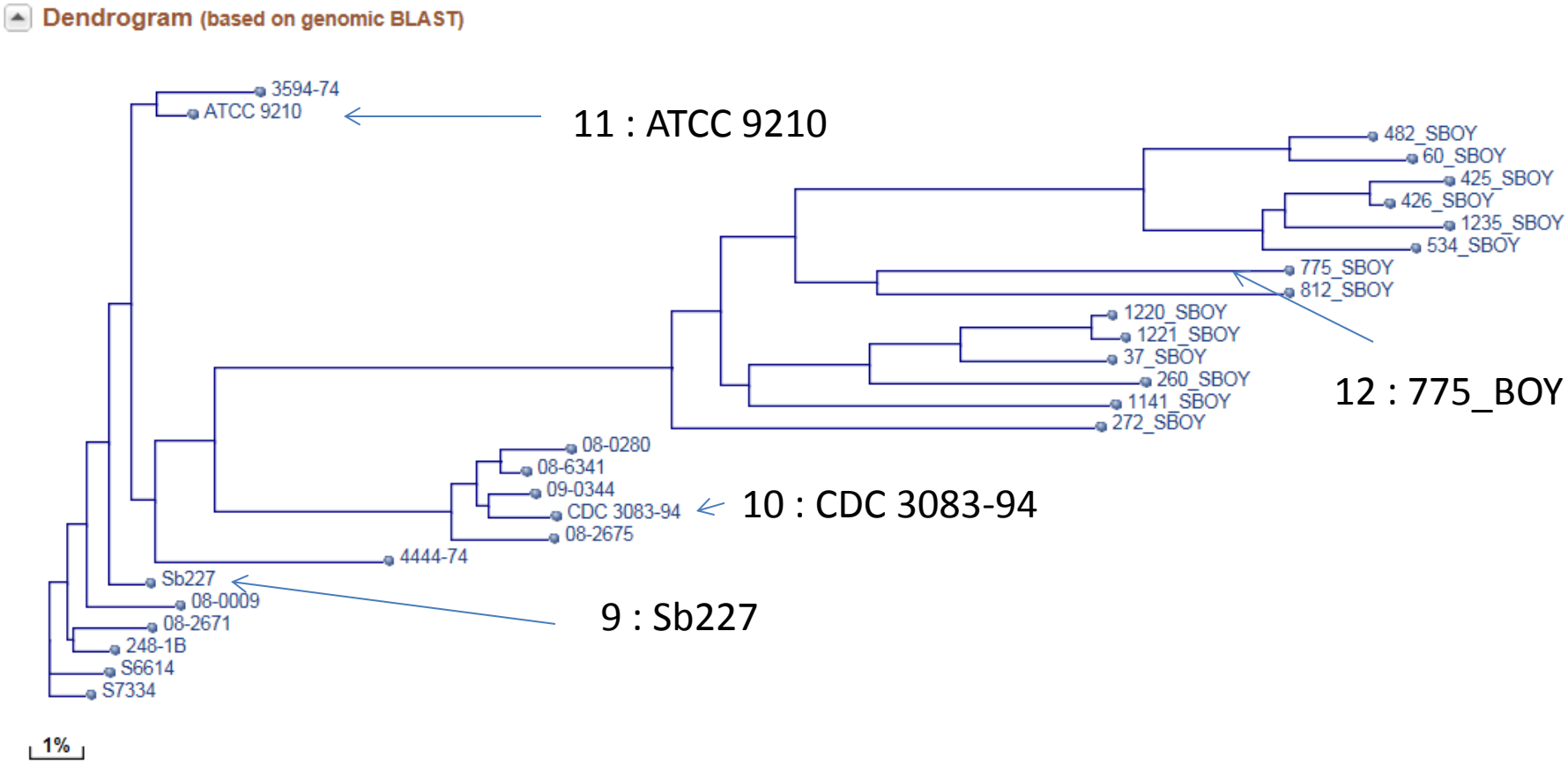
- MALDI can type to sequenced bacteria
- Same method types any sequenced bacteria to all other sequenced bacteria
- Limitations in distinguishing bacteria largely due to taxonomic confusion
  - Same principle works with entire proteome database, but is trickier
  - Can screen for modifications to ribosomal proteins

- Collaborators
  - Kent Voorhees and Chris Cox, Colorado School of Mines
  - Martin Polz, MIT
  - SimulTOF Systems

Shigella boydii

http://www.ncbi.nlm.nih.gov/genome/?term=Shigella+boydii

1-28-2016



Shigella dysenteriae
http://www.ncbi.nlm.nih.gov/genome/?term=Shigella+boydii
1-28-2016

Reference genome: [see all organisms]

- ***Shigella dysenteriae Sd197***

Submitter: Microbial Genome Center of ChMPH

Human Pathogen

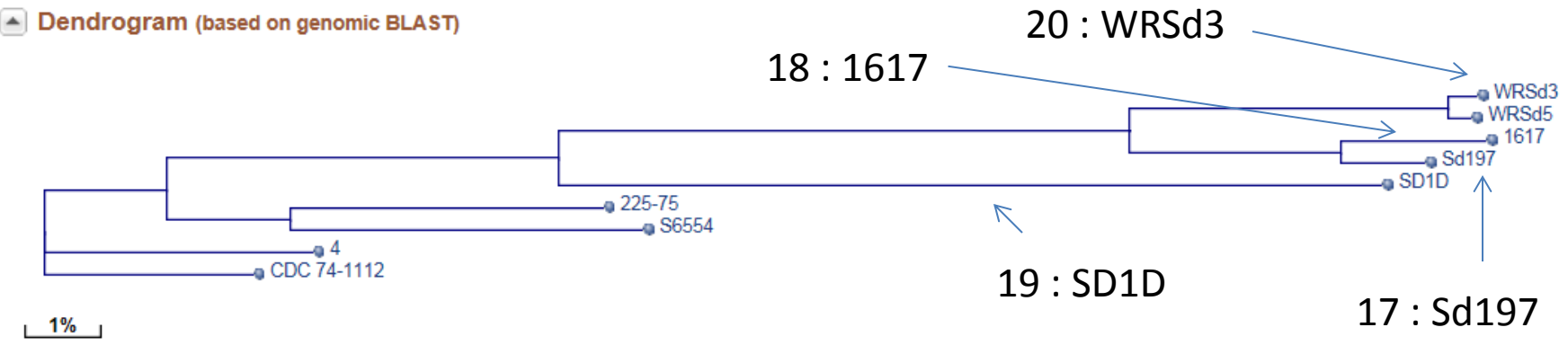
Morphology: Gram:Negative, Shape:Bacilli, Motility:No

Environment: Salinity:NonHalophilic, OxygenReq:Facultative, OptimumTemperature:37, TemperatureRange:Mesophilic, Habitat:HostAssociated

Phenotype: Disease:Dysentery

Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Gene	Pseudogene
Chr	-	NC_007606.1	CP000034.1	4.37	51.2	4,063	22	85	4,602	427
Plsm	pSD1_197	NC_007607.1	CP000035.1	0.182726	44.8	223	-	-	224	1
Plsm	pSD197_spA	NC_009344.1	CP000640.1	0.008953	39.7	8	-	-	8	-

Dendrogram (based on genomic BLAST)

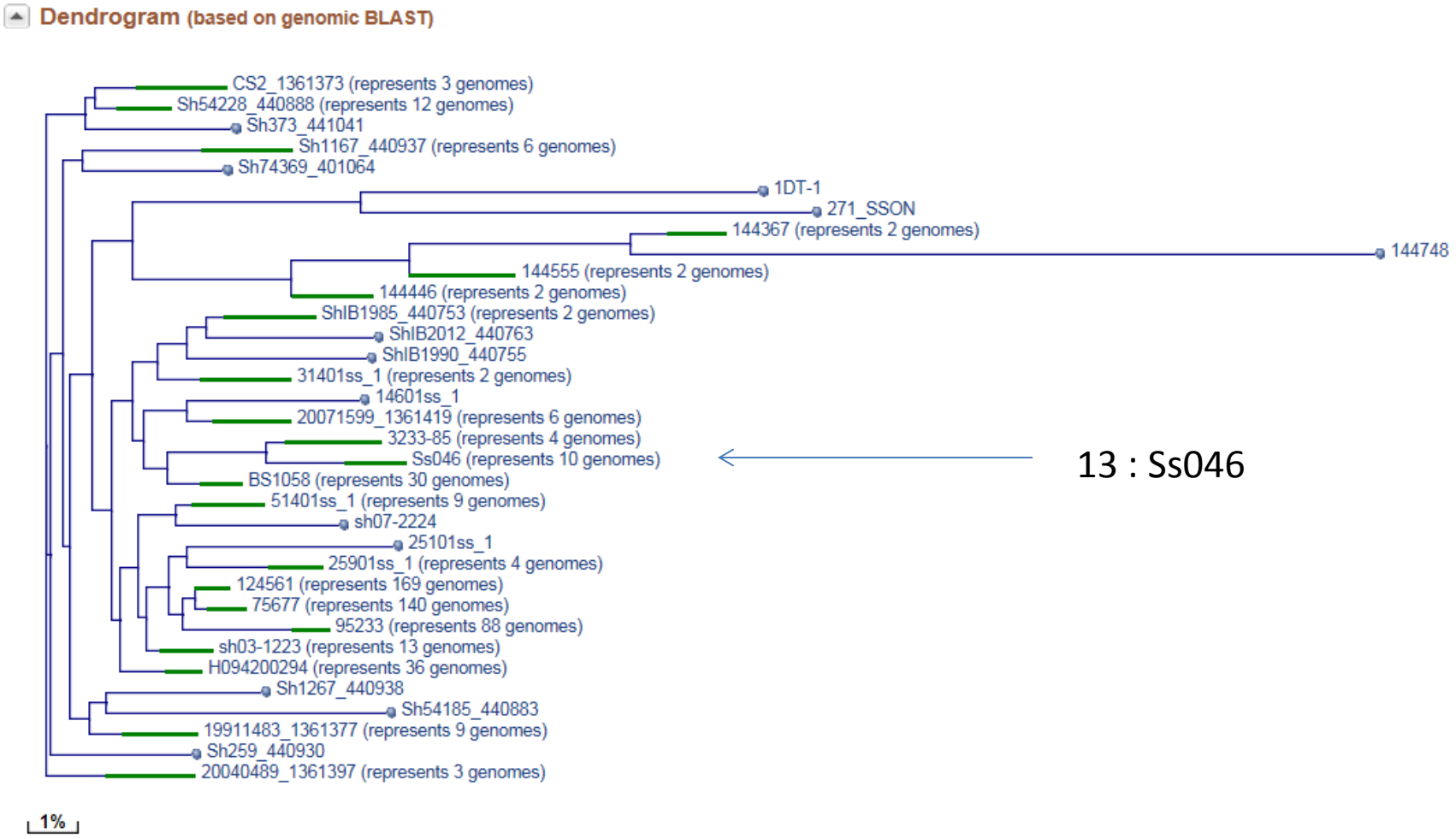


Shigella sonnei

http://www.ncbi.nlm.nih.gov/genome/?term=Shigella+sonnei

1-28-2016

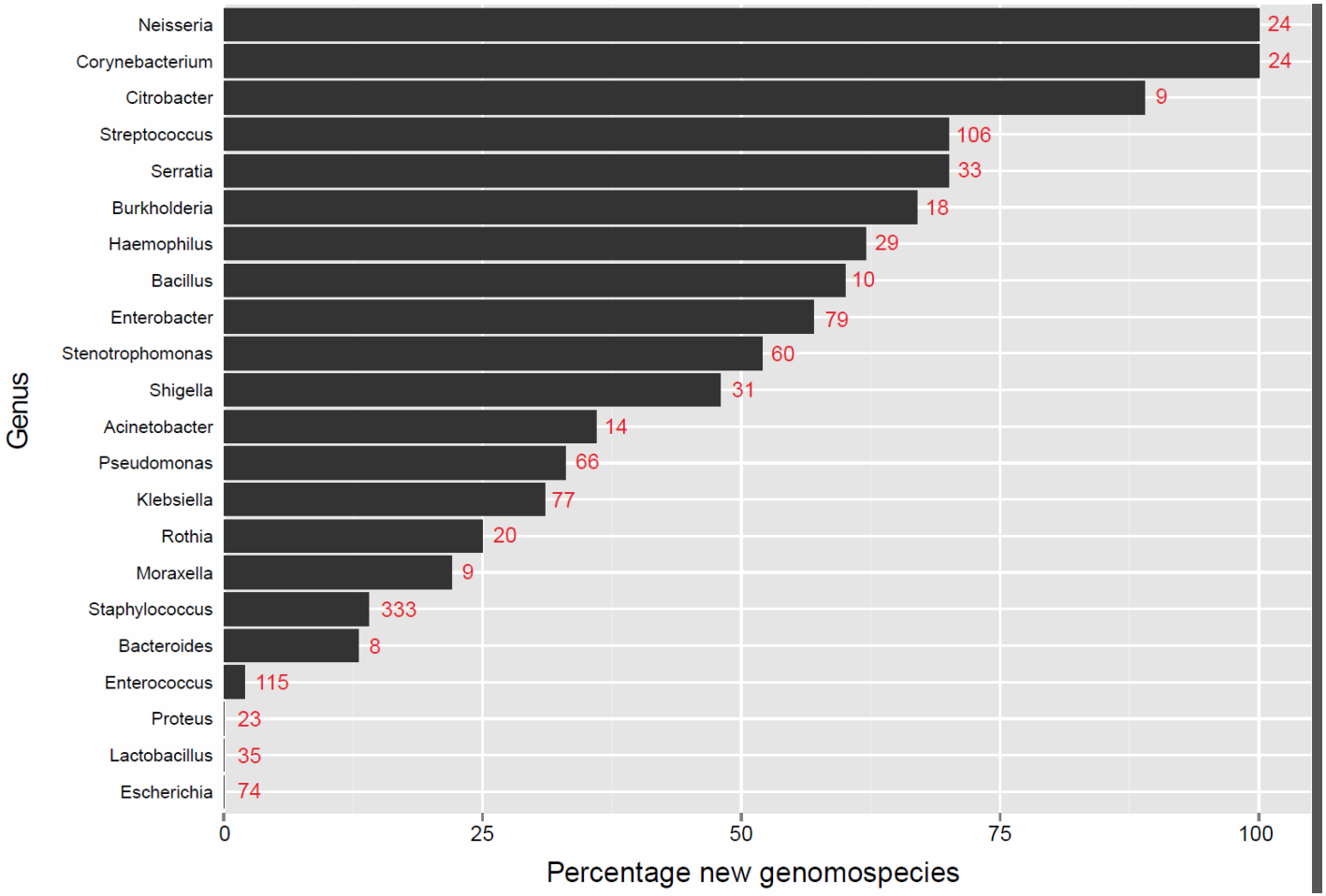
Can't locate the other 3 S. sonnei!



13 : Ss046

A Year of Infection in the Intensive Care Unit:
Prospective Whole Genome Sequencing of Bacterial Clinical Isolates Reveals Cryptic Transmissions and Novel Microbiota Figure S3
David J. Roach et al; http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4521703/

All genera with >5 isolates sequenced are indicated.
The x-axis indicates the proportion of each genus that consisted of novel genomospecies by ANIb analysis.
The red number to the right of each bar indicates the total number of isolates sampled from each genus in this study.



Serratia marcescens 290 genomes 2-9-2016

- ***Serratia marcescens FGI94***

Submitter: UW-Madison

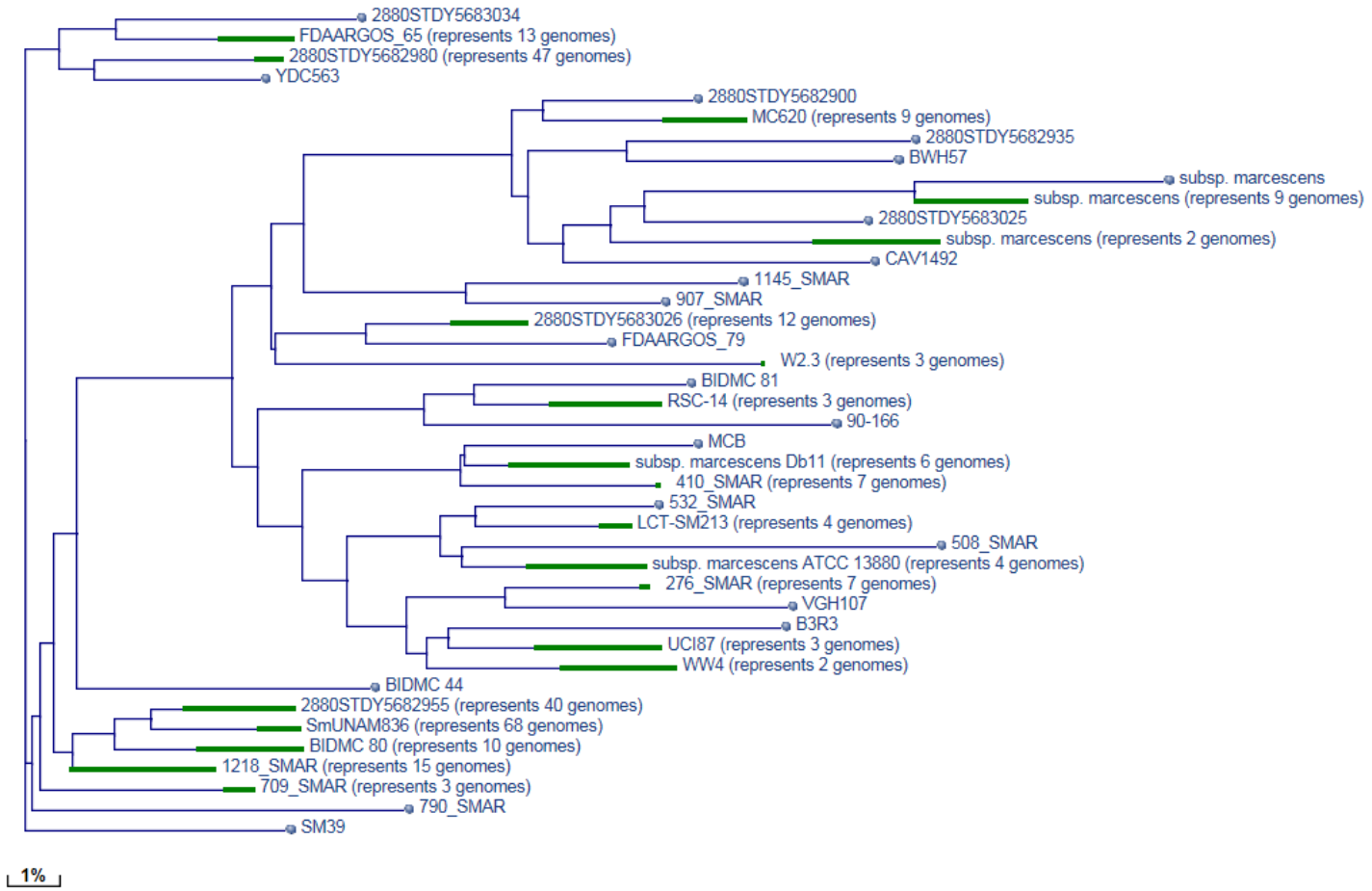
Morphology: Gram:Negative, Shape:Bacilli, Endospores:No, Motility:Yes

Environment: Salinity:NonHalophilic, OxygenReq:Facultative, TemperatureRange:Mesophilic, Habitat:HostAssociated

Phenotype: BioticRelationship:Episymbiont, TrophicLevel:Heterotroph, Disease:NA

Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Other RNA	Gene	Pseudogene
	-	NC_020064.1	CP003942.1	4.86	58.9	4,290	22	83	3	4,436	38

Dendrogram (based on genomic BLAST)

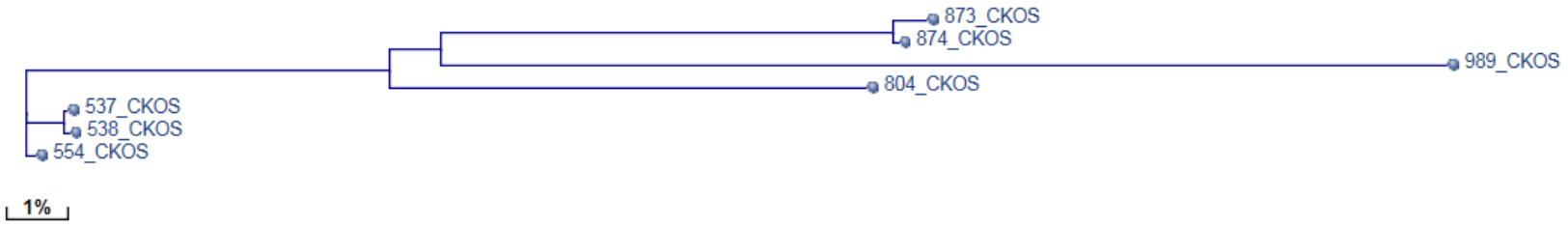


Citrobacter koseri 14 genomes 2-9-2016

Replicon Info

Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Gene	Pseudogene
	-	NC_009792.1	CP000822.1	4.72	53.8	4,222	22	83	4,377	50
	pCKO3	NC_009793.1	CP000823.1	0.01	55.2	11	-	-	11	-
	pCKO2	NC_009794.1	CP000824.1	0.01	51.3	8	-	-	8	-

Dendrogram (based on genomic BLAST)



Citrobacter freundii 52 genomes 2-9-2016

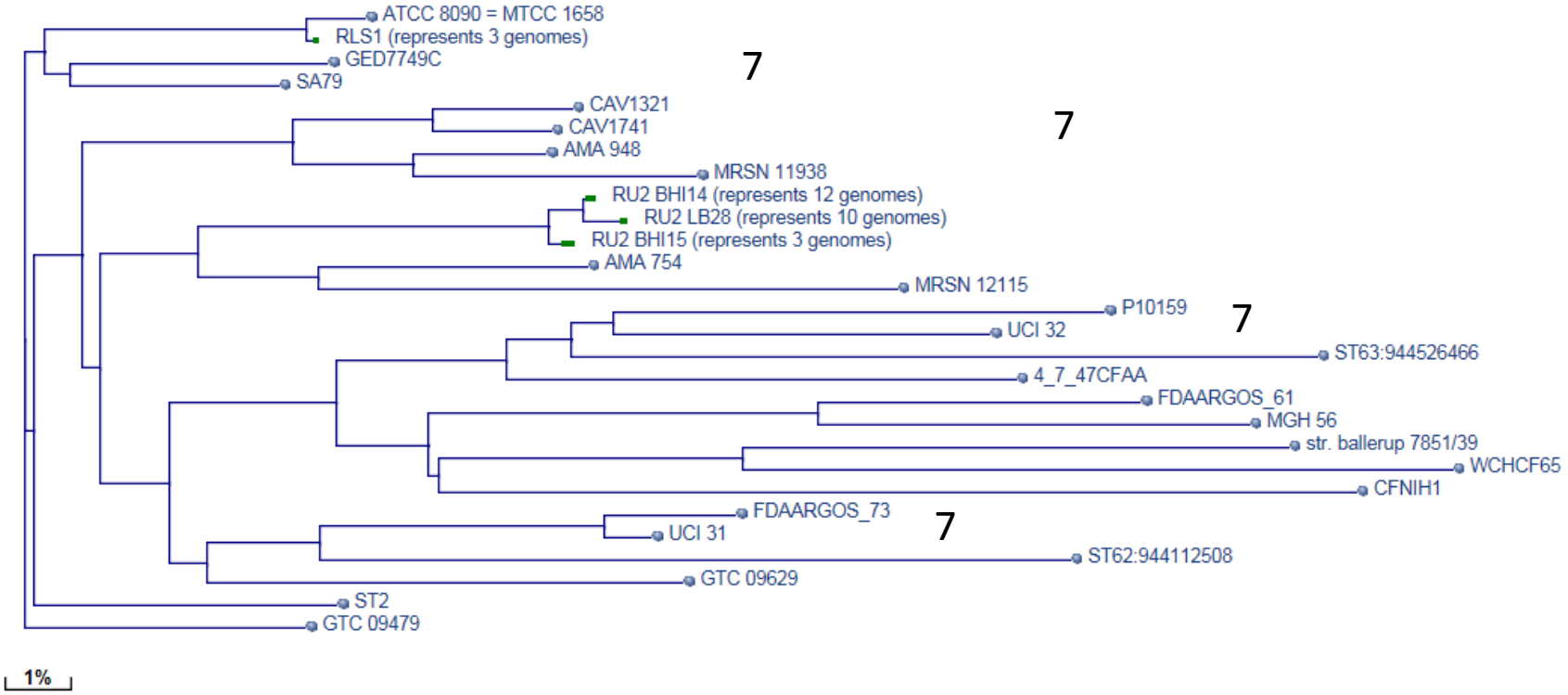
Representative genome: [see all organisms]

- ***Citrobacter freundii CFNIH1***

Submitter: NISC

Type	Name	RefSeq	INSDC	Size (Mb)	GC%	Protein	rRNA	tRNA	Other RNA	Gene	Pseudogene
Chr	-	NZ_CP007557.1	CP007557.1	5.1	52.2	4,672	25	83	1	4,847	66
Plsm	pKEC-a3c	NZ_CP007558.1	CP007558.1	0.272297	52.7	298	-	-	-	309	11

Dendrogram (based on genomic BLAST)



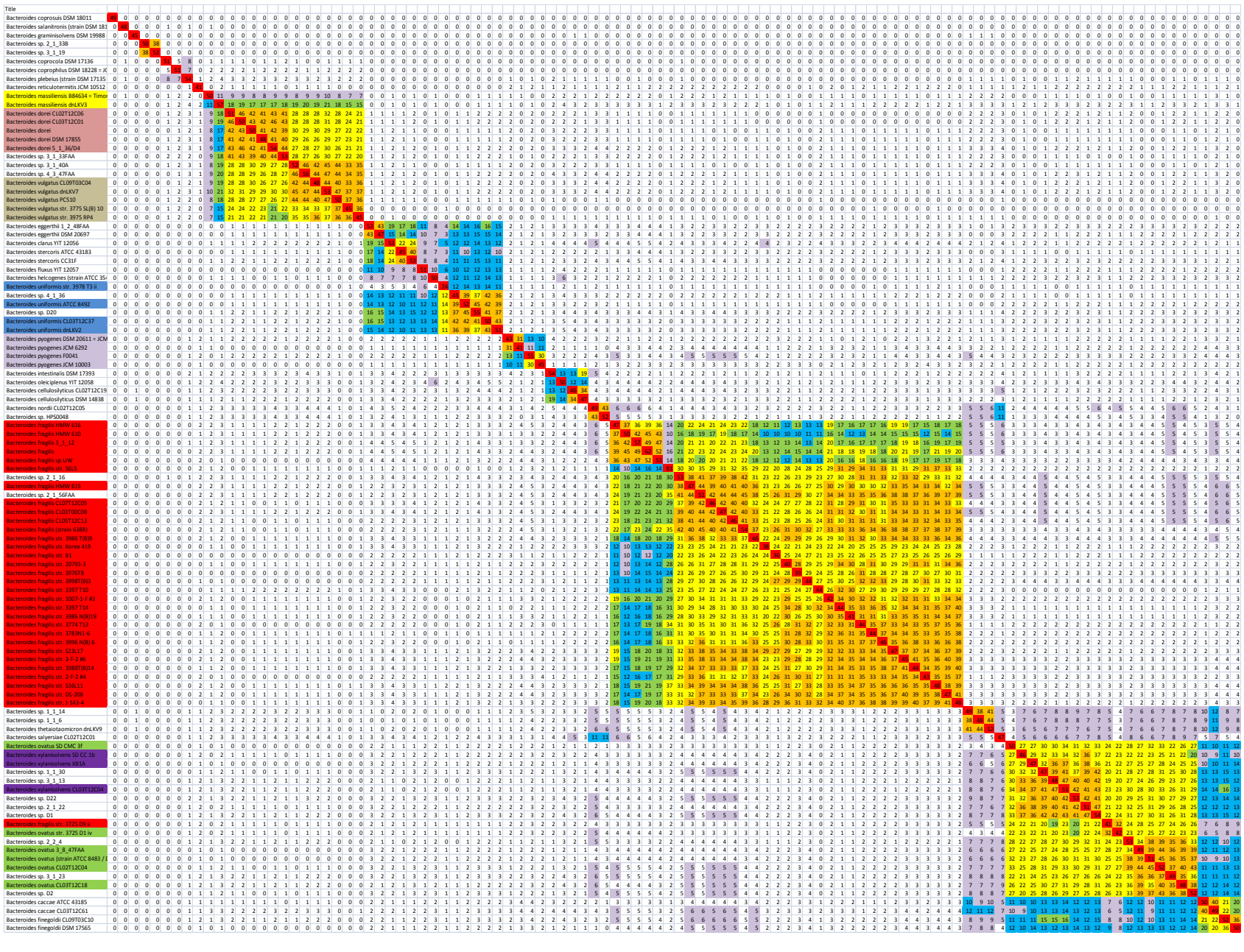
Intermingling of strains mapped to Citrobacte species at NCBI

I	Title	NCBI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	ribo
1	Citrobacter koseri		59	52	54	54	54	22	19	25	23	11	11	12	12	12	10	11	11	11	11	11	11	11	11	11	11	11	11	10	13	10	12	11	1
2	Citrobacter koseri 661_CKOS	19669	52	59	57	57	57	24	20	27	24	11	9	11	11	11	8	10	10	10	10	10	10	10	10	10	10	10	10	9	8	12	9	11	11	1
3	Citrobacter koseri FDAARGOS_86	19669	54	57	59	59	59	24	20	27	24	12	10	11	12	12	9	11	11	11	11	11	11	11	11	11	11	11	11	10	9	12	9	11	11	1
4	Citrobacter koseri ATCC BAA-895	19669	54	57	59	59	59	24	20	27	24	12	10	11	12	12	9	11	11	11	11	11	11	11	11	11	11	11	11	10	9	12	9	11	11	1
5	Citrobacter koseri FDAARGOS_164	19669	54	57	59	59	59	24	20	27	24	12	10	11	12	12	9	11	11	11	11	11	11	11	11	11	11	11	11	10	9	12	9	11	11	1
6	Citrobacter rodentium (strain ICC168)		22	24	24	24	24	54	21	28	23	12	10	11	10	10	9	12	12	12	12	12	12	12	12	12	12	12	12	11	13	11	12	13	2
7	Citrobacter amalonaticus Y19		19	20	20	20	20	21	56	36	35	7	12	13	13	13	9	15	15	15	15	15	15	15	15	15	15	15	15	14	13	14	9	13	14	3
8	Citrobacter farmeri GTC 1319		25	27	27	27	27	28	36	56	38	9	12	13	12	12	9	13	13	13	13	13	13	13	13	13	13	13	13	13	12	17	13	13	14	3
9	Citrobacter sp. MGH 55		23	24	24	24	24	23	35	38	55	6	12	13	12	12	9	13	13	13	13	13	13	13	13	13	13	13	13	13	12	15	10	13	14	3
10	Citrobacter koseri 1243_CKOS (K. pneumoniae)	19672	11	11	12	12	12	12	7	9	6	59	7	7	8	8	6	9	9	9	9	9	9	9	9	9	9	9	9	9	9	7	5	6	6	4
11	Citrobacter freundii MGH 56	21126	11	9	10	10	10	10	12	12	12	7	53	46	44	44	37	42	41	41	40	41	41	41	41	41	41	41	41	41	39	27	32	33	34	5
12	Citrobacter sp. KTE151		12	11	11	11	11	11	13	13	13	7	46	58	52	52	38	44	43	43	42	43	43	43	43	43	42	43	43	43	40	29	34	36	37	5
13	Citrobacter freundii 980_SSON	21126	12	11	12	12	12	10	13	12	12	8	44	52	57	57	36	46	45	45	44	45	45	44	44	44	44	45	45	40	37	27	31	34	38	5
14	Citrobacter koseri 989_CKOS	21126	12	11	12	12	12	10	13	12	12	8	44	52	57	57	36	46	45	45	44	45	45	44	44	44	44	45	45	40	37	27	31	34	38	5
15	Citrobacter sp. CIP 55.13		10	8	9	9	9	9	9	9	9	6	37	38	36	36	50	39	39	39	38	39	39	40	40	40	40	40	40	38	33	25	32	31	32	6
16	Citrobacter freundii P10159	21126	11	10	11	11	11	12	15	13	13	9	42	44	46	46	39	59	54	54	54	55	55	55	55	55	55	56	56	49	46	30	34	37	41	7
17	Citrobacter freundii CAV1321	21126	11	10	11	11	11	12	15	13	13	9	41	43	45	45	39	54	58	58	56	57	57	55	55	55	55	56	56	50	44	29	35	37	41	7
18	Citrobacter freundii CAV1741	21126	11	10	11	11	11	12	15	13	13	9	41	43	45	45	39	54	58	58	56	57	57	55	55	55	55	56	56	50	44	29	35	37	41	7
19	Citrobacter freundii ST2	21126	11	10	11	11	11	12	15	13	13	9	40	42	44	44	38	54	56	56	57	57	57	55	55	55	55	56	56	50	43	28	34	36	40	7
20	Citrobacter koseri 873_CKOS	21126	11	10	11	11	11	12	15	13	13	9	41	43	45	45	39	55	57	57	57	59	59	56	56	56	56	57	57	51	44	29	35	37	41	7
21	Citrobacter koseri 874_CKOS	21126	11	10	11	11	11	12	15	13	13	9	41	43	45	45	39	55	57	57	57	59	59	56	56	56	56	57	57	51	44	29	35	37	41	7
22	Citrobacter koseri 554_CKOS	21126	11	10	11	11	11	12	15	13	13	9	41	43	44	44	40	55	55	55	55	56	56	59	59	59	57	58	58	50	44	30	35	37	40	7
23	Citrobacter koseri 537_CKOS	21126	11	10	11	11	11	12	15	13	13	9	41	43	44	44	40	55	55	55	55	56	56	59	59	59	57	58	58	50	44	30	35	37	40	7
24	Citrobacter koseri 538_CKOS	21126	11	10	11	11	11	12	15	13	13	9	41	43	44	44	40	55	55	55	55	56	56	59	59	59	57	58	58	50	44	30	35	37	40	7
25	Citrobacter freundii SA79	21126	11	10	11	11	11	12	15	13	13	9	41	42	44	44	40	55	55	55	55	56	56	57	57	57	59	58	58	49	43	29	34	36	40	7
26	Citrobacter freundii GTC09479	21126	11	10	11	11	11	12	15	13	13	9	41	43	45	45	40	56	56	56	56	57	57	58	58	58	58	59	59	50	44	30	35	37	41	7
27	Citrobacter koseri 804_CKOS	21126	11	10	11	11	11	12	15	13	13	9	41	43	45	45	40	56	56	56	56	57	57	58	58	58	58	59	59	50	44	30	35	37	41	7
28	Citrobacter freundii UCI 31	21126	11	9	10	10	10	12	14	13	13	9	41	43	40	40	38	49	50	50	50	51	51	50	50	50	49	50	50	55	45	28	34	34	35	7
29	Citrobacter freundii UCI 32	21126	10	8	9	9	9	11	13	12	12	9	39	40	37	37	33	46	44	44	43	44	44	44	44	44	43	44	44	45	49	27	31	33	34	8
30	Citrobacter youngae ATCC 29220		13	12	12	12	12	13	14	17	15	7	27	29	27	27	25	30	29	29	28	29	29	30	30	30	29	30	30	28	27	51	33	32	34	9
31	Citrobacter werkmanii NBRC 105721		10	9	9	9	9	11	9	13	10	5	32	34	31	31	32	34	35	35	34	35	35	35	35	35	34	35	35	34	31	33	52	43	43	10
32	Citrobacter freundii CFNIH1	21126	12	11	11	11	11	12	13	13	13	6	33	36	34	34	31	37	37	37	36	37	37	37	37	37	36	37	37	34	33	32	43	56	53	10
33	Citrobacter freundii CFNIH1 2-8-2016	21126	11	11	11	11	11	13	14	14	14	6	34	37	38	38	32	41	41	41	40	41	41	40	40	40	40	41	41	35	34	34	43	53	58	10

Bacteroides species differentiation by ribosomal proteins

Roughly, orange means same species

fragilis



Neisseria meningitidis and relatives