Bacterial taxonomy: Current developments

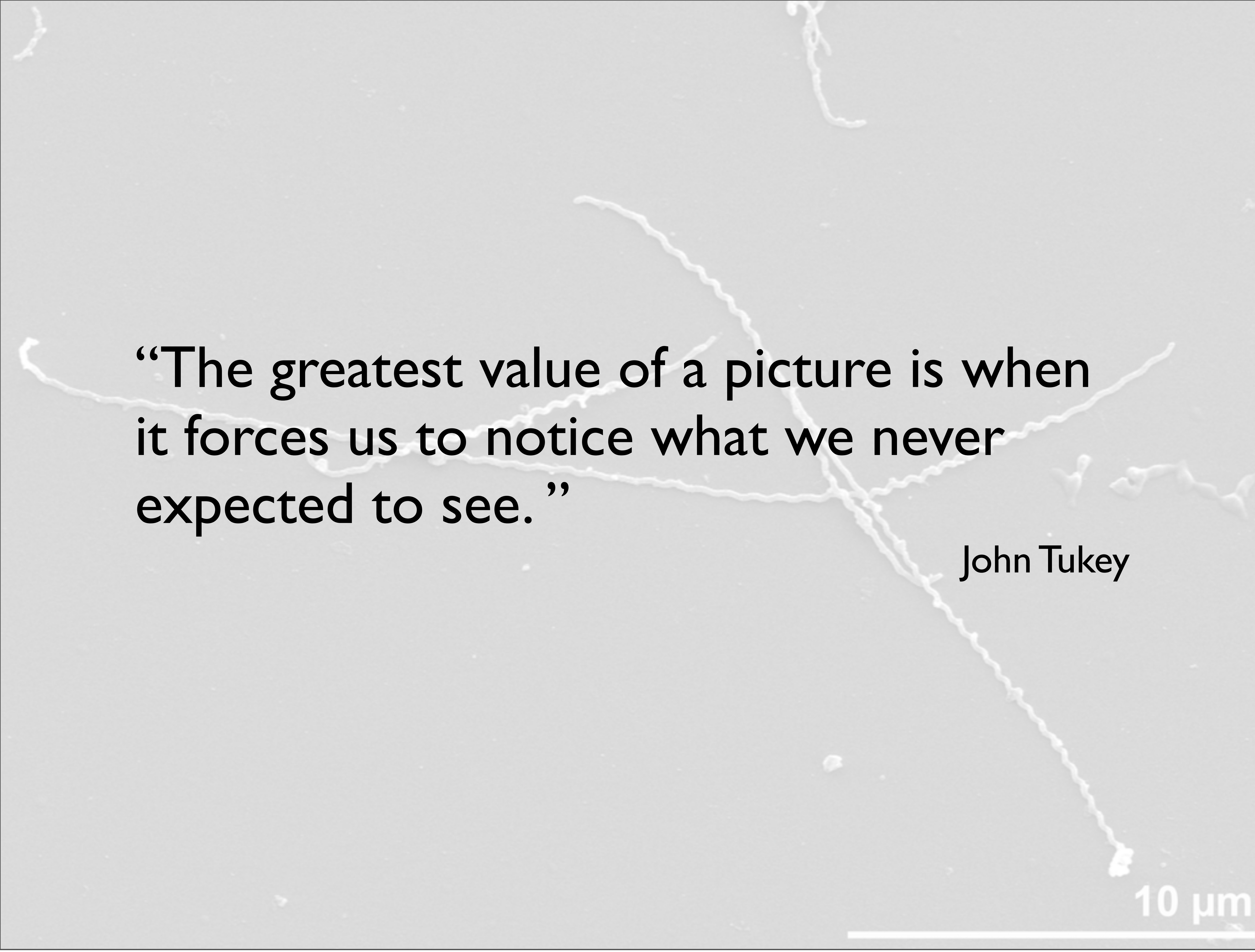
George M. Garrity

MICHIGAN STATE UNIVERSITY

John W. Tukey
1915 - 2000

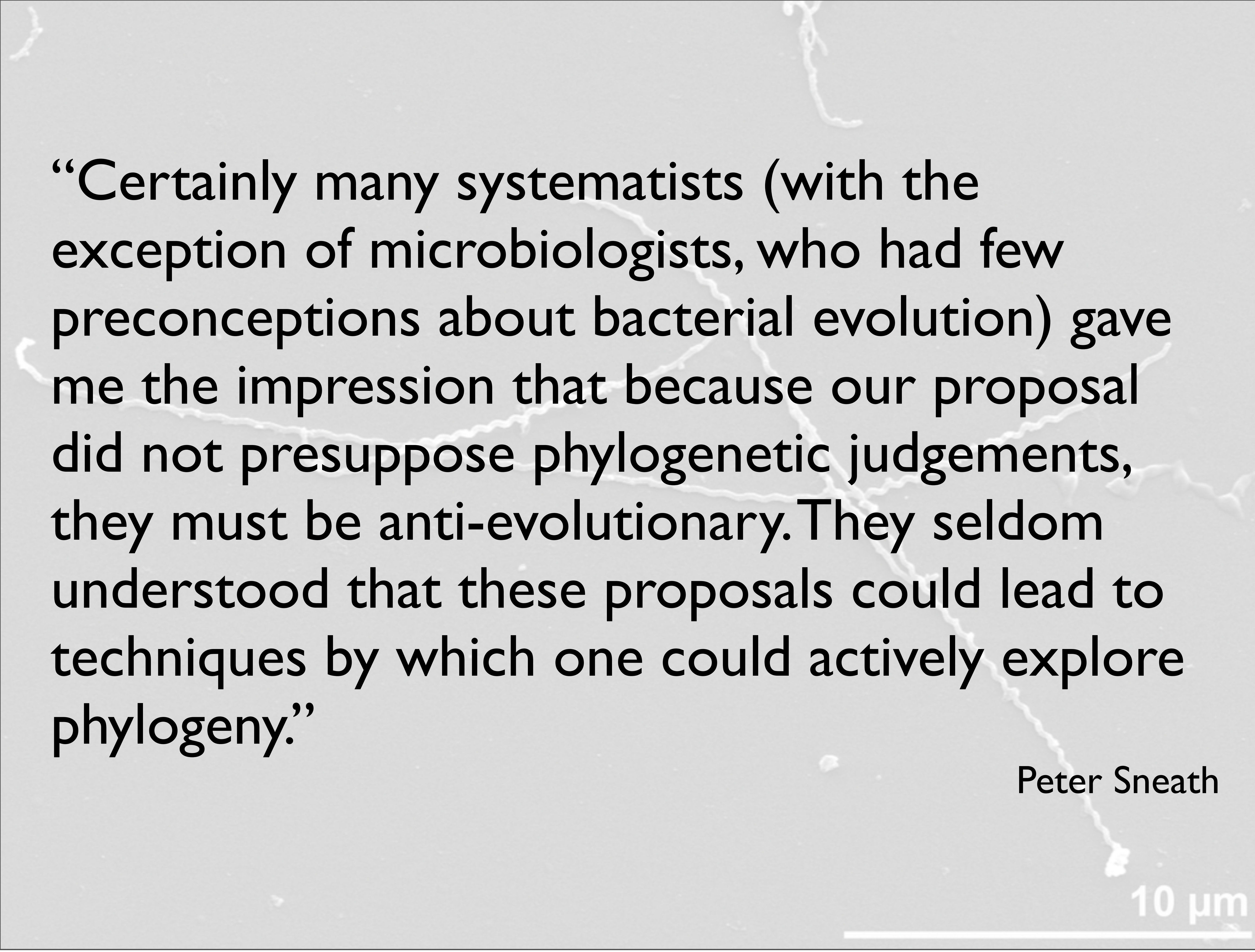
Peter H.A. Sneath
1923 - 2011

Samuel T. Cowan
1905 - 1976



“The greatest value of a picture is when it forces us to notice what we never expected to see.”

John Tukey



"Certainly many systematists (with the exception of microbiologists, who had few preconceptions about bacterial evolution) gave me the impression that because our proposal did not presuppose phylogenetic judgements, they must be anti-evolutionary. They seldom understood that these proposals could lead to techniques by which one could actively explore phylogeny."

Peter Sneath

"Sam Cowan was determined to show that it was possible to distribute cultures that were authentic, pure and typical: the fact that this is now broadly accepted as the normal practice is a measure of his success."

"The proposal to make a new start for bacterial names in 1980 gave Cowan much quiet satisfaction."

"He had the unusual ability to make taxonomy interesting to other people."

Peter Sneath

Crowdsourcing the global prokaryotic taxonomy

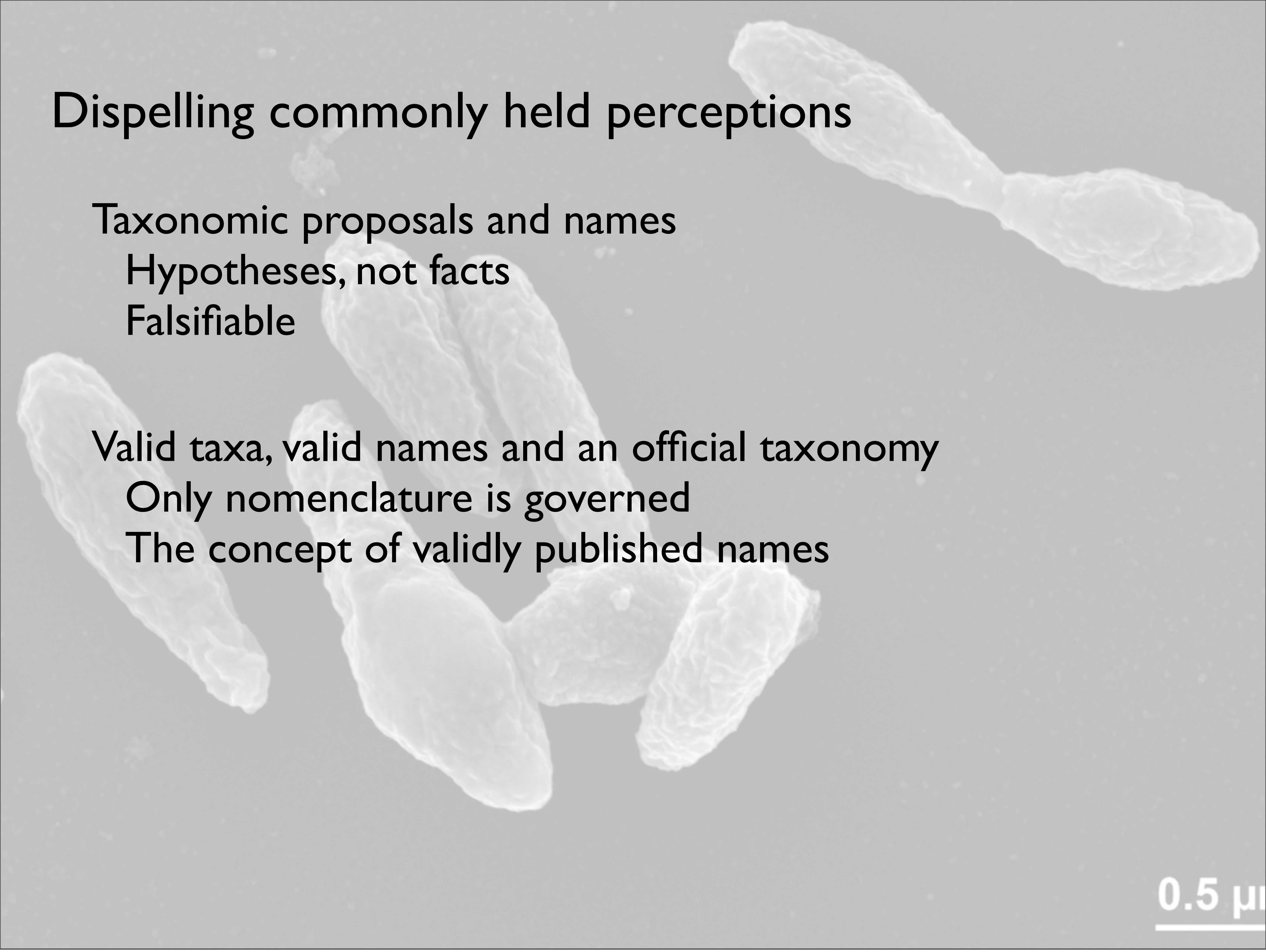
The current 16S rRNA based taxonomy

- In development for 35 years
- Work product of > 17,500 authors
- > 20,600 taxonomic description
- 12,195 effective publications
- Earliest taxa > 175 yrs
  - Average age 16.8 yrs
  - 13.8% synonyms
  - 7.8% explicitly emended

Dispelling commonly held perceptions

Taxonomic proposals and names
- Hypotheses, not facts
- Falsifiable

Valid taxa, valid names and an official taxonomy
- Only nomenclature is governed
- The concept of validly published names



Antonie van Leeuwenhoek (2014) 106:43–56
DOI 10.1007/s10482-013-0084-1

INVITED REVIEW ARTICLE

Antonie van Leeuwenhoek 80th Anniversary Issue

Then and now: a systematic review of the systematics of prokaryotes in the last 80 years

Aharon Oren · George M. Garrity

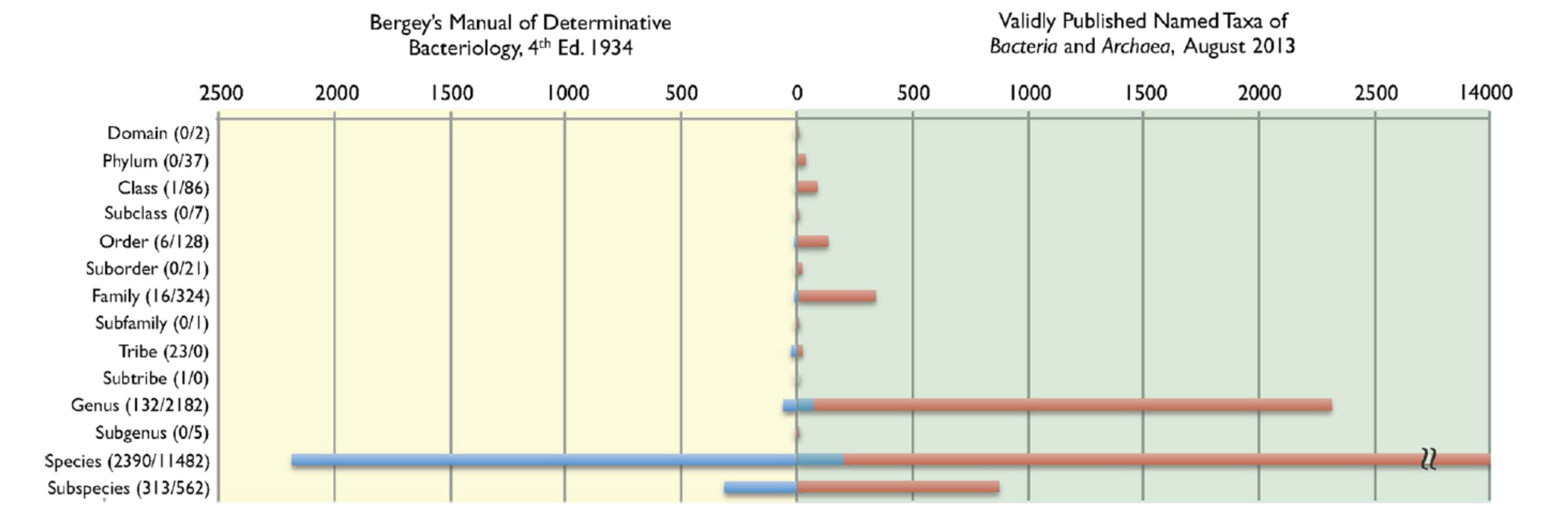
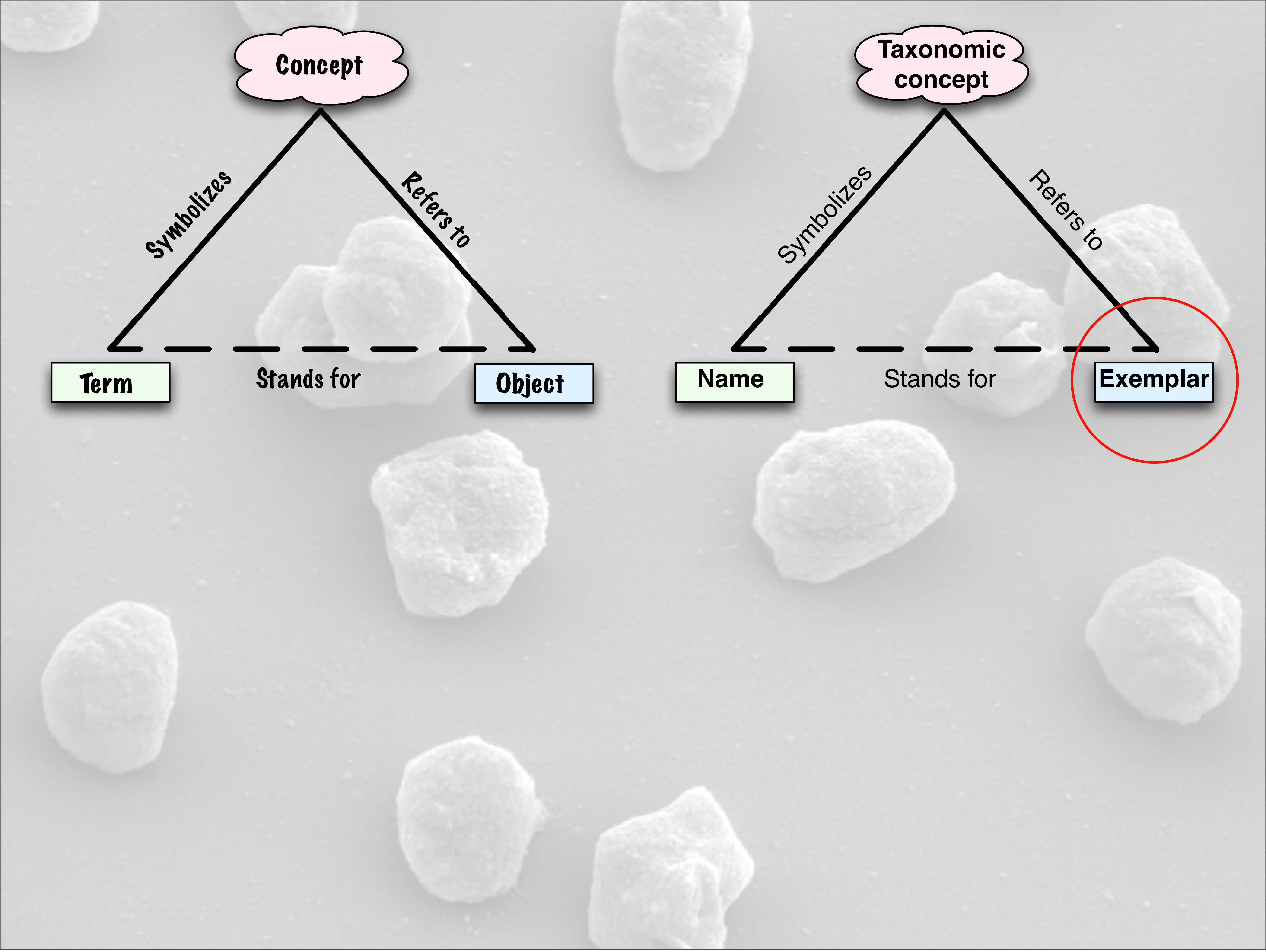
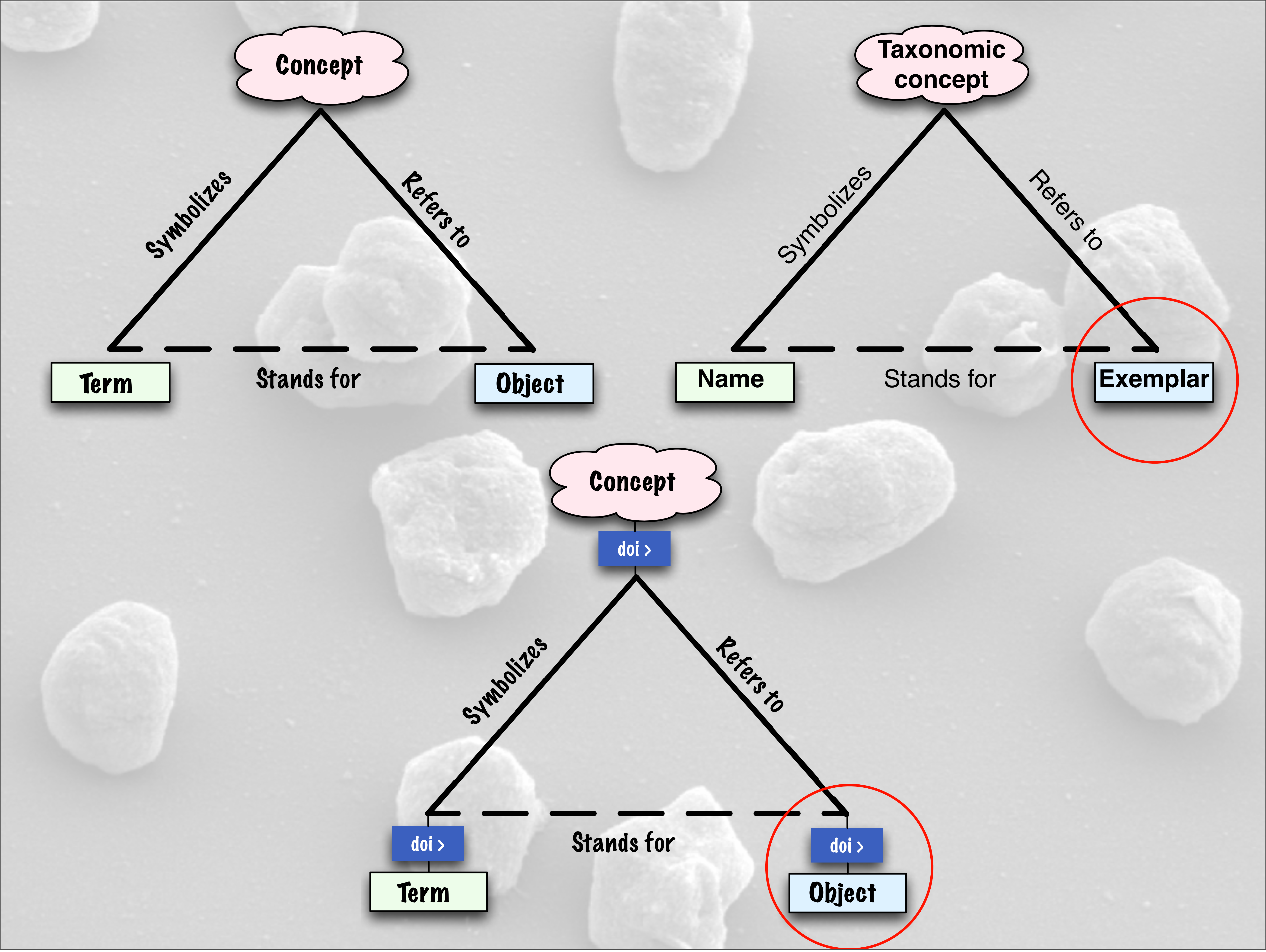
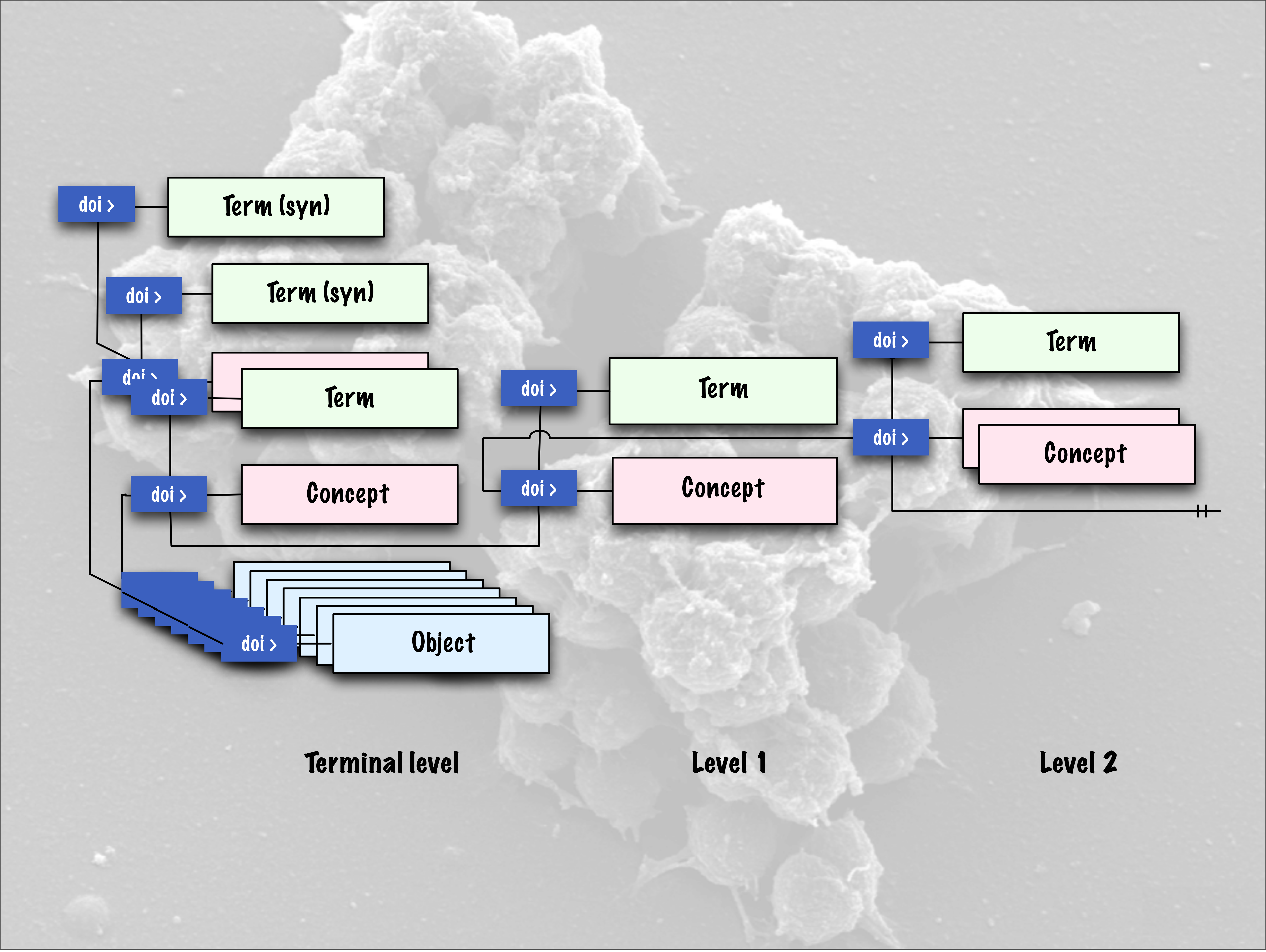


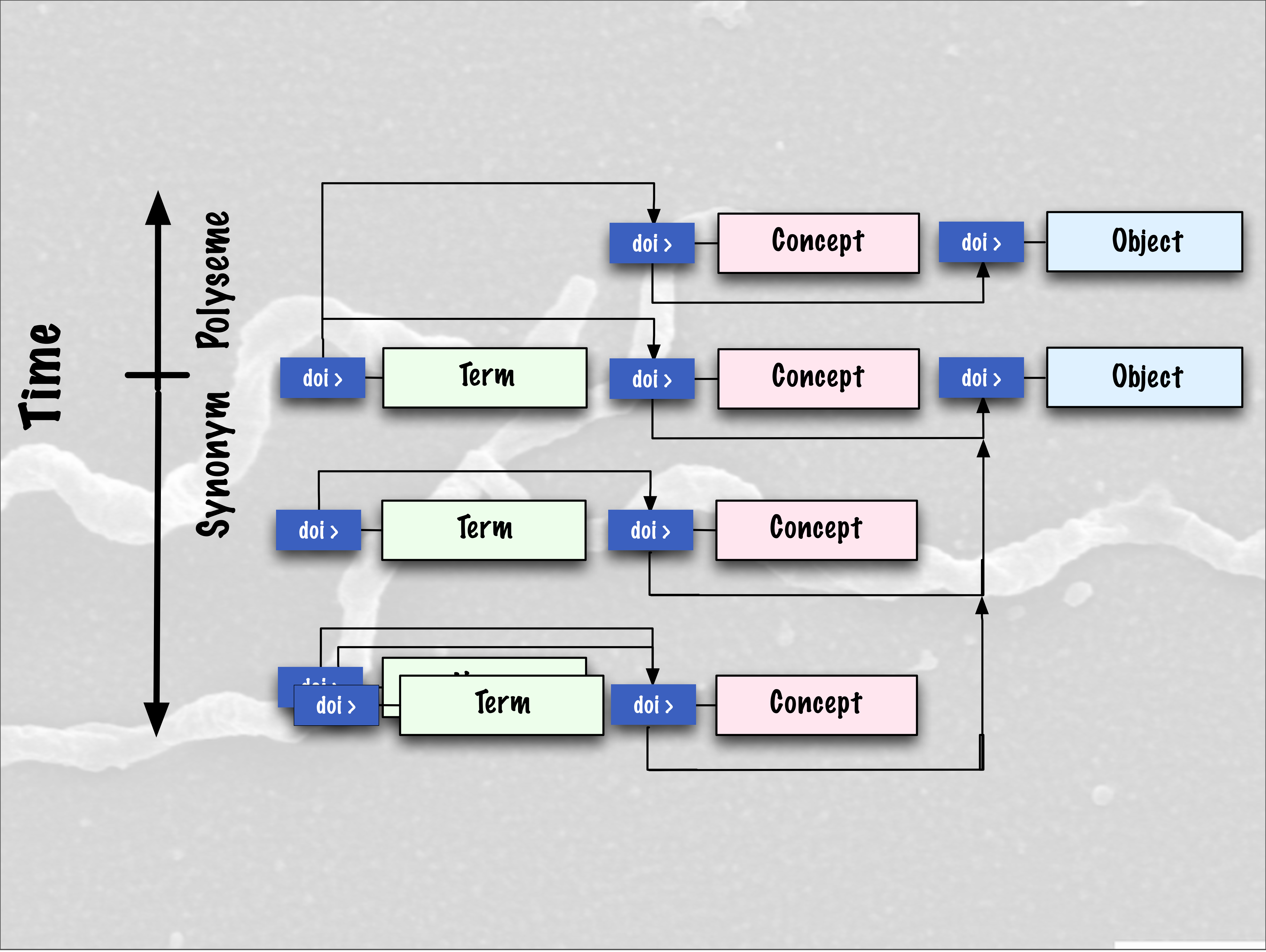
Fig. 1 Then and now. A comparison of the accepted taxonomies in use at the time of the first and current issues of *Antonie van Leeuwenhoek*. Printed with permission.

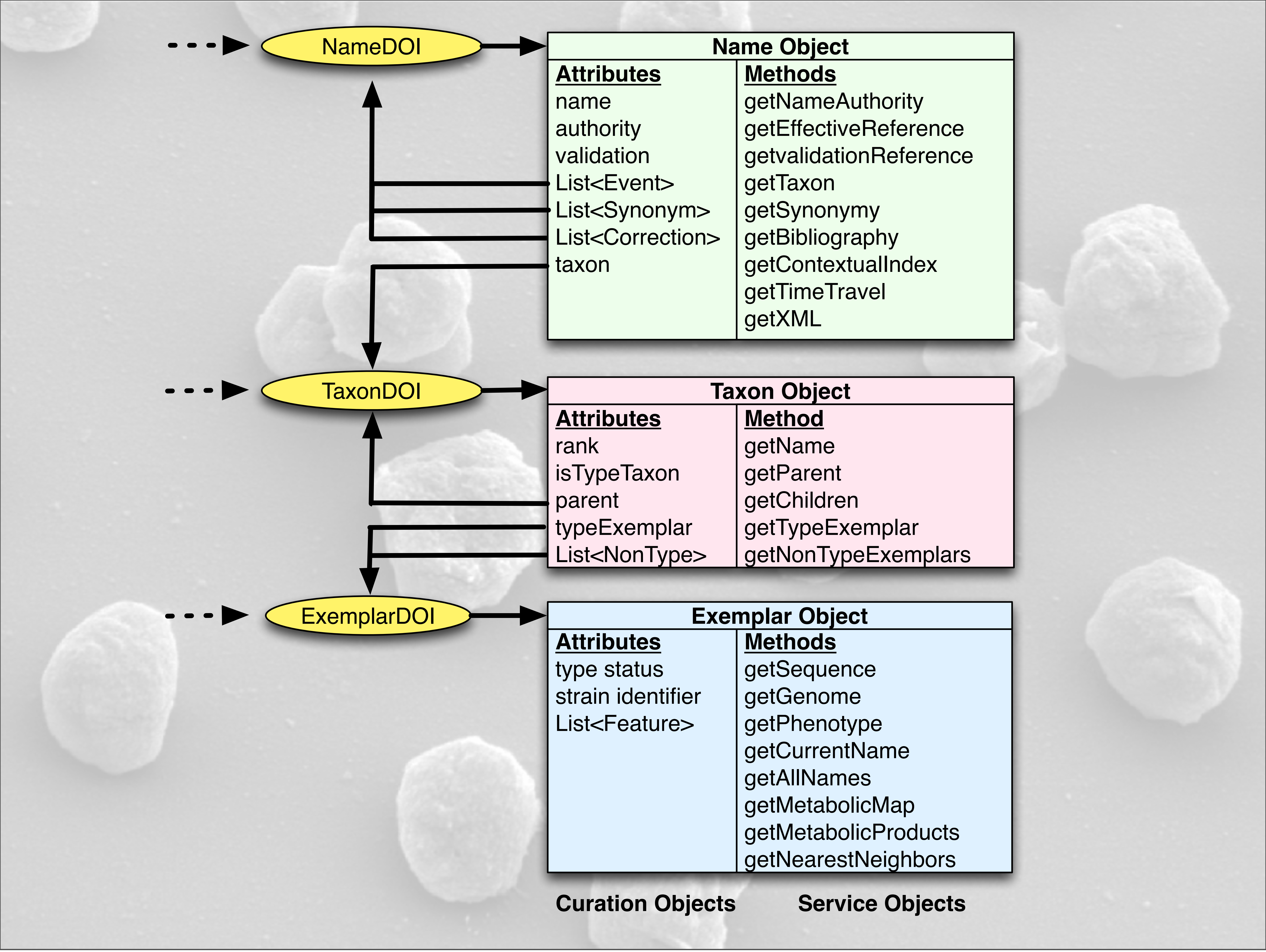
2 µm


Concept
Symbolizes
Refers to
Term
Stands for
Object
Taxonomic concept
Symbolizes
Refers to
Name
Stands for
Exemplar

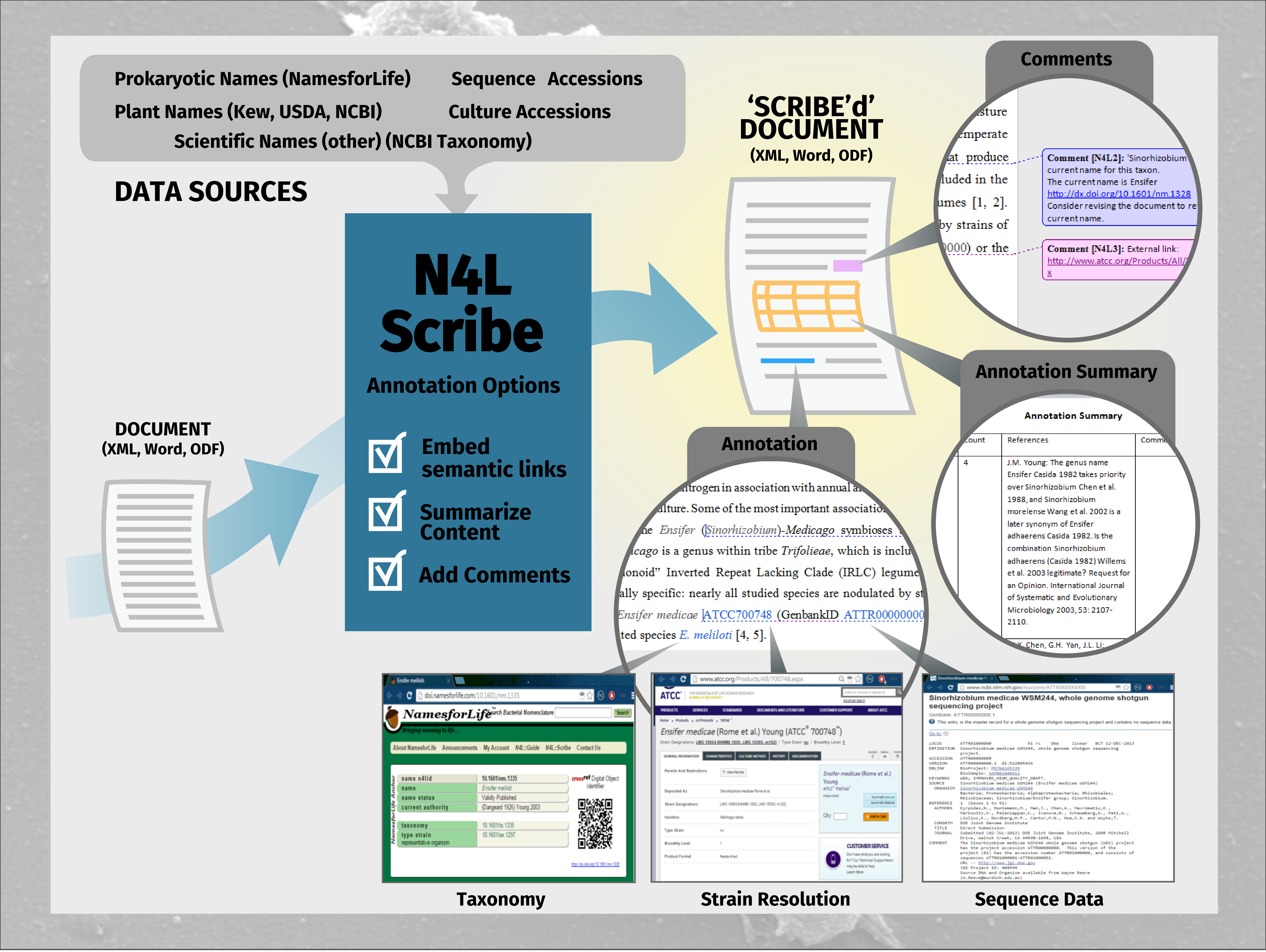

Concept
Symbolizes
Refers to
Term
Stands for
Object
Taxonomic concept
Symbolizes
Refers to
Name
Stands for
Exemplar
Concept
doi >
Symbolizes
Refers to
doi >
Term
Stands for
doi >
Object


doi >
Term (syn)
doi >
Term (syn)
doi >
Term
doi >
Concept
doi >
Object
doi >
Term
doi >
Concept
doi >
Term
doi >
Concept
Terminal level
Level 1
Level 2


Time
Polyseme
Synonym
doi >
Concept
doi >
Object
doi >
Term
doi >
Concept
doi >
Object
doi >
Term
doi >
Concept
doi >
Term
doi >
Concept



NameDOI
Name Object
Attributes
name
authority
validation
List<Event>
List<Synonym>
List<Correction>
taxon
Methods
getNameAuthority
getEffectiveReference
getvalidationReference
getTaxon
getSynonymy
getBibliography
getContextualIndex
getTimeTravel
getXML
TaxonDOI
Taxon Object
Attributes
rank
isTypeTaxon
parent
typeExemplar
List<NonType>
Method
getName
getParent
getChildren
getTypeExemplar
getNonTypeExemplars
ExemplarDOI
Exemplar Object
Attributes
type status
strain identifier
List<Feature>
Methods
getSequence
getGenome
getPhenotype
getCurrentName
getAllNames
getMetabolicMap
getMetabolicProducts
getNearestNeighbors
Curation Objects
Service Objects



Prokaryotic Names (NamesforLife)
Sequence Accessions
Plant Names (Kew, USDA, NCBI)
Culture Accessions
Scientific Names (other) (NCBI Taxonomy)
DATA SOURCES
DOCUMENT
(XML, Word, ODF)
N4L
Scribe
Annotation Options
Embed semantic links
Summarize Content
Add Comments
'SCRIBE'd' DOCUMENT
(XML, Word, ODF)
Comments
Comment [N4L2]: 'Sinorhizobium current name for this taxon. The current name is Ensifer http://dx.doi.org/10.1601/nm.1328 Consider revising the document to re current name.
Comment [N4L3]: External link: http://www.atcc.org/Products/All/
Annotation Summary
Annotation Summary
Count
References
4
J.M. Young: The genus name Ensifer Casida 1982 takes priority over Sinorhizobium Chen et al. 1988, and Sinorhizobium morelense Wang et al. 2002 is a later synonym of Ensifer adhaerens Casida 1982. Is the combination Sinorhizobium adhaerens (Casida 1982) Willems et al. 2003 legitimate? Request for an Opinion. International Journal of Systematic and Evolutionary Microbiology 2003, 53: 2107-2110.
Annotation
Ensifer medicae ATCC700748 (GenbankID ATTR00000000
ted species E. meliloti [4, 5].
Taxonomy
Strain Resolution
Sequence Data
Sinorhizobium medicae WSM244, whole genome shotgun sequencing project
Ensifer medicae (Rome et al.) Young (ATCC® 700748™)

Letter

Nature **462**, 1056-1060 (24 December 2009) | doi:10.1038/nature08656; Received 3 June 2009; Accepted 30 October 2009

A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea

Dongying Wu[1,2], Philip Hugenholtz[1], Konstantinos Mavromatis[1], Rüdiger Pukall[3], Eileen Dalin[1], Natalia N. Ivanova[1], Victor Kunin[1], Lynne Goodwin[4], Martin Wu[5], Brian J. Tindall[3], Sean D. Hooper[1], Amrita Pati[1], Athanasios Lykidis[1], Stefan Spring[3], Iain J. Anderson[1], Patrik D'haeseleer[1,6], Adam Zemla[6], Mitchell Singer[2], Alla Lapidus[1], Matt Nolan[1], Alex Copeland[1], Cliff Han[4], Feng Chen[1], Jan-Fang Cheng[1], Susan Lucas[1], Cheryl Kerfeld[1], Elke Lang[3], Sabine Gronow[3], Patrick Chain[1,4], David Bruce[4], Edward M. Rubin[1], Nikos C. Kyrpides[1], Hans-Peter Klenk[3] & Jonathan A. Eisen[1,2]

Letter

Nature **462**, 1056-1060 (24 December 2009) | doi:10.1038/nature08656; Received 3 June 2009; Accepted 30 October 2009

A phylogeny-driven genomic encyclopaedia of Bacteria

Standards in Genomic Sciences (2014) **9:**1278-1296 DOI:10.4056/sigs.5068949

Genomic Encyclopedia of Type Strains, Phase I: The one thousand microbial genomes (KMG-I) project

Nikos C. Kyrpides[1], Tanja Woyke[1], Jonathan A. Eisen[2], George Garrity[3,4], Timothy G. Lilburn[5], Brian J. Beck[5], William B. Whitman[6], Phil Hugenholtz[7], and Hans-Peter Klenk[8]

[1] DOE Joint Genome Institute, Walnut Creek, CA
[2] University of California, Davis, CA
[3] Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, USA
[4] NamesforLife, LLC, East Lansing, MI, USA
[5] American Type Culture Collection, Manassas, VA
[6] Department of Microbiology, University of Georgia, Athens, GA
[7] Australian Centre for Ecogenomics, The University of Queensland, Brisbane QLD 4072, Australia
[8] Leibniz Institute DSMZ — German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

Letter

Nature **462**, 1056-1060 (24 December 2009) | doi:10.1038/nature08656; Received 3 June 2009;
Accepted 30 October 2009

A phylogeny-driven genomic encyclopaedia of Bacteria

Standards in Genomic Sciences (2014) **9:**1278-1296 DOI:10.4056/sigs.5068949

Genomic Encyclopedia of Type Strains, Phase I: The one thousand microbial genomes (KMG-I) project

OPEN ACCESS Freely available online PLOS | BIOLOGY

Community Page

Genomic Encyclopedia of Bacteria and Archaea: Sequencing a Myriad of Type Strains

Nikos C. Kyrpides[1,2]*, Philip Hugenholtz[3], Jonathan A. Eisen[4], Tanja Woyke[1], Markus Göker[5], Charles T. Parker[6], Rudolf Amann[7], Brian J. Beck[8], Patrick S. G. Chain[9], Jongsik Chun[10], Rita R. Colwell[11,12], Antoine Danchin[13], Peter Dawyndt[14], Tom Dedeurwaerdere[15], Edward F. DeLong[16], John C. Detter[9], Paul De Vos[14,17], Timothy J. Donohue[18], Xiu-Zhu Dong[19], Dusko S. Ehrlich[20], Claire Fraser[21], Richard Gibbs[22], Jack Gilbert[23], Paul Gilna[24], Frank Oliver Glöckner[7,25], Janet K. Jansson[26], Jay D. Keasling[26,27], Rob Knight[28], David Labeda[29], Alla Lapidus[30,31], Jung-Sook Lee[32], Wen-Jun Li[33], Juncai MA[34], Victor Markowitz[1,26], Edward R. B. Moore[35], Mark Morrison[36], Folker Meyer[37], Karen E. Nelson[38], Moriya Ohkuma[39], Christos A. Ouzounis[40,41], Norman Pace[42], Julian Parkhill[43], Nan Qin[44], Ramon Rossello-Mora[45], Johannes Sikorski[5], David Smith[46], Mitch Sogin[47], Rick Stevens[37], Uli Stingl[48], Ken-ichiro Suzuki[49], Dorothea Taylor[6], Jim M. Tiedje[50], Brian Tindall[5], Michael Wagner[51], George Weinstock[52], Jean Weissenbach[53], Owen White[21], Jun Wang[44,54], Lixin Zhang[19,55], Yu-Guang Zhou[34], Dawn Field[56], William B. Whitman[57], George M. Garrity[6,50], Hans-Peter Klenk[5]*

Letter

Nature **462**, 1056-1060 (24 December 2009) | doi:10.1038/nature08656; Received 3 June 2009;
Accepted 30 October 2009

A phylogeny-driven genomic encyclopaedia of Bacteria

Standards in Genomic Sciences (2014) **9**:1278-1296 DOI:10.4056/sigs.5068949

Genomic Encyclopedia of Type Strains, Phase I: The one thousand microbial genomes (KMG-I) project

OPEN ACCESS Freely available online

PLOS | BIOLOGY

Community Page

Genomic Encyclopedia of Bacteria and Archaea: Se

Nik
Ch
Rit
Joh
Cla
Jar
Jur
Ma
No
Mi
Bri
Jur
Ge

Whitman *et al. Standards in Genomic Sciences* (2015) 10:26
DOI 10.1186/s40793-015-0017-x

Standards in Genomic Sciences

COMMENTARY Open Access

Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase III: the genomes of soil and plant-associated and newly described type strains

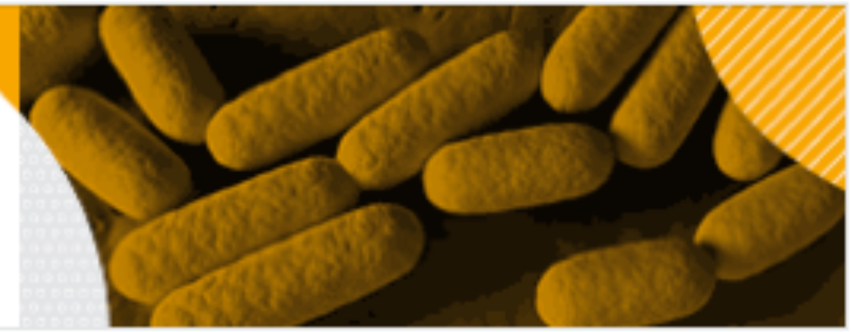
William B Whitman[1*], Tanja Woyke[2], Hans-Peter Klenk[3], Yuguang Zhou[4], Timothy G Lilburn[5,11], Brian J Beck[5,10], Paul De Vos[6], Peter Vandamme[6], Jonathan A Eisen[7], George Garrity[8], Philip Hugenholtz[9] and Nikos C Kyrpides[2]

Validly published named prokaryotes used in the analyses

Rank	Validly Published	Coverage (16S)[a]	Genomes	Outliers	Singleton	2 to 3	4 to 9	≥10
species	12,981	11,800	4092[d]	274	11,800	0	0	0
genus	2,716	2422	1333	131	1,166	616	437	203
family	389	451[b]	343	86	121	71	79	180
order	174	202[b]	161	47	52	28	34	90
class	83	85[b]	69	19	23	13	14	35
phylum	0	34[c]	33	9	4	4	6	20
domain	0	2	2	2	0	0	0	2

Source - NamesforLife, LLC, based on validly published named prokaryotic species, subspecies and higher taxa as of August 5, 2015; adjusted for synonyms, basonyms, orthographic corrections and rejected names arising from Judicial Opinions.

INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY
Official publication of the ICSP and the BAM Division of the IUMS



About | Editorial Board | Accepted Papers | Current Issue | Archive | Collections | Submit a Paper

F International Code of Nomenclature of Prokaryotes

Authors: Charles T Parker[1], Brian J Tindall[2], George M Garrity[3]

VIEW AFFILIATIONS

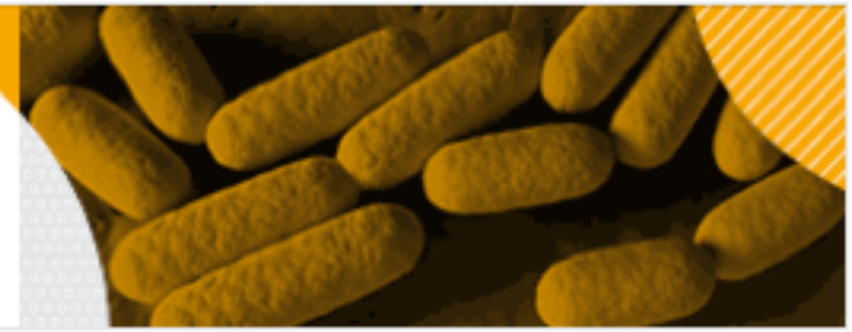
PDF
4.18 MB

Published Ahead of Print: 20 November, 2015 International Journal of Systematic and Evolutionary Microbiology doi: 10.1099/ijsem.0.000778

Published Online: 20/11/2015

INTERNATIONAL JOURNAL OF **SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY**

Official publication of the ICSP and the BAM Division of the IUMS



About | Editorial Board | Accepted Papers | Current Issue | Archive | Collections | Submit a Paper

F International Code of Nomenclature of Prokaryotes

Authors: Charles T Parker[1], Brian J Tindall[2], George M Garrity[3]

VIEW AFFILIATIONS

PDF 4.18 MB

Published Ahead of Print: 20 November, 2015 International Journal of Systematic and Evolutionary Microbiology doi: 10.1099/ijsem.0.000778

Published Online: 20/11/2015

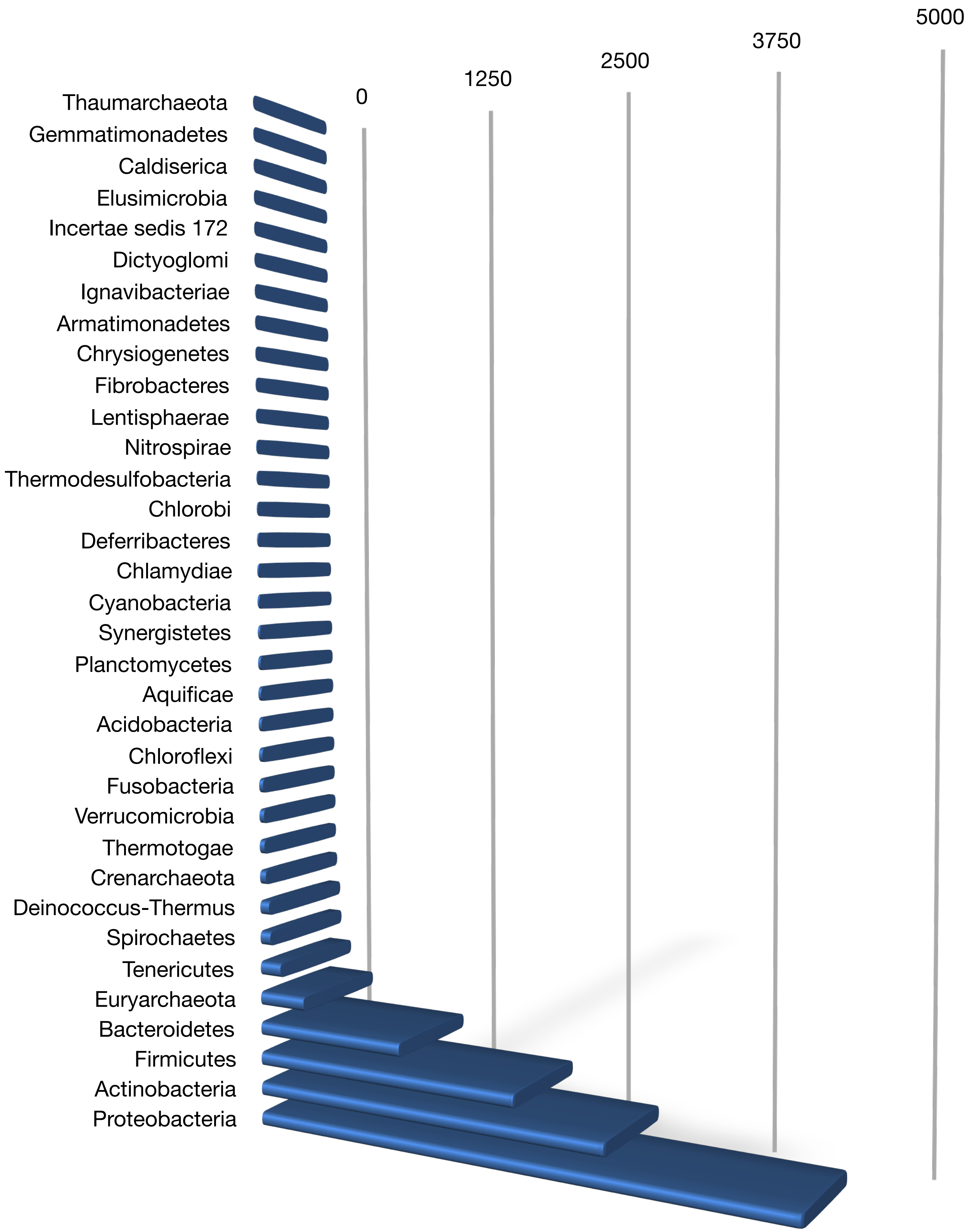
International Journal of Systematic and Evolutionary Microbiology (2015), **65**, 4284–4287 DOI 10.1099/ijsem.0.000664

Taxonomic Note

Proposal to include the rank of phylum in the International Code of Nomenclature of Prokaryotes

Aharon Oren,[1] Milton S. da Costa,[2] George M. Garrity,[3] Fred A. Rainey,[4] Ramon Rosselló-Móra,[5] Bernhard Schink,[6] Iain Sutcliffe,[7] Martha E. Trujillo[8] and William B. Whitman[9]

The cultivated bacterial and archaeal phyla



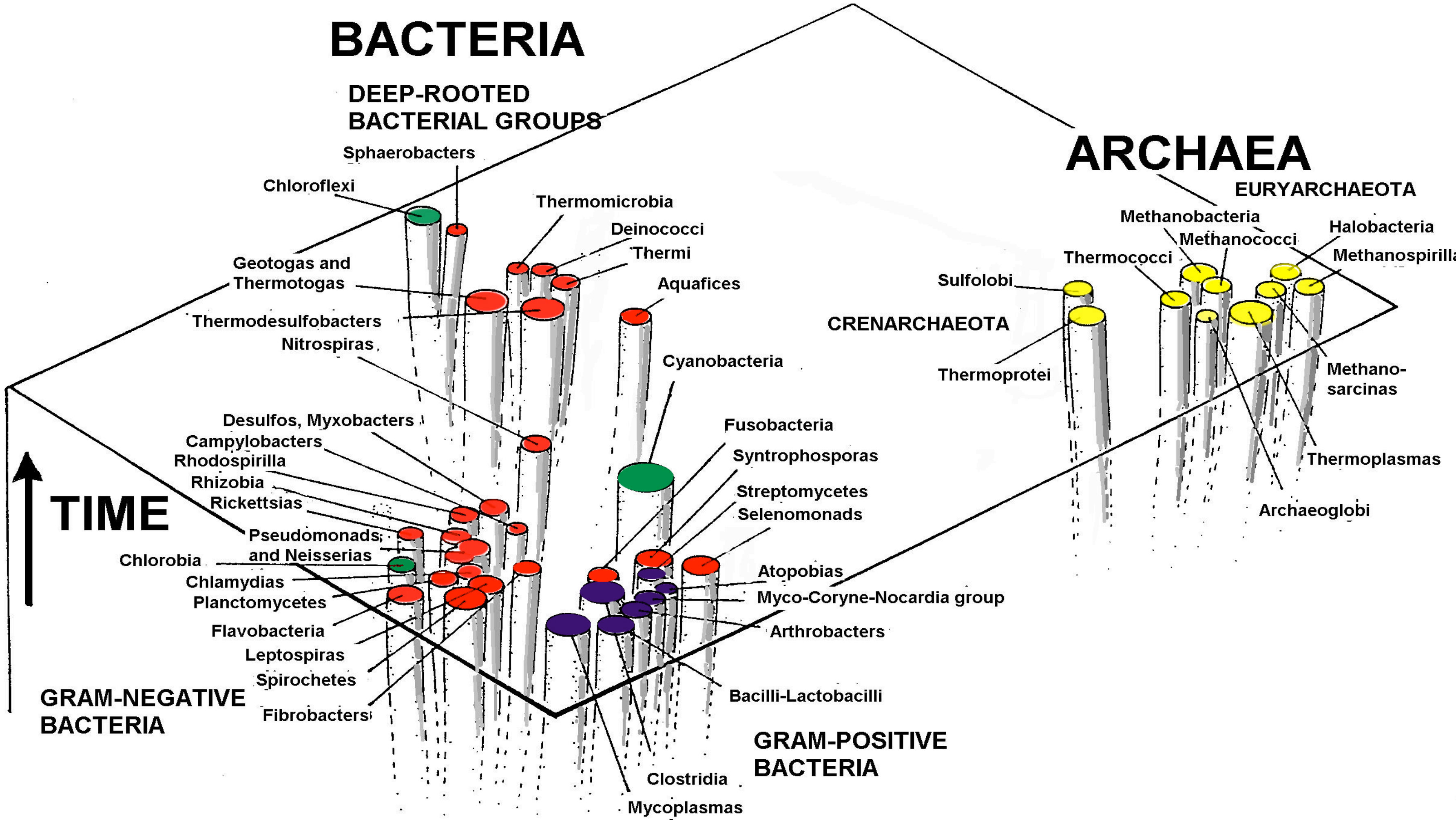
CRUCIAL QUESTIONS

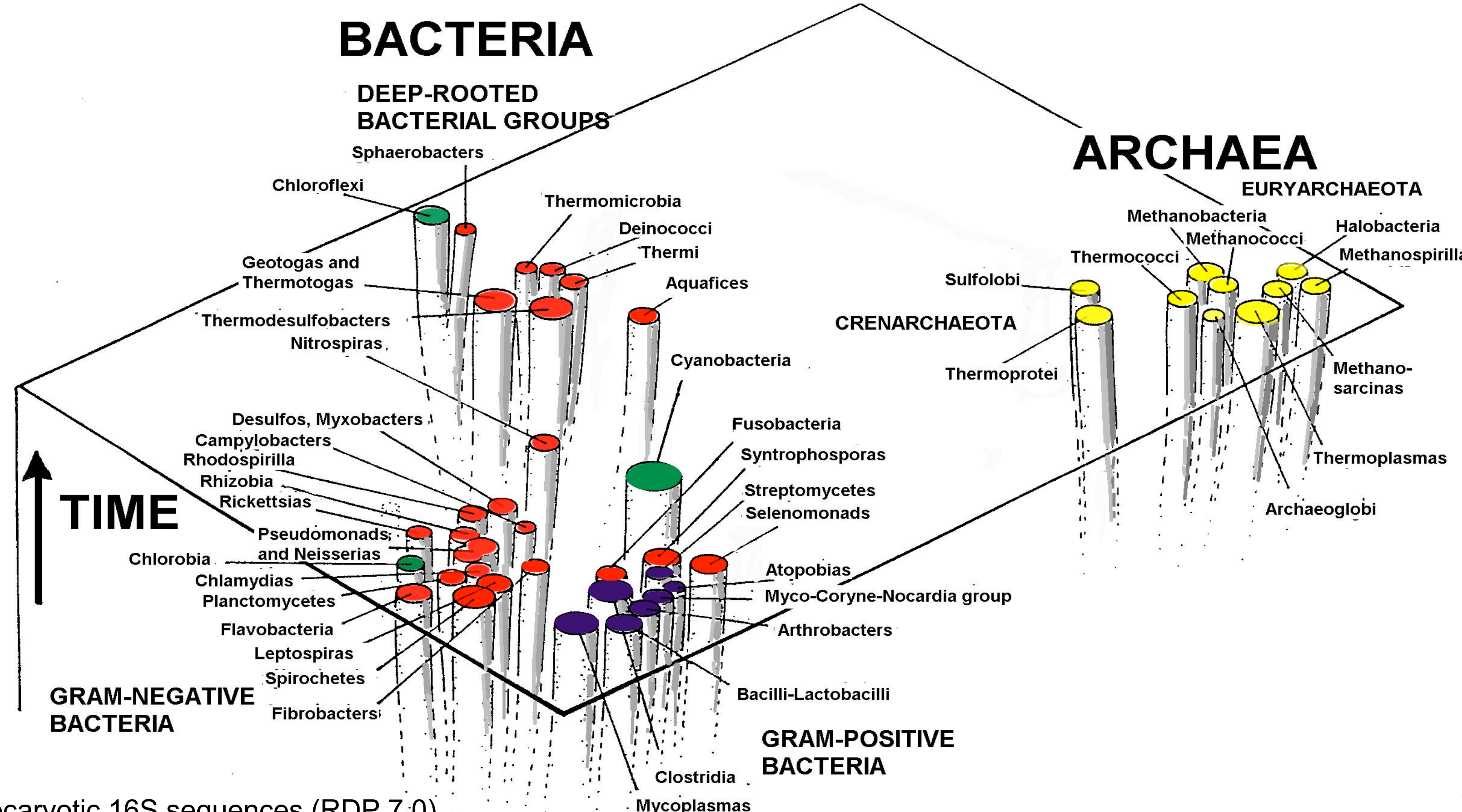
Here are five questions that anyone conducting or evaluating this research should ask to keep from getting carried away by hype.

Can experiments detect differences that matter? Profiling a microbiome could produce a catalogue at the level of phyla, species or genes. Much work relies on analysis of *16S rRNA*, an ancient gene that tolerates little variation and so is reliably found across the bacterial kingdom. But this allows only a coarse sorting. For example, microbiomes associated with obesity have been distinguished by different ratios of bacterial phyla, which encompass a staggering range of diversity. If this criterion were used to characterize animal communities, an aviary of 100 birds and 25 snails would be considered identical to an aquarium with 8 fish and 2 squid, because each has four times as many vertebrates as molluscs. Even within a single species, strains often differ greatly in the genes they contain.

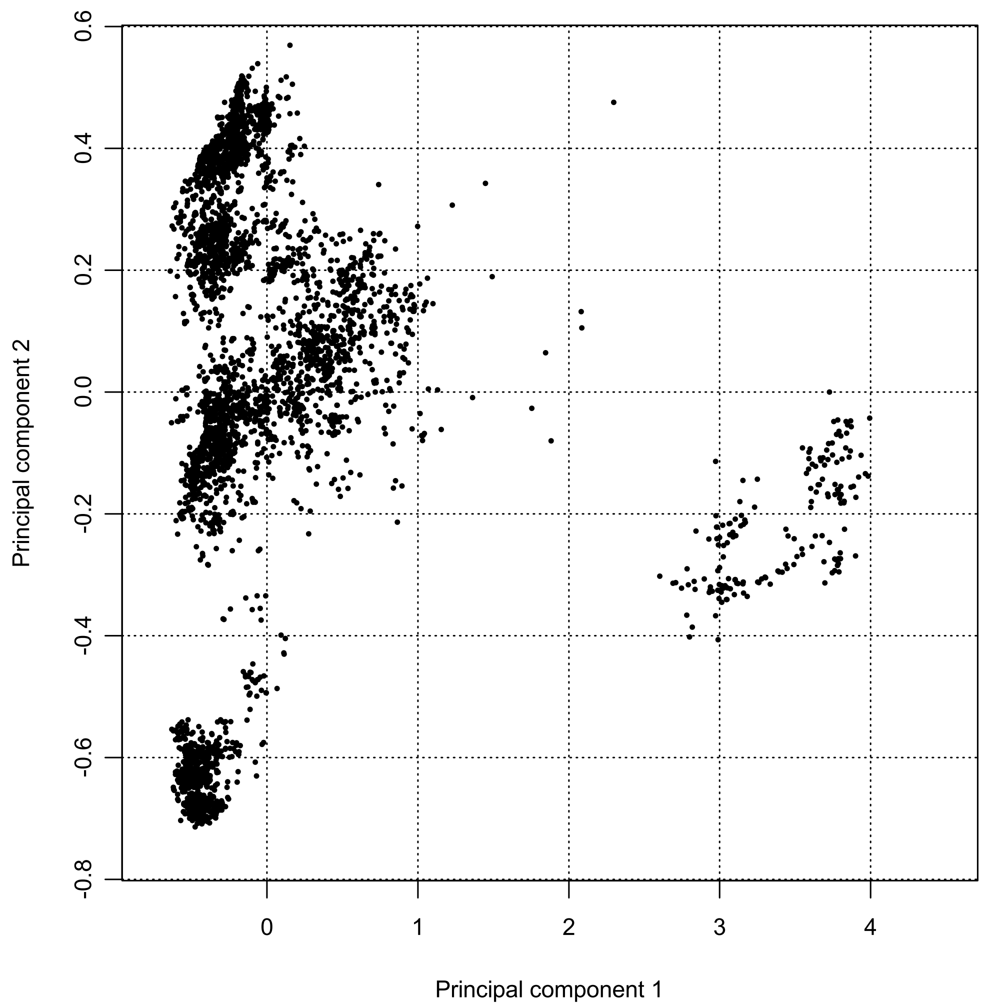
William P. Hanage

Nature 512, 247–248 (21 August 2014)
doi:10.1038/512247a

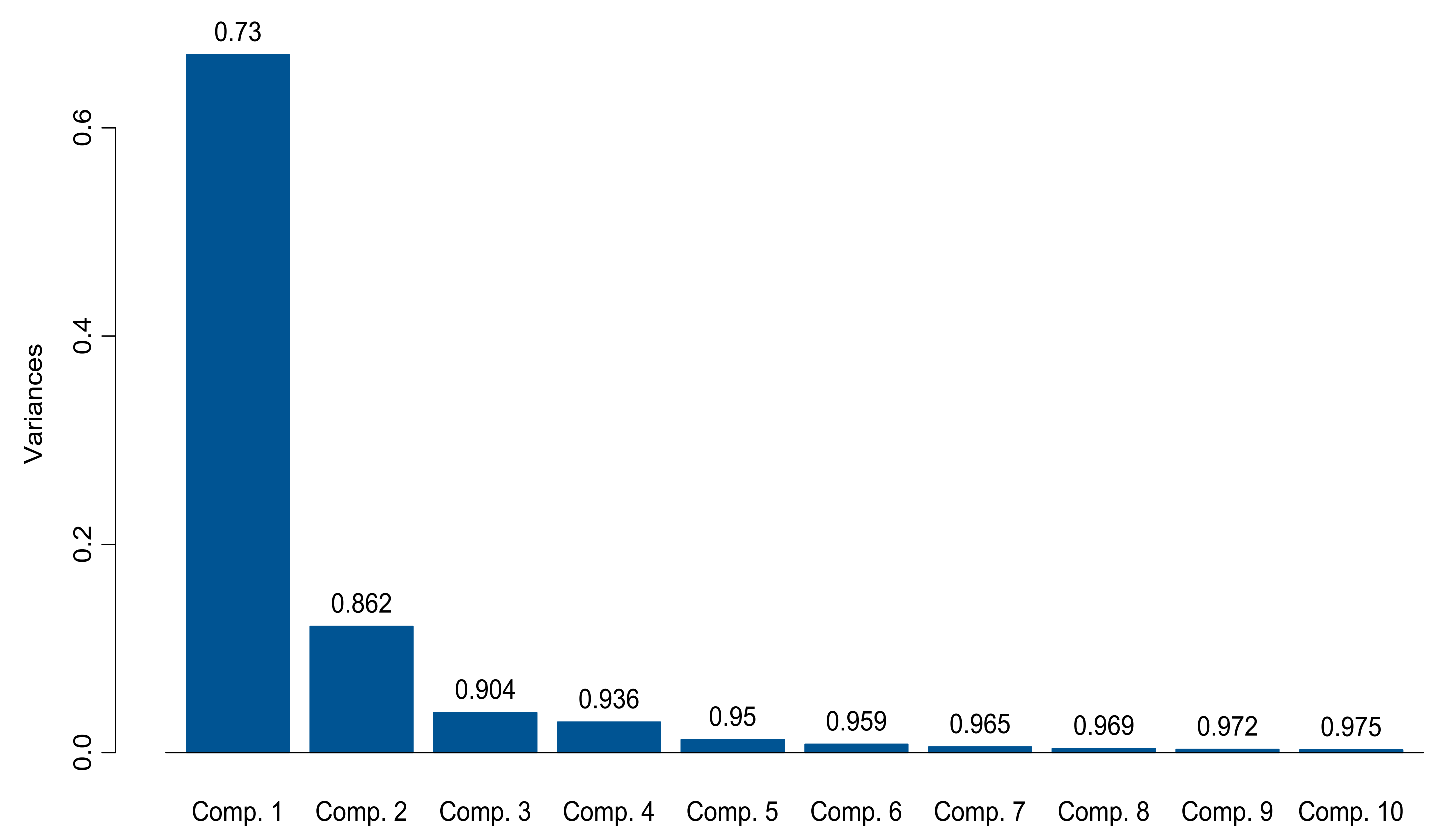

BACTERIA
DEEP-ROOTED BACTERIAL GROUPS
Sphaerobacters
Chloroflexi
Thermomicrobia
Deinococci
Thermi
Geotogas and Thermotogas
Aquafices
Thermodesulfobacters
Nitrospiras
Cyanobacteria
Desulfos, Myxobacters
Fusobacteria
Campylobacters
Rhodospirilla
Syntrophosporas
Rhizobia
Streptomycetes
Rickettsias
Selenomonads
TIME
Pseudomonads, and Neisserias
Chlorobia
Atopobias
Chlamydias
Myco-Coryne-Nocardia group
Planctomycetes
Flavobacteria
Arthrobacters
Leptospiras
Spirochetes
Bacilli-Lactobacilli
GRAM-NEGATIVE BACTERIA
Fibrobacters
GRAM-POSITIVE BACTERIA
Clostridia
Mycoplasmas
ARCHAEA
EURYARCHAEOTA
Methanobacteria
Methanococci
Halobacteria
Thermococci
Methanospirilla
Sulfolobi
CRENARCHAEOTA
Thermoprotei
Methano-sarcinas
Thermoplasmas
Archaeoglobi



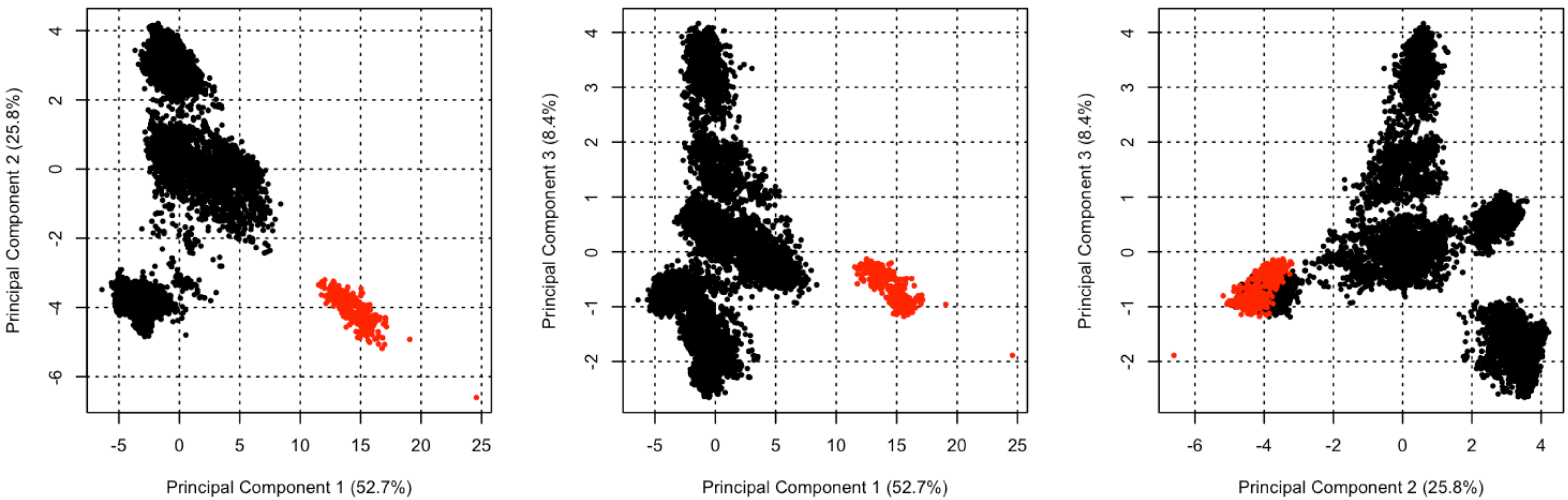
Map of the procaryotic 16S sequences (RDP 7.0)



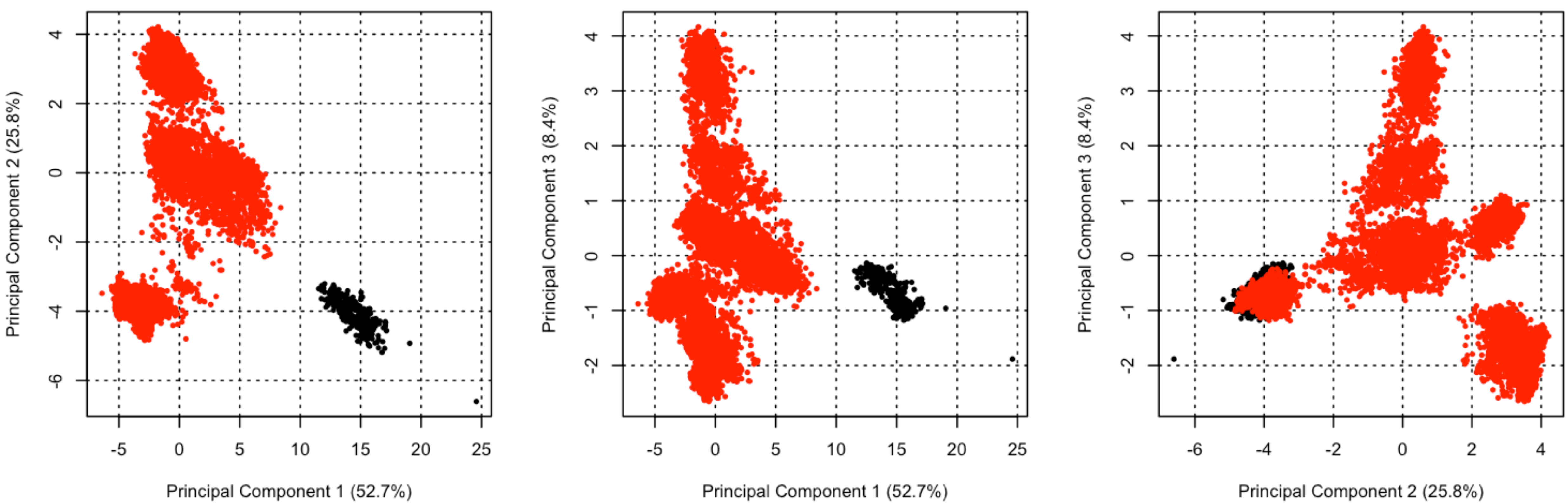
Scree plot for 16S sequences (n=4405)



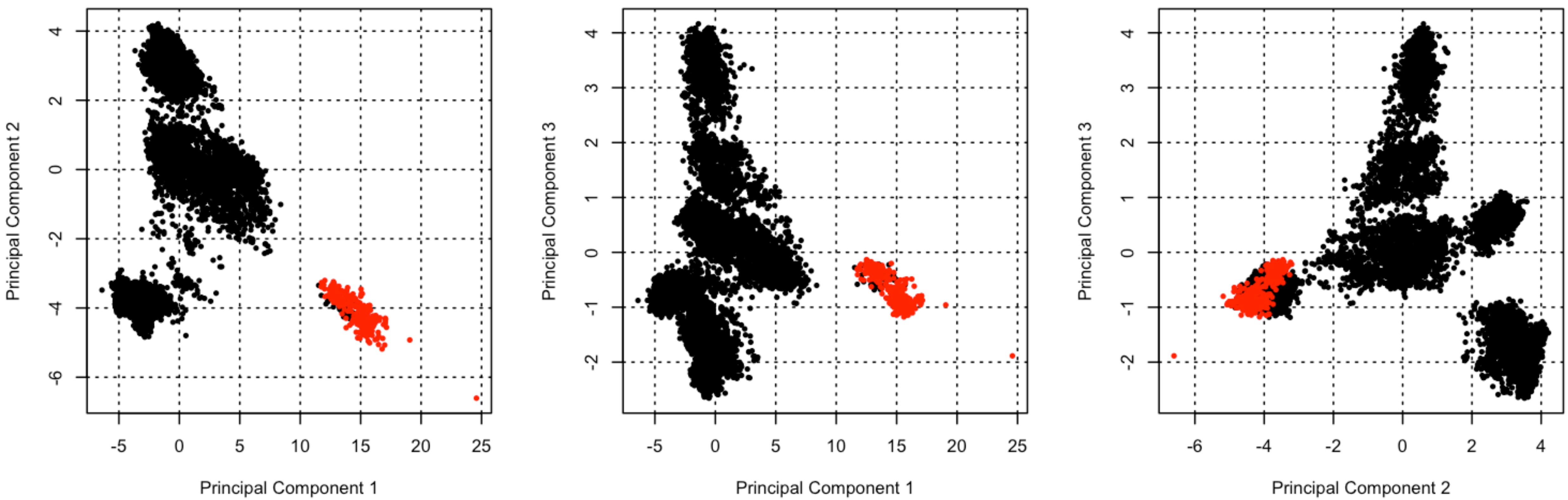
Archaea



Bacteria

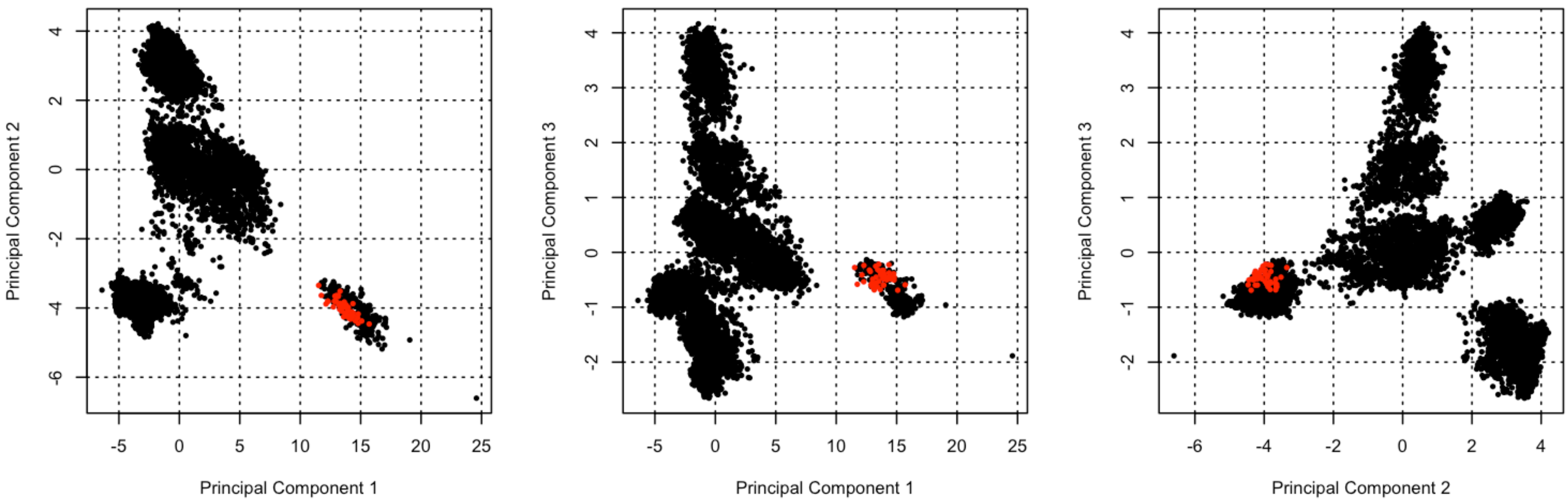


Euryarchaeota



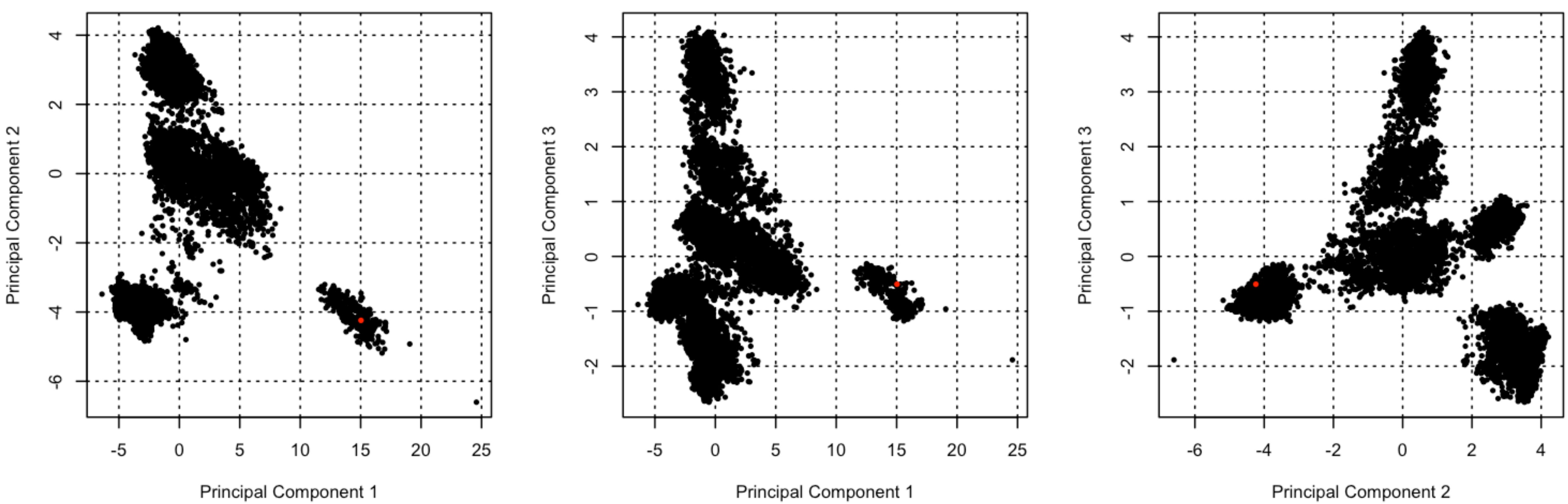
Archaea : Euryarchaeota

Crenarchaeota



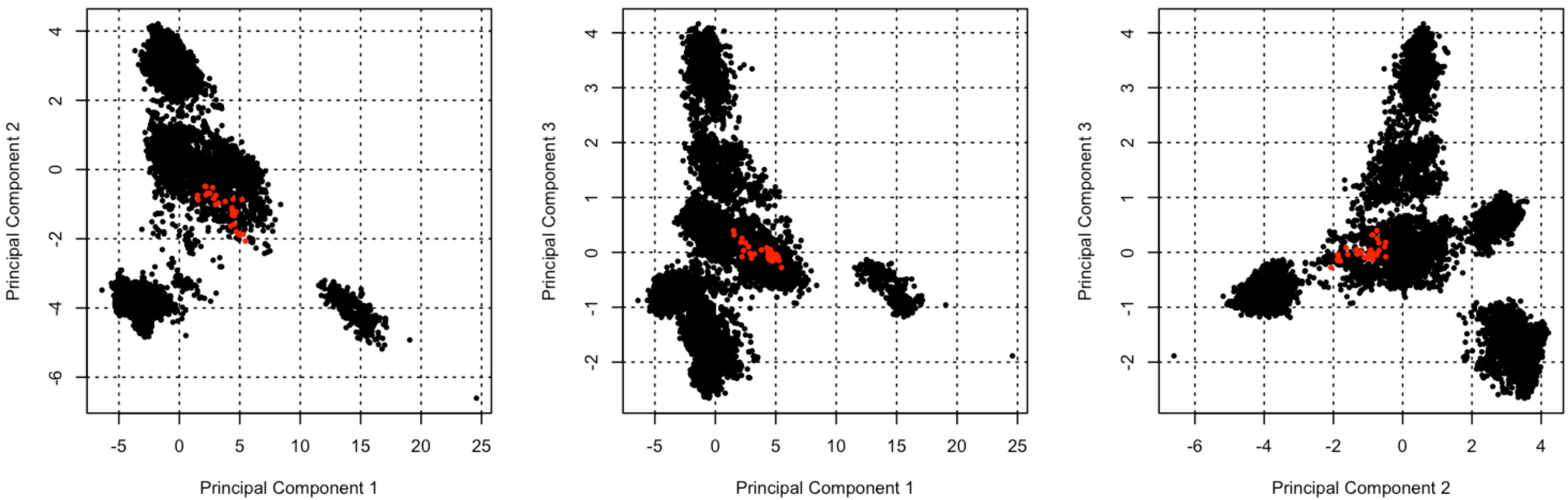
Archaea : Crenarchaeota

Thaumarchaeota



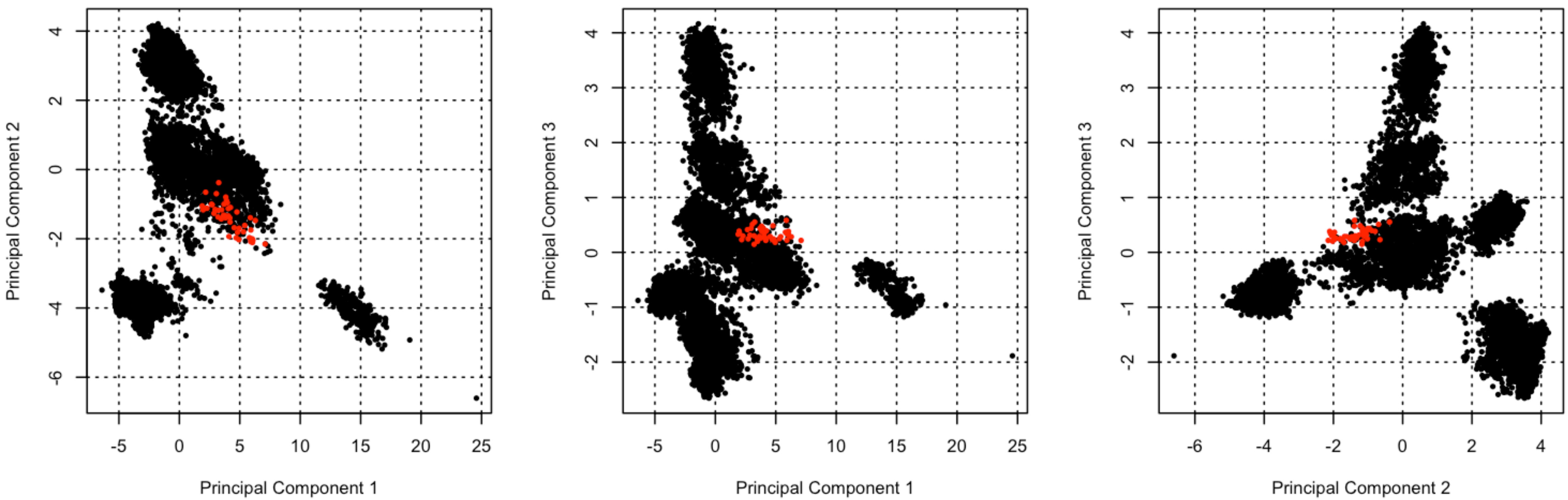
Archaea : Thaumarchaeota

Aquificae



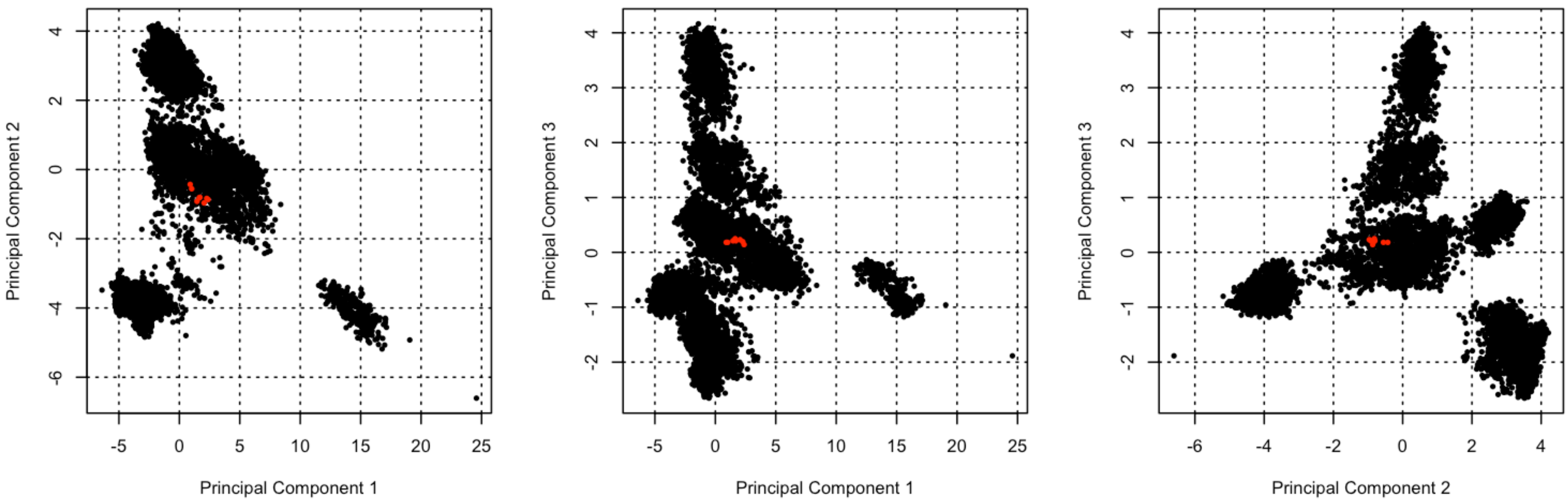
Bacteria : Aquificae

Thermotogae



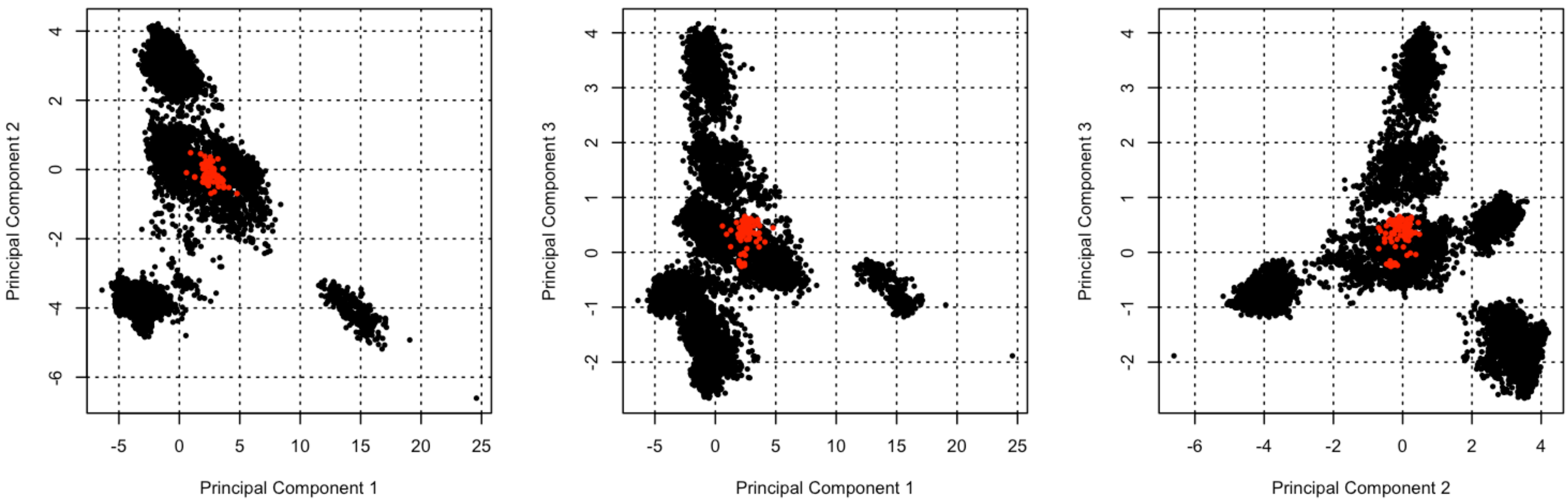
Bacteria : Thermotogae

Thermodesulfobacteria



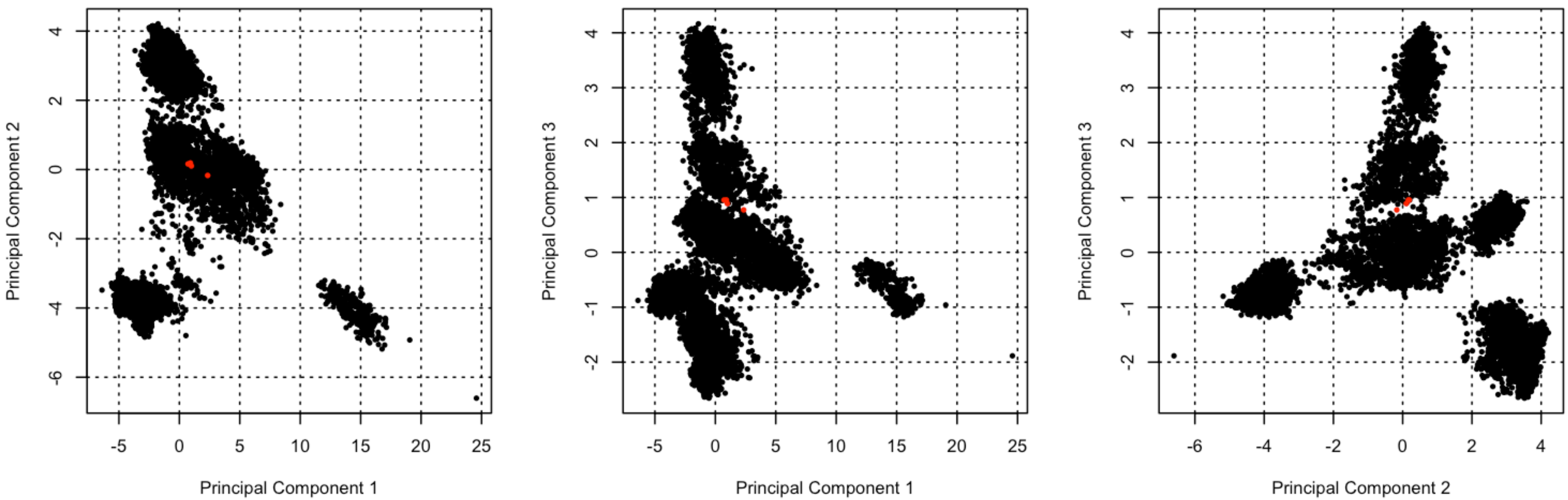
Bacteria : Thermodesulfobacteria

Deinococcus-Thermus



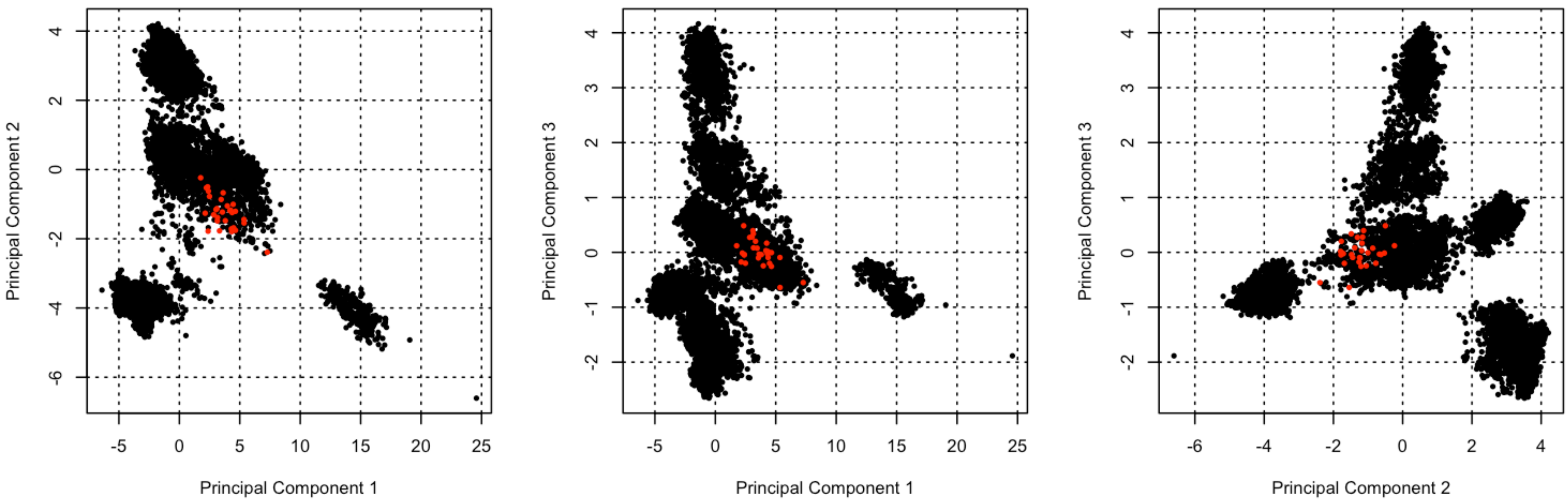
Bacteria : Deinococcus-Thermus

Chrysiogenetes



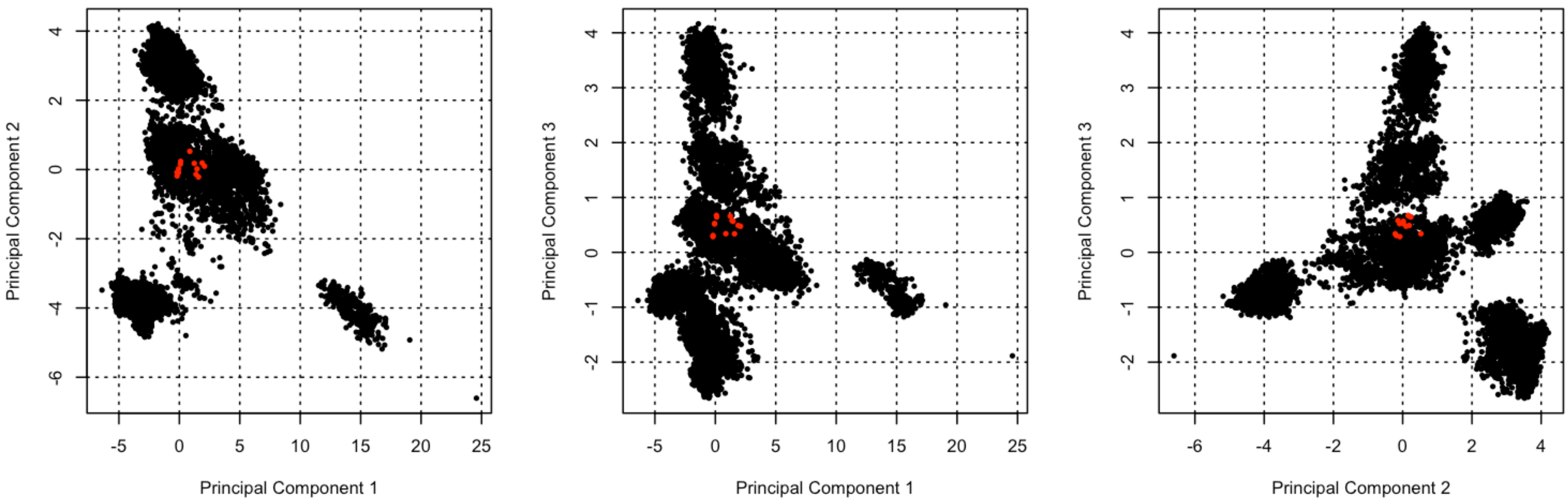
Bacteria : Chrysiogenetes

Chloroflexi



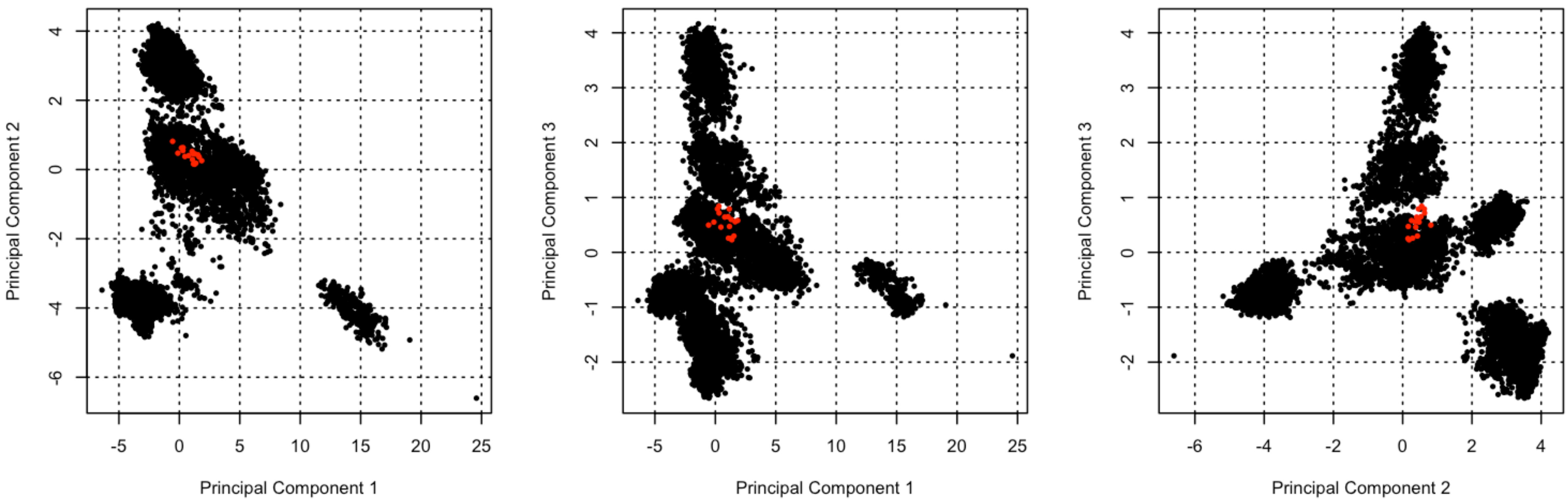
Bacteria : Chloroflexi

Deferribacteres



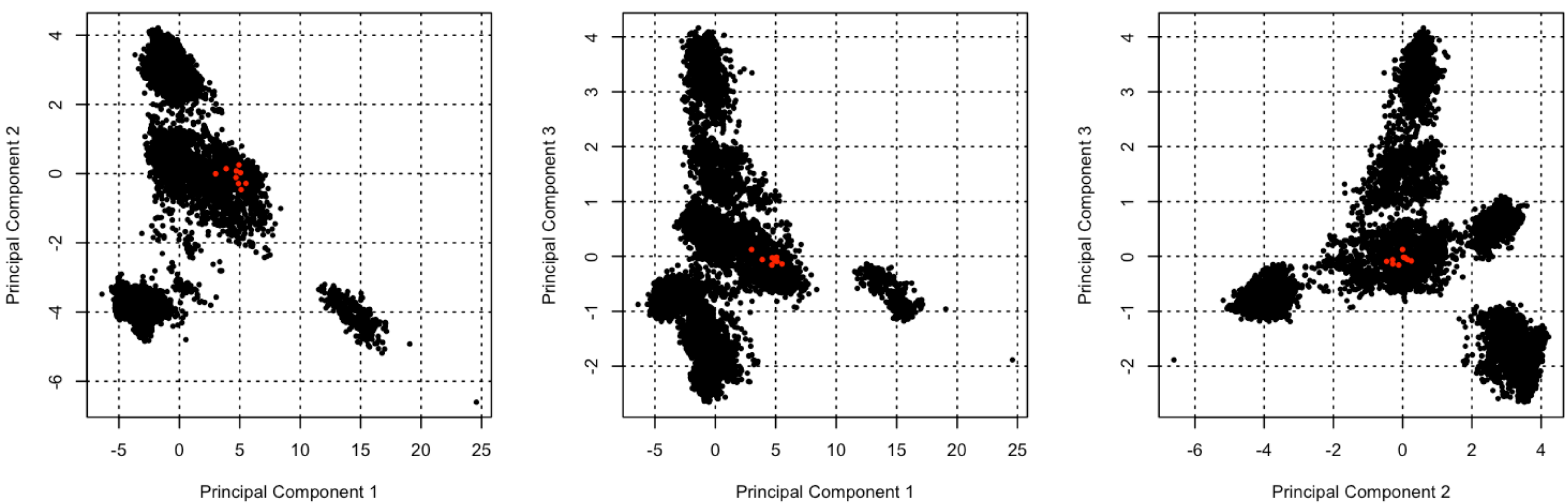
Bacteria : Deferribacteres

Cyanobacteria



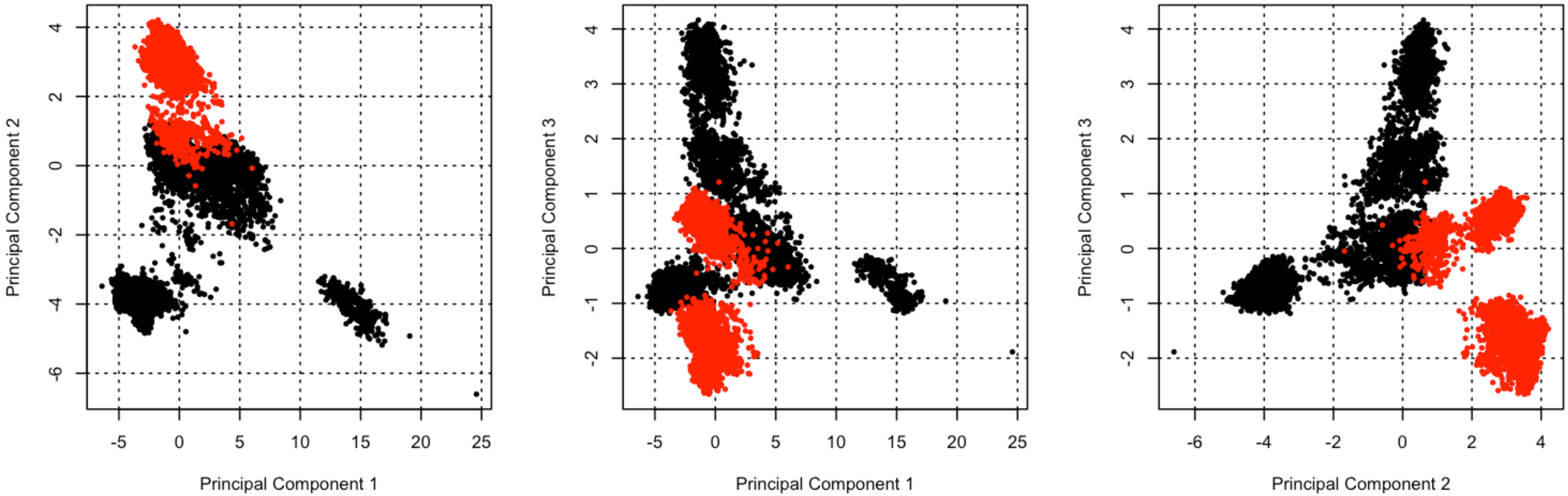
Bacteria : Cyanobacteria

Chlorobi



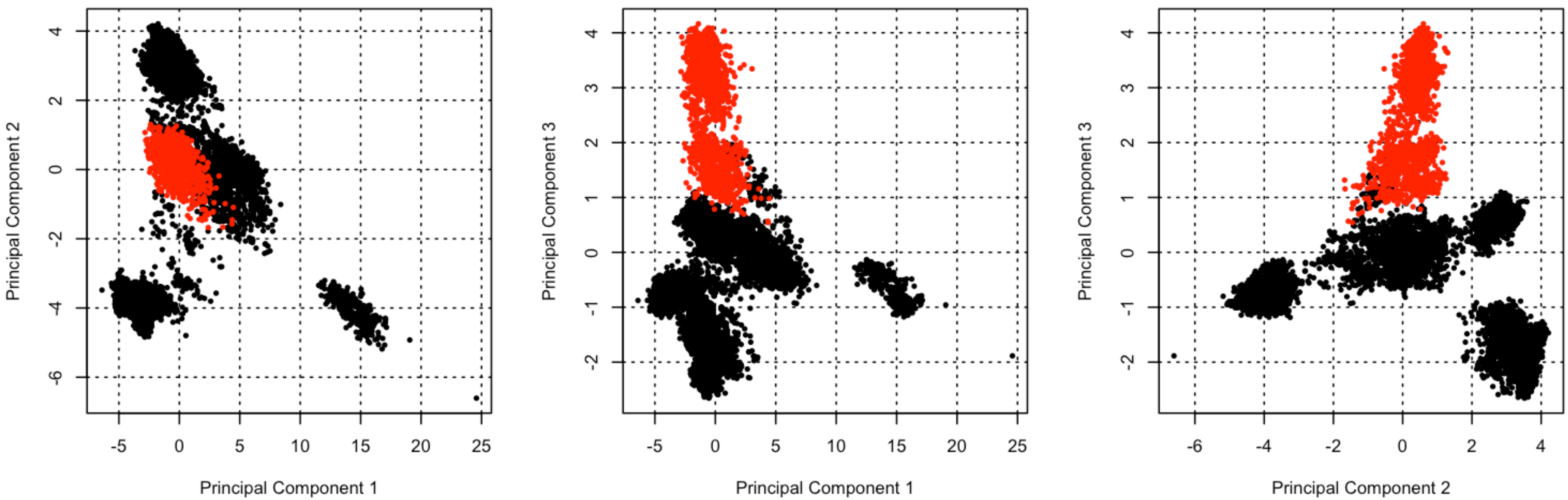
Bacteria : Chlorobi

Proteobacteria



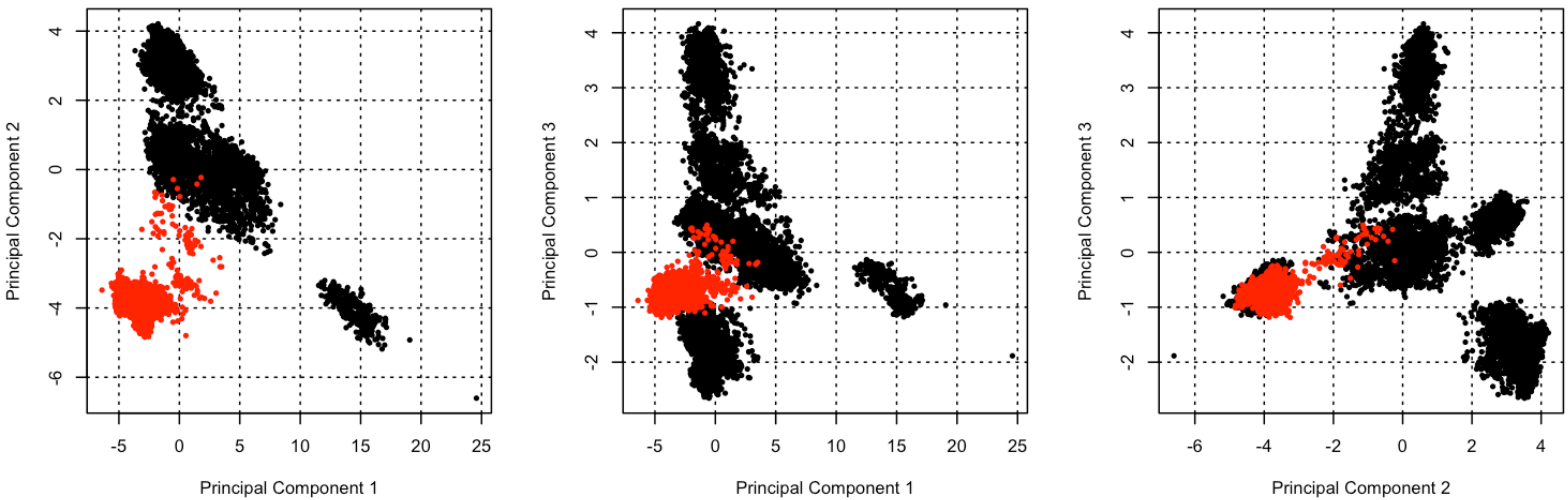
Bacteria : Proteobacteria

Firmicutes



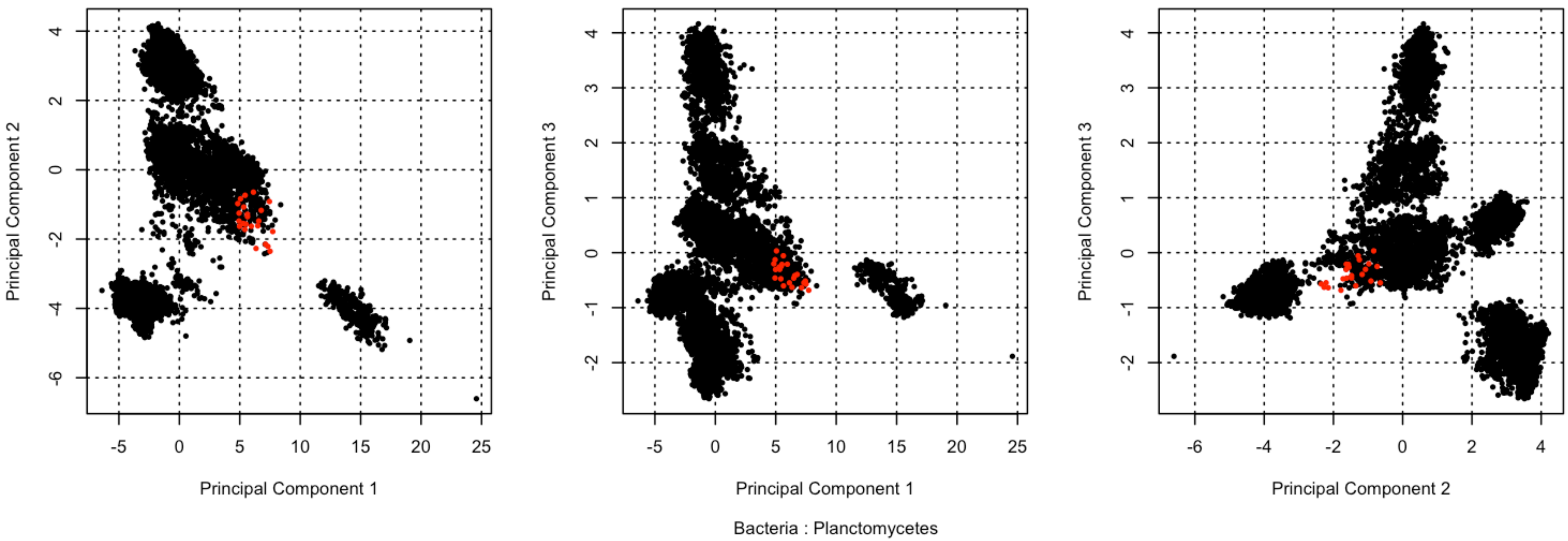
Bacteria : Firmicutes

Actinobacteria

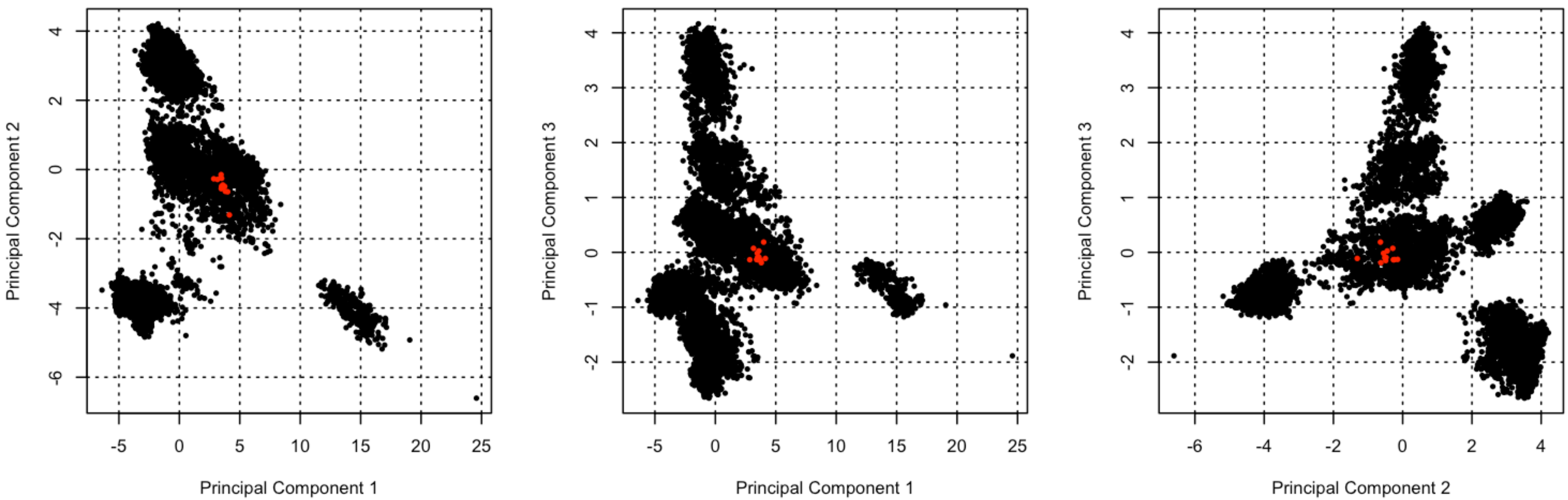


Bacteria : Actinobacteria

Planctomycetes

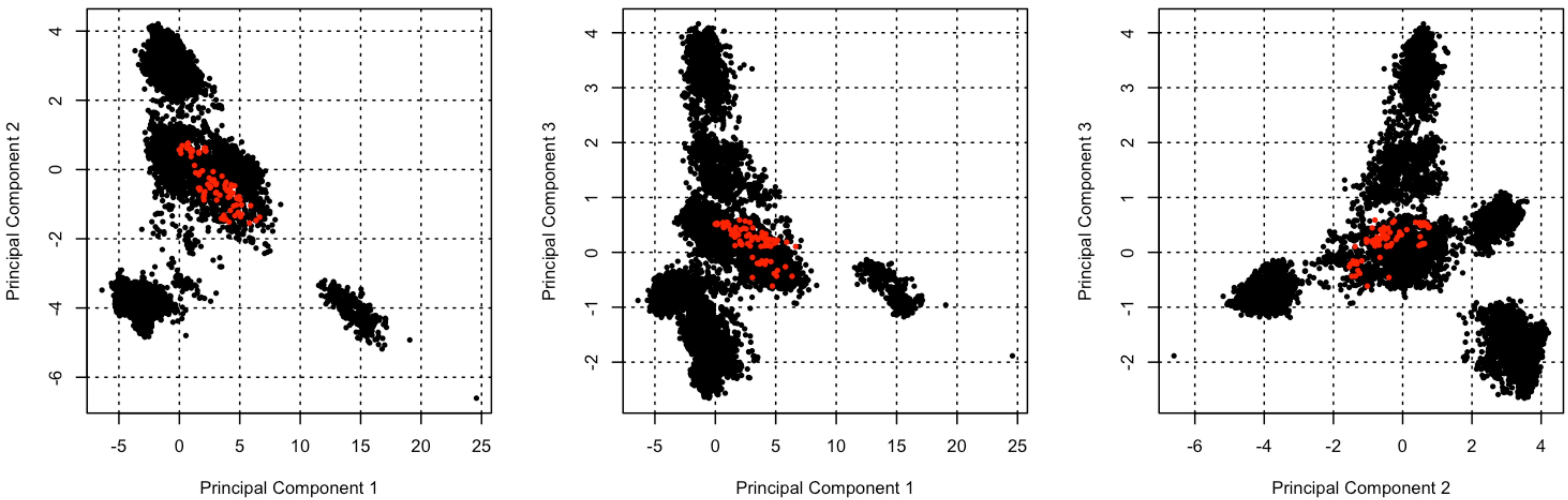


Chlamydiae



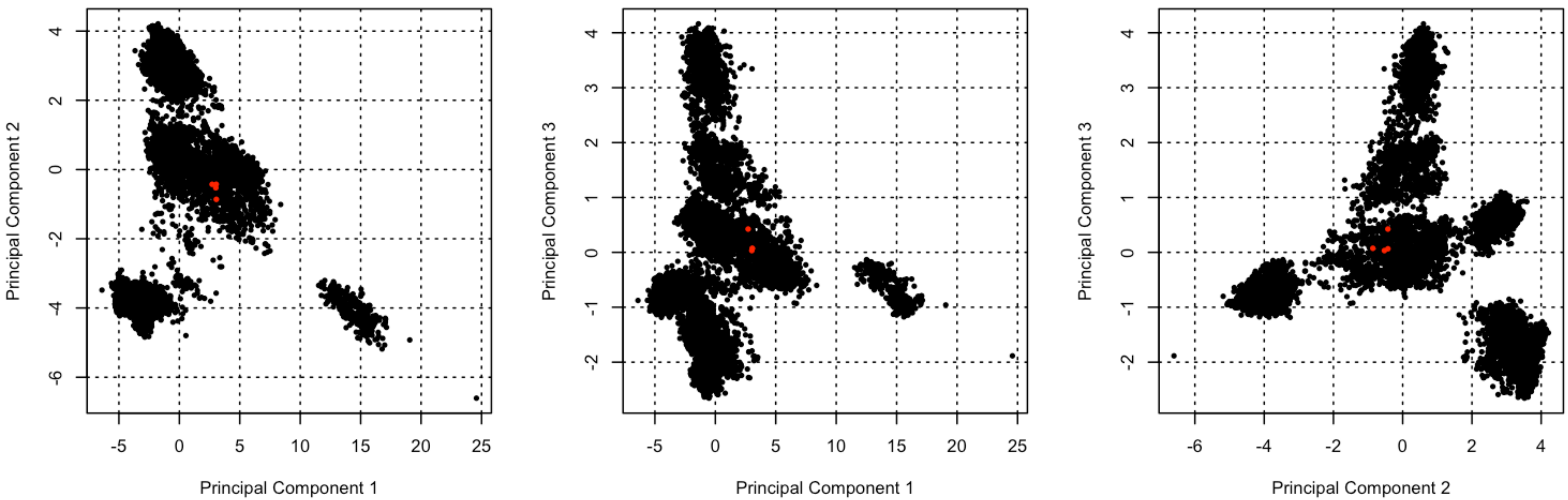
Bacteria : Chlamydiae

Spirochaetes



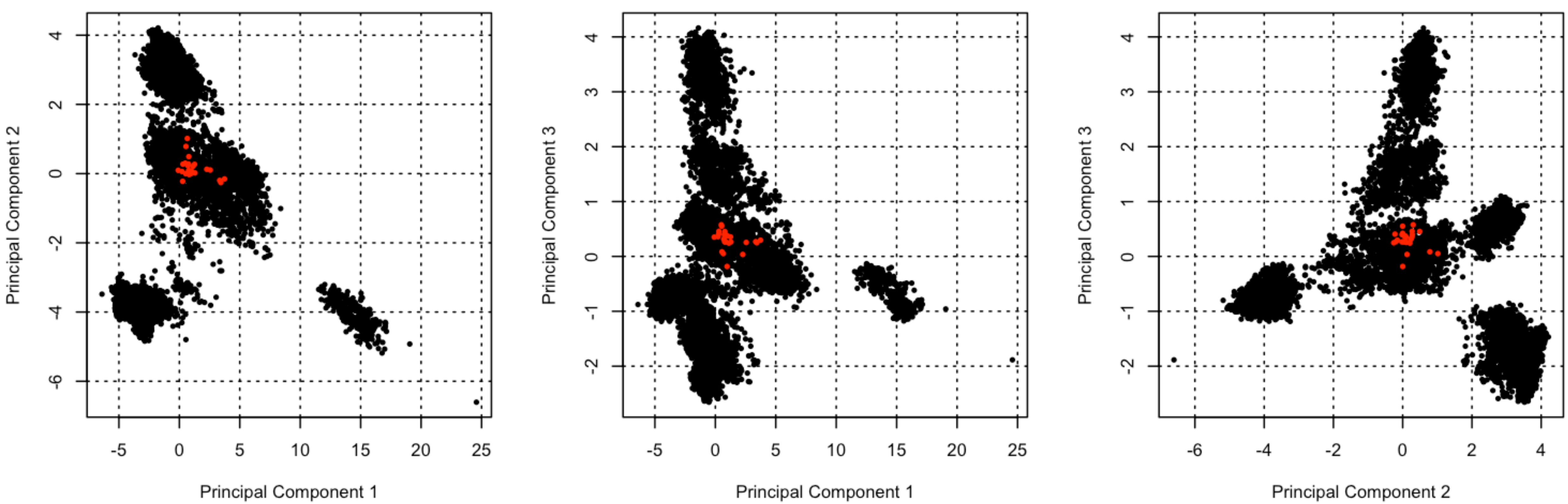
Bacteria : Spirochaetes

Fibrobacteres



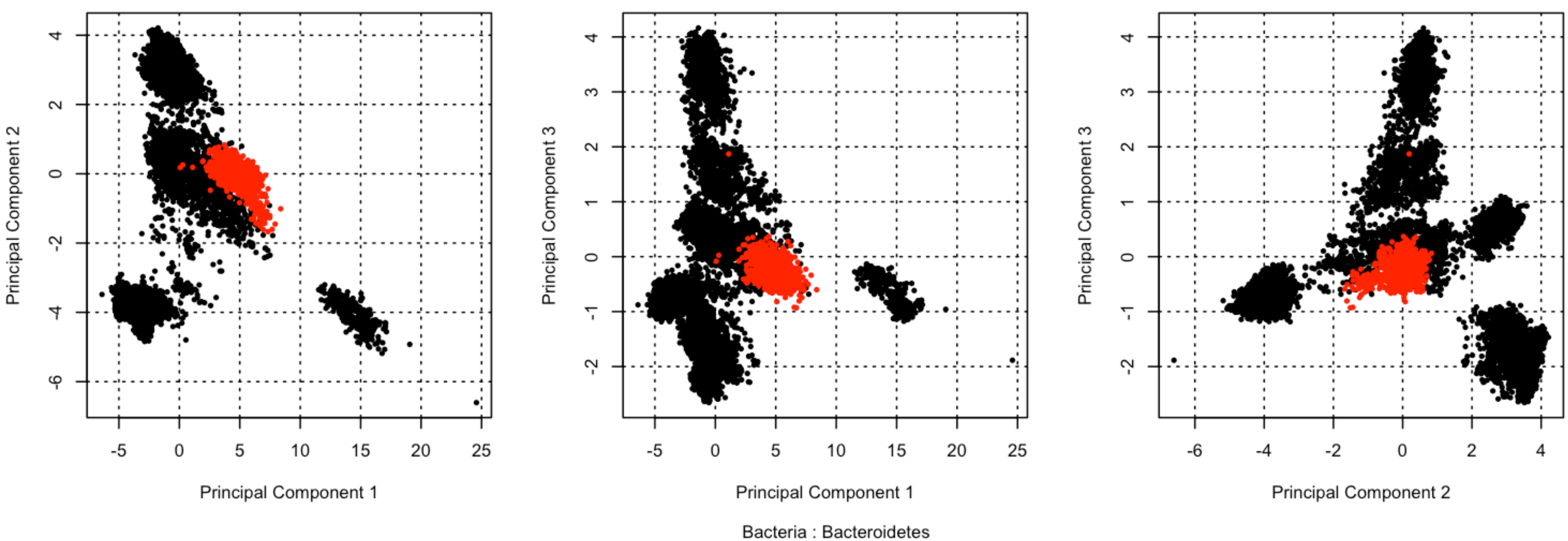
Bacteria : Fibrobacteres

Acidobacteria

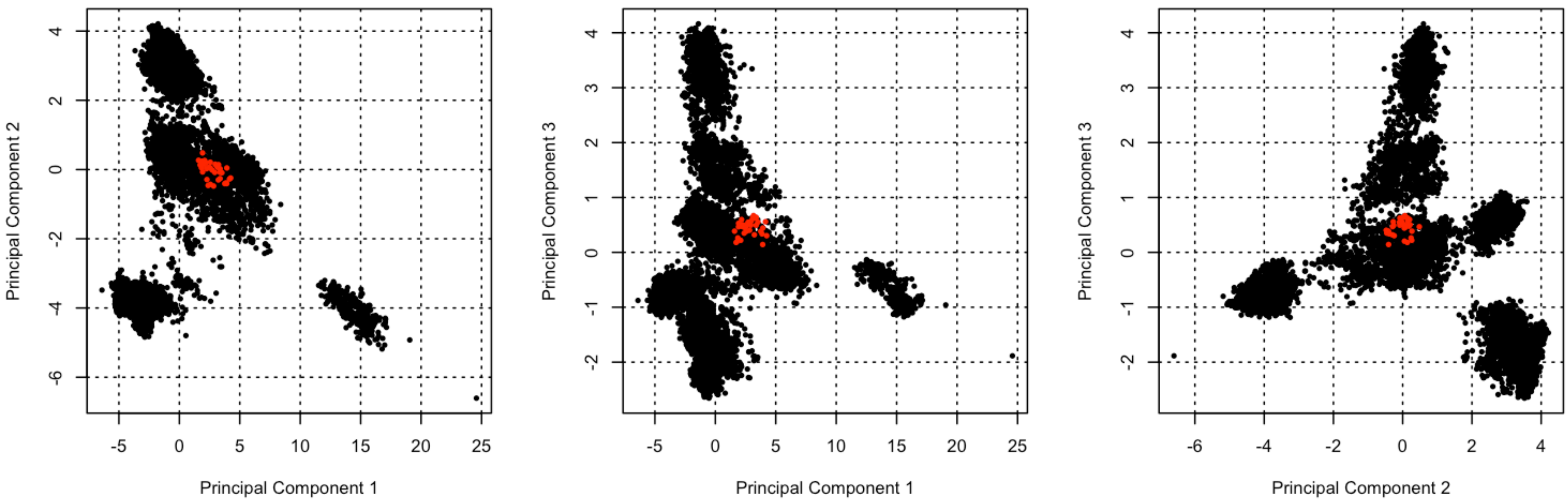


Bacteria : Acidobacteria

Bacteroidetes

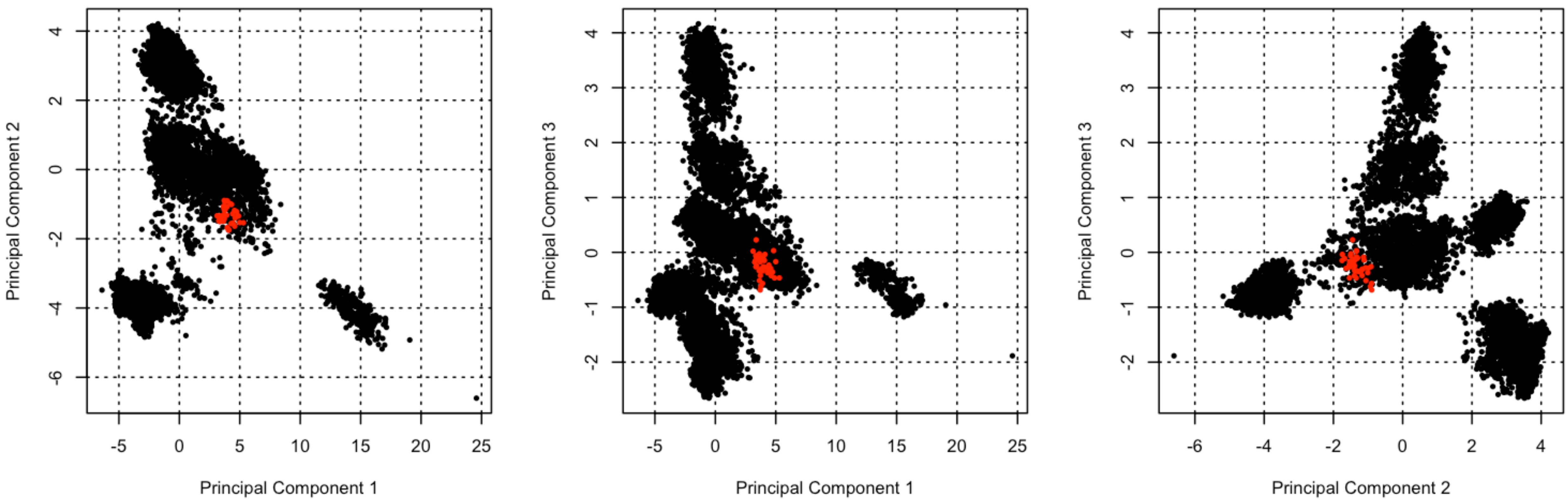


Fusobacteria



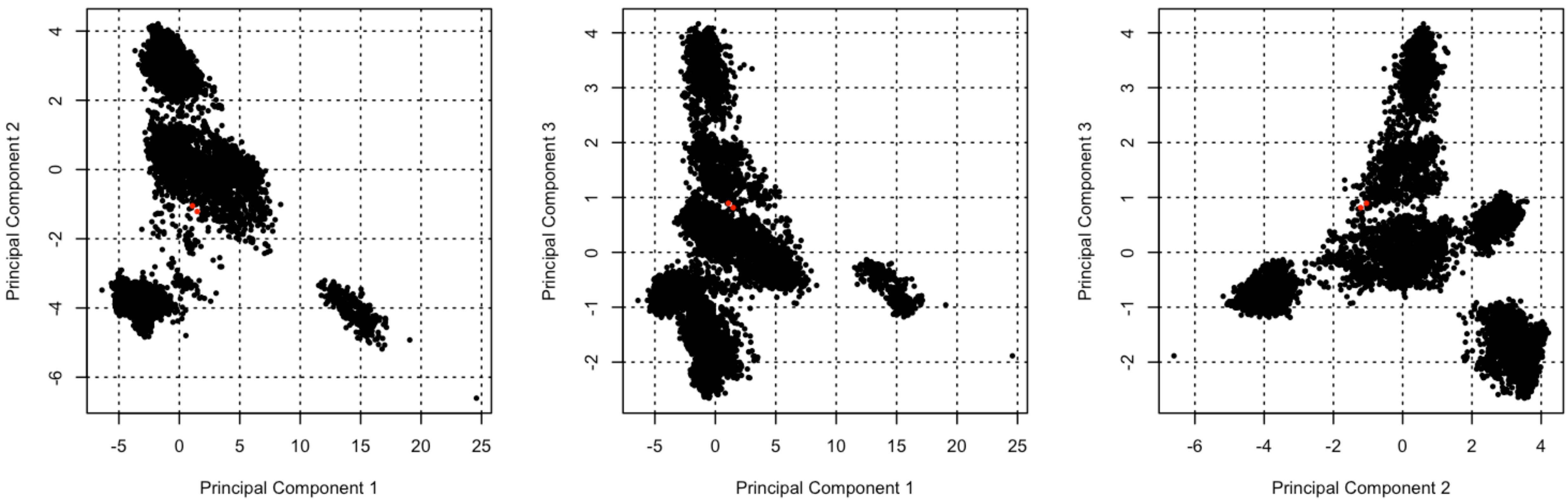
Bacteria : Fusobacteria

Verrucomicrobia



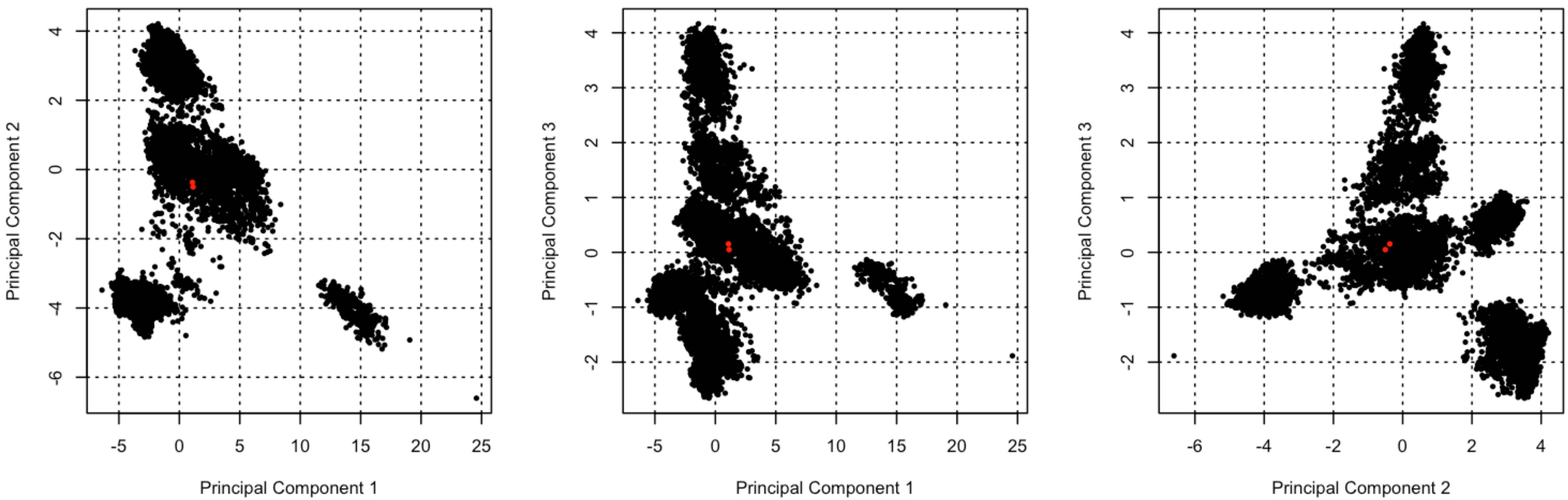
Bacteria : Verrucomicrobia

Dictyoglomi



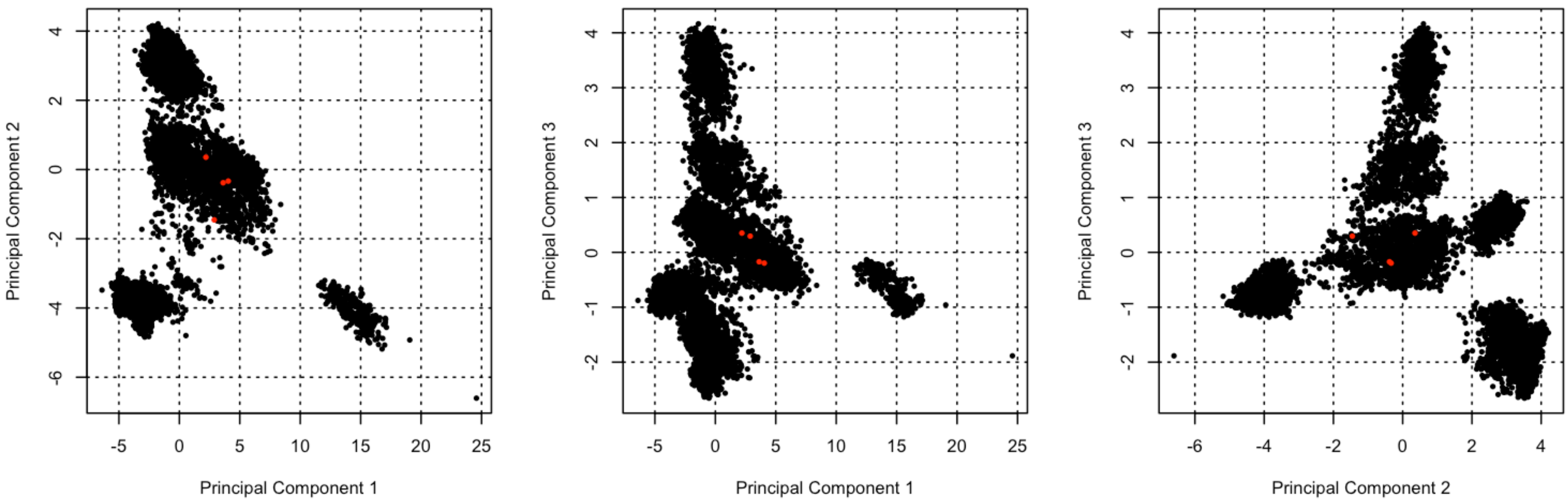
Bacteria : Dictyoglomi

Gemmatimonadetes



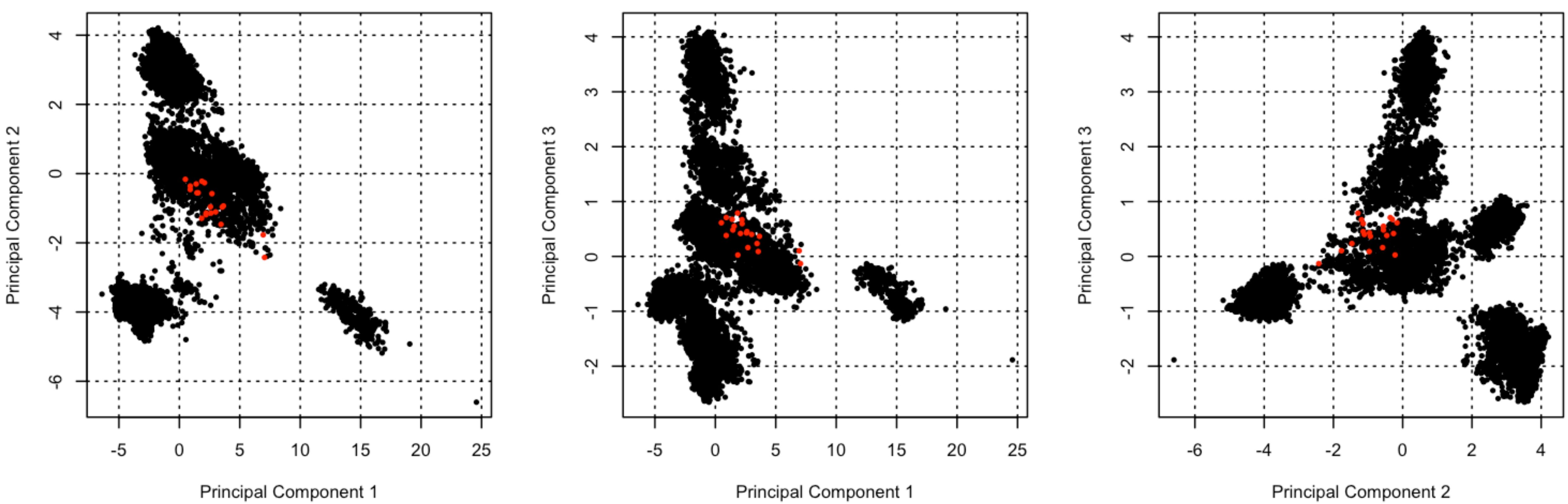
Bacteria : Gemmatimonadetes

Lentisphaerae



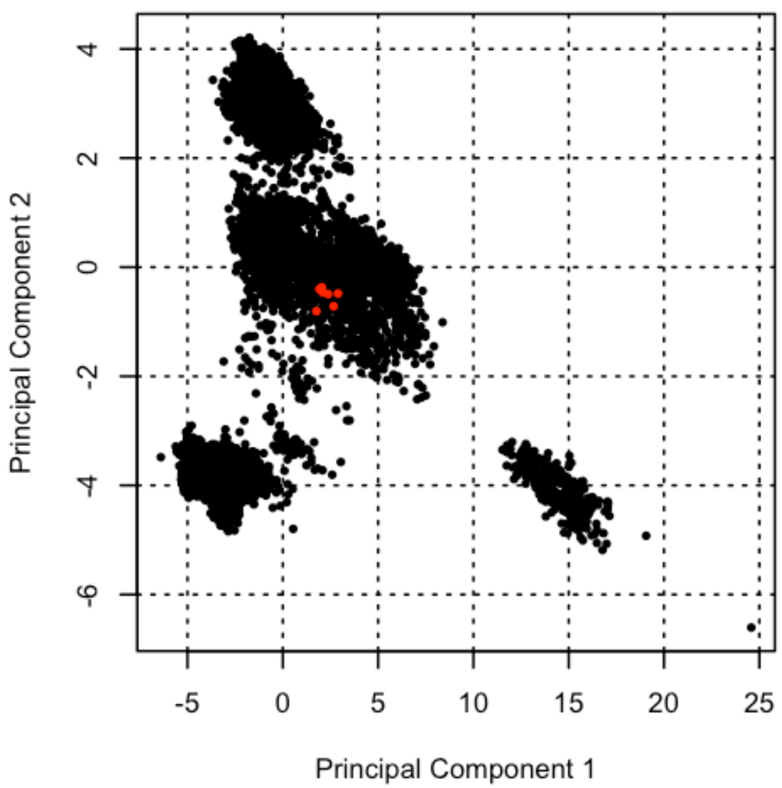
Bacteria : Lentisphaerae

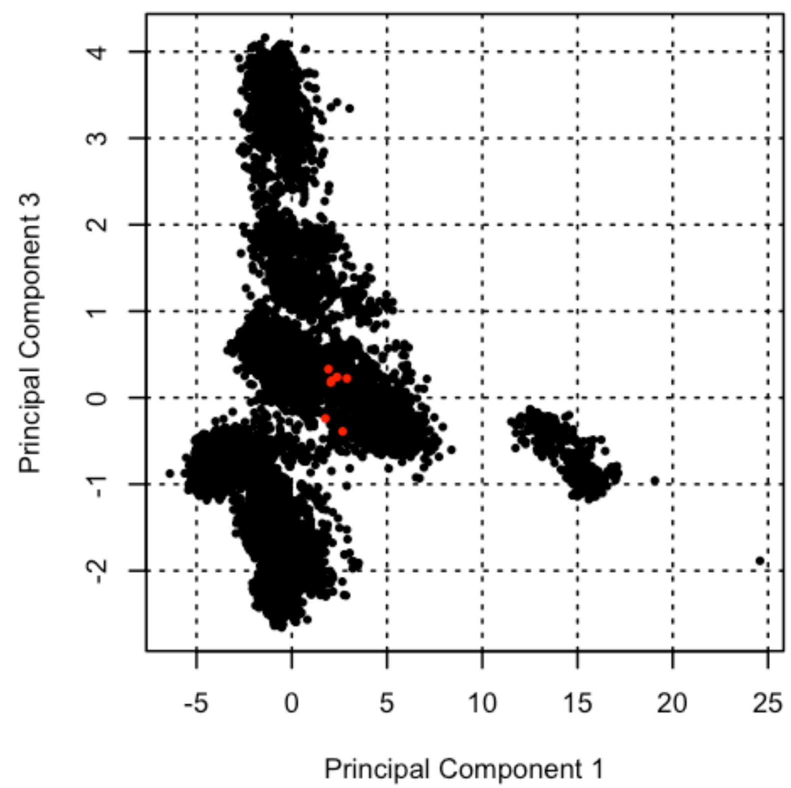
Synergistetes

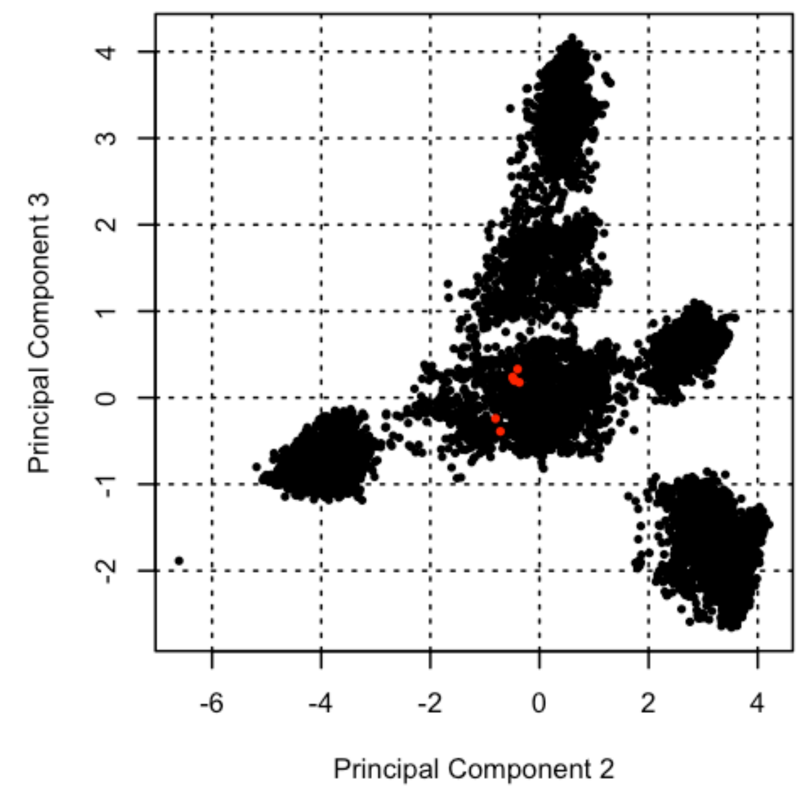


Bacteria : Synergistetes

Nitrospirae

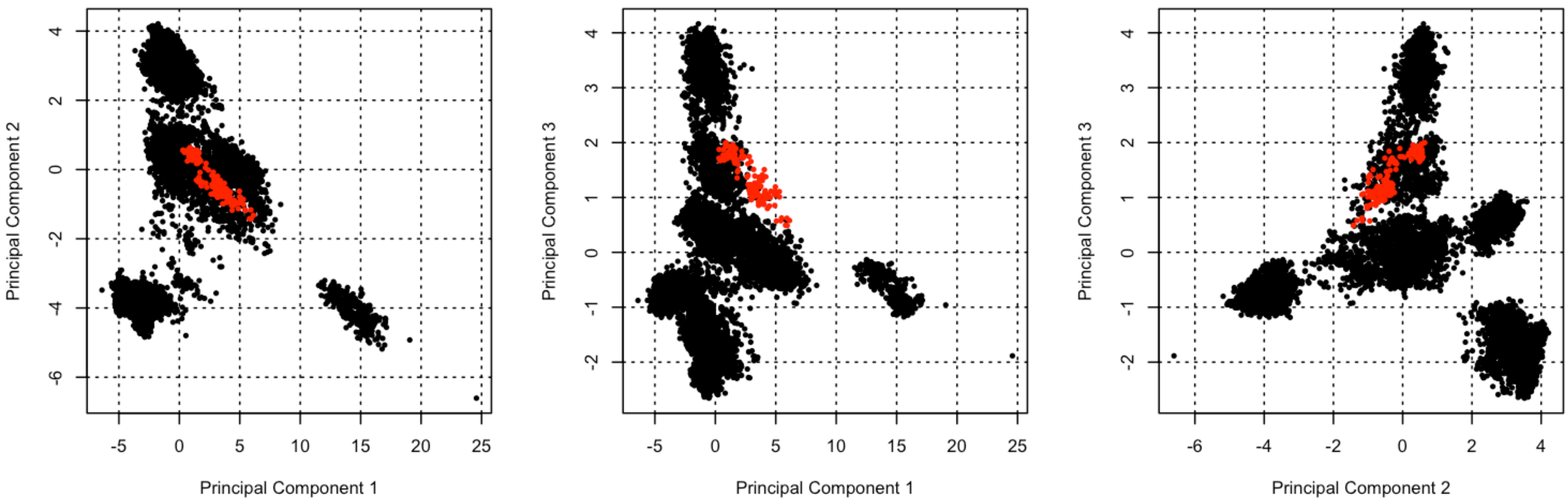






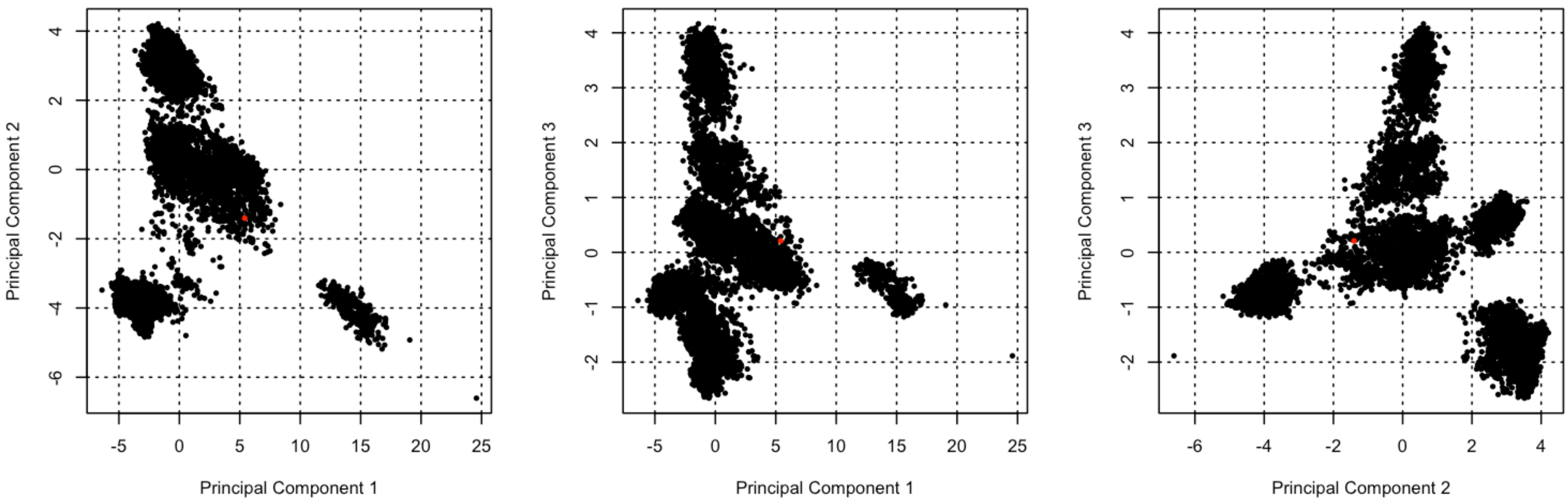
Bacteria : Nitrospirae

Tenericutes



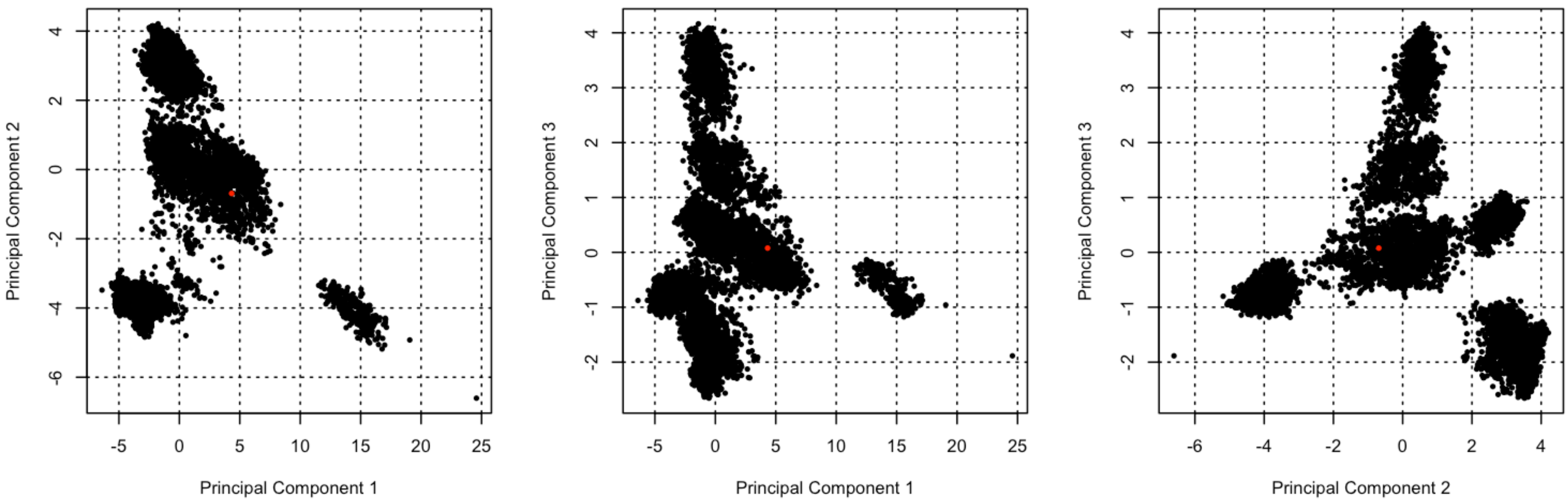
Bacteria : Tenericutes

Caldiserica



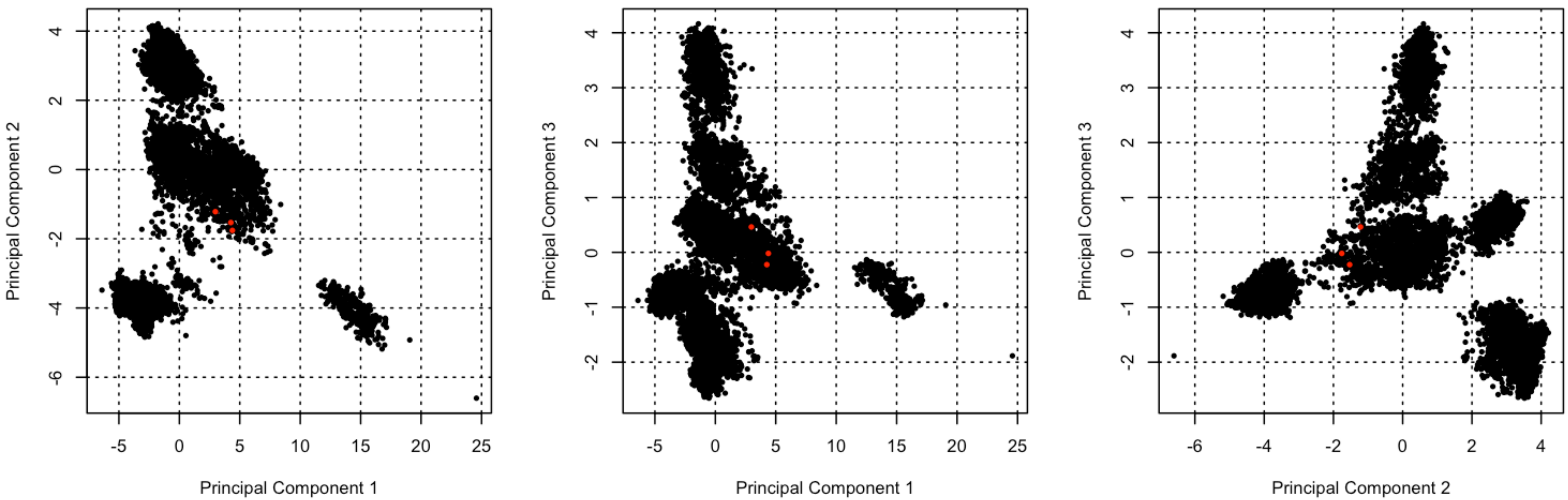
Bacteria : Caldiserica

Elusimicrobia



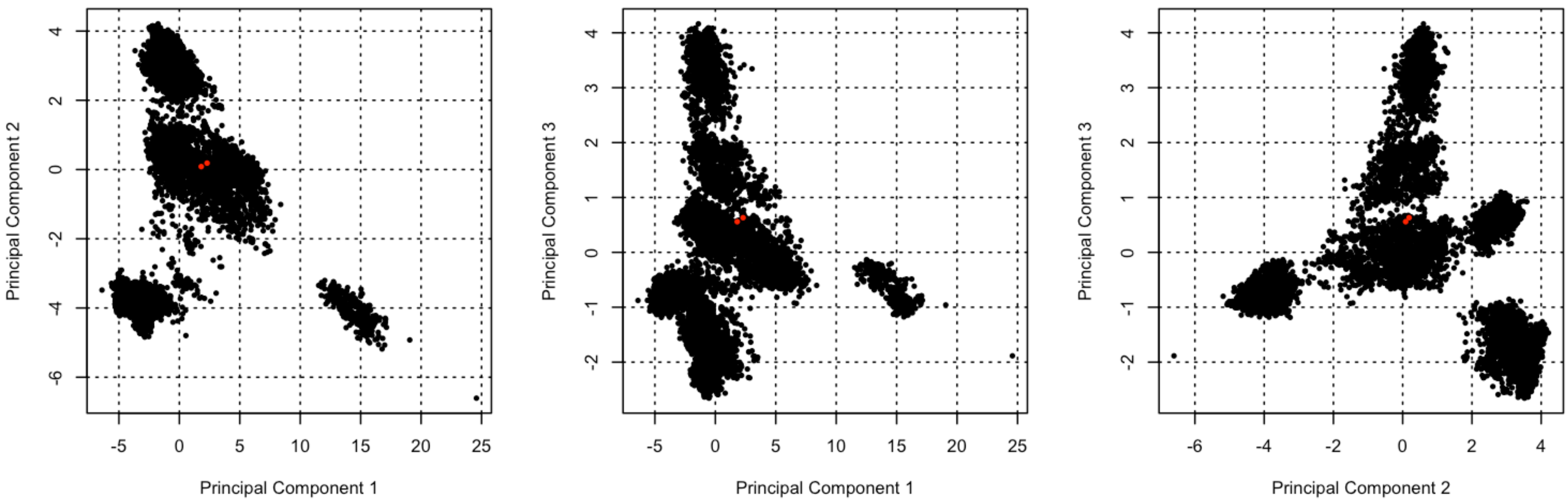
Bacteria : Elusimicrobia

Armatimonadetes



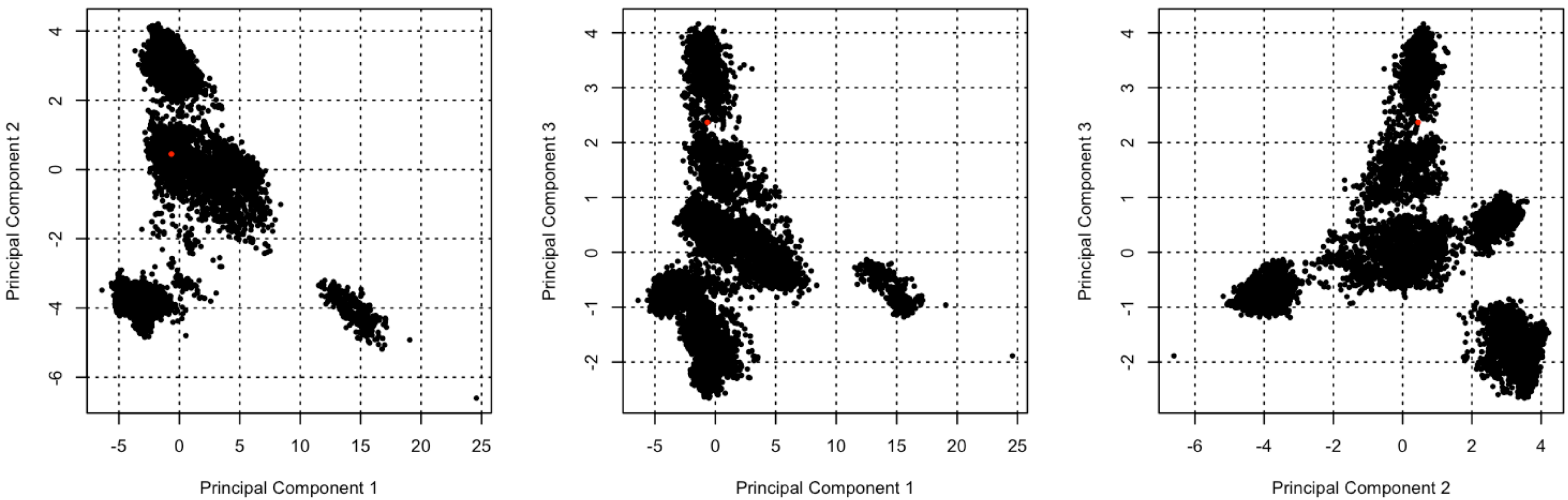
Bacteria : Armatimonadetes

Ignavibacteriae



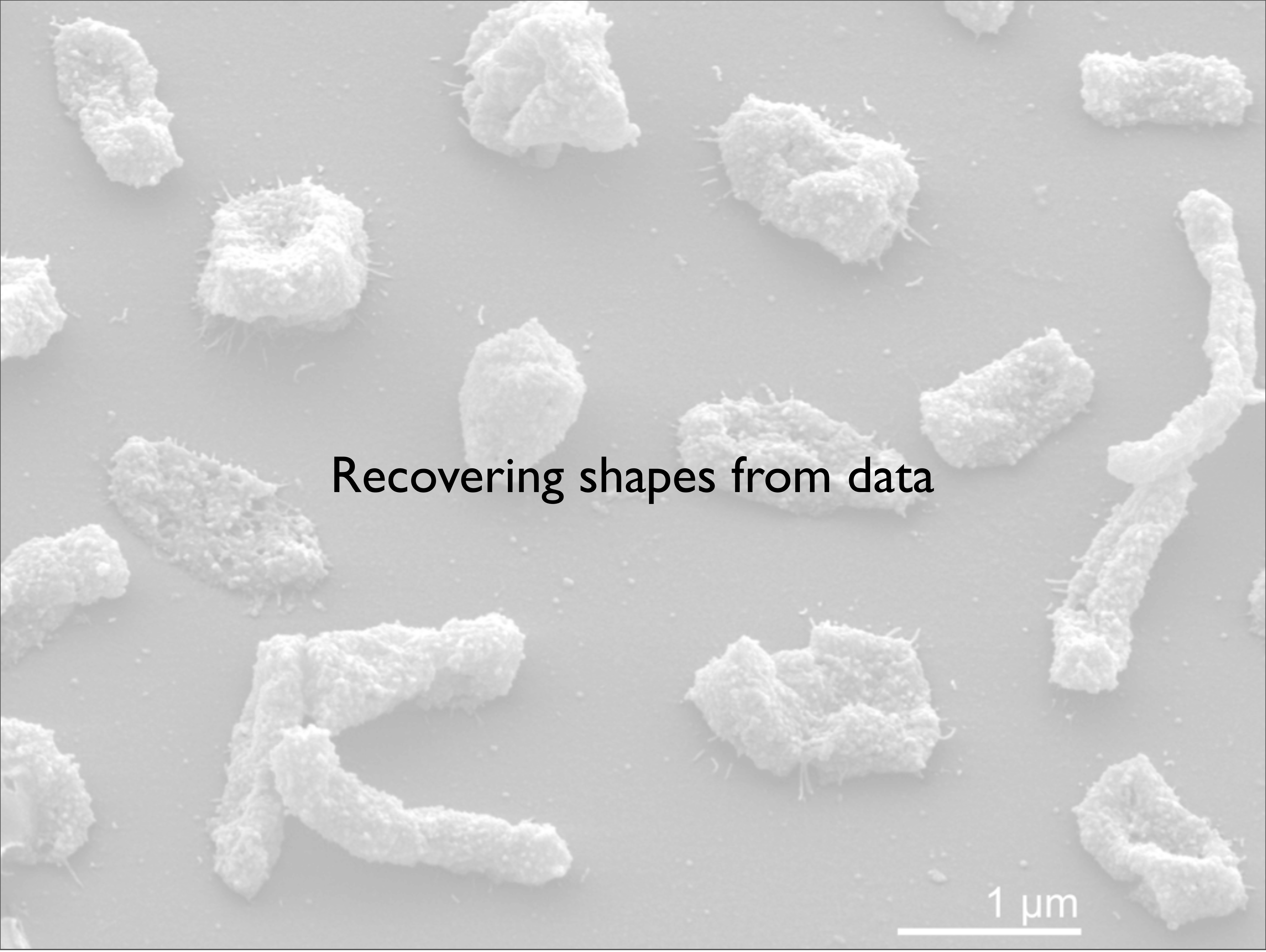
Bacteria : Ignavibacteriae

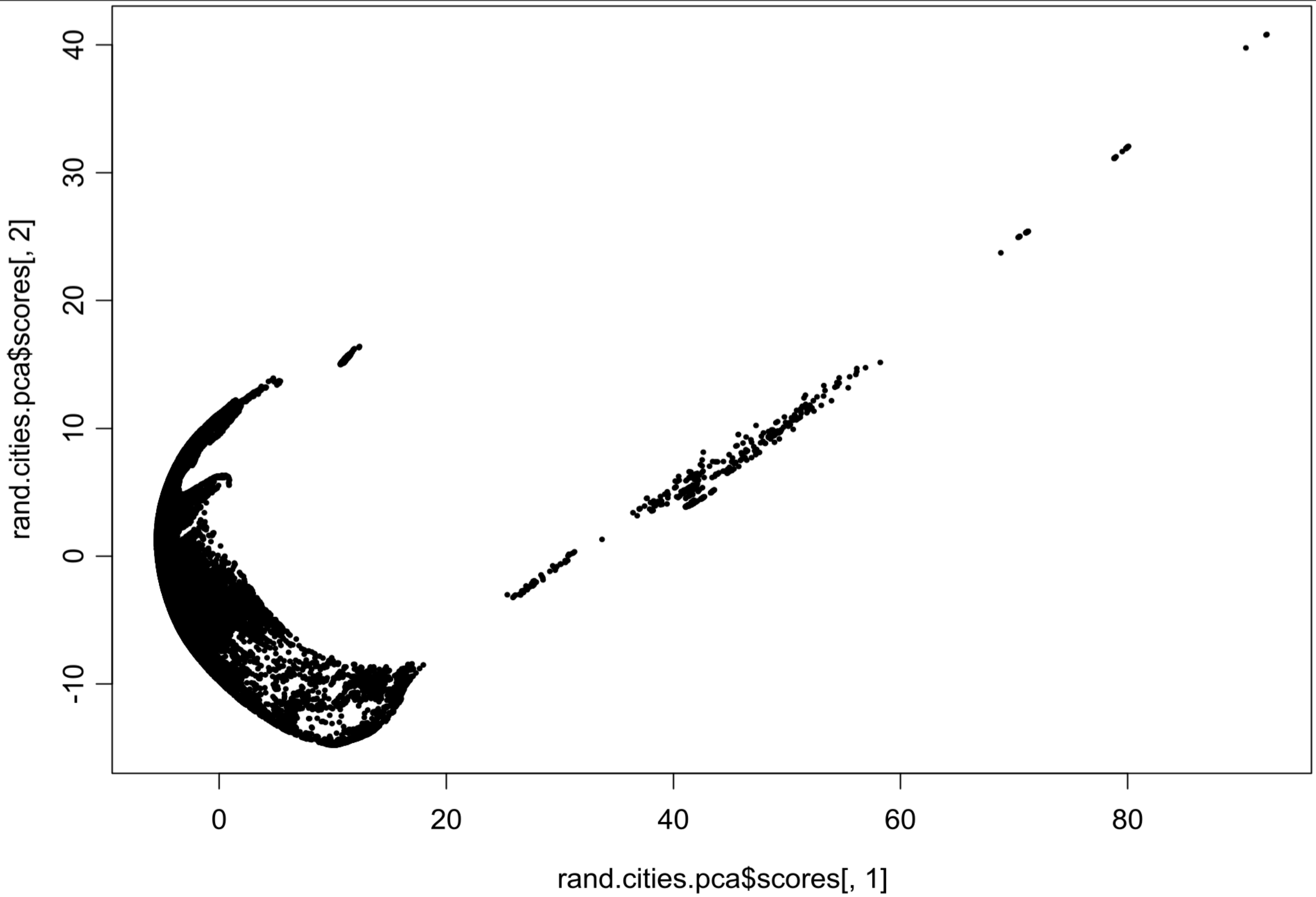
Incertae sedis 172



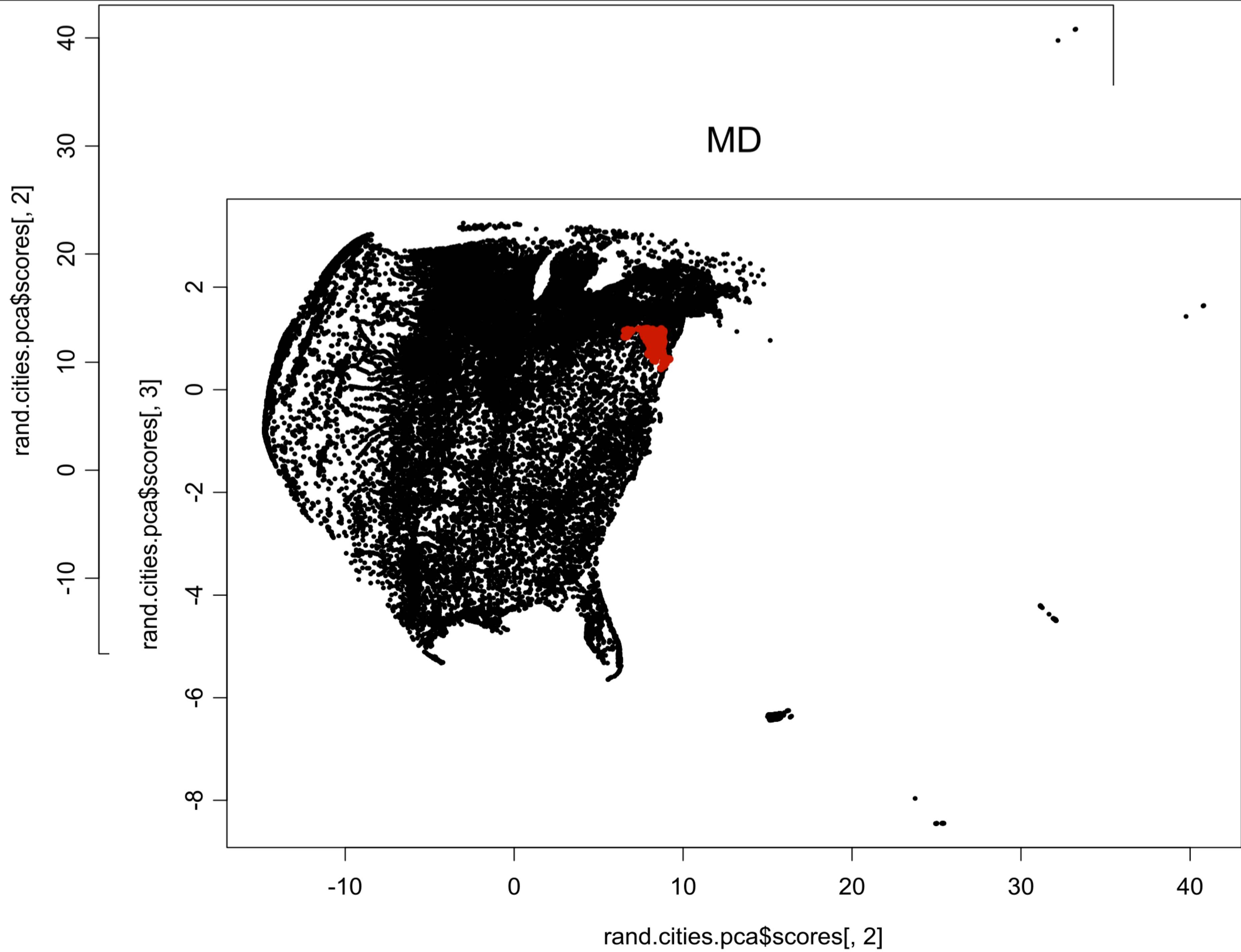
Bacteria : Incertae sedis 172

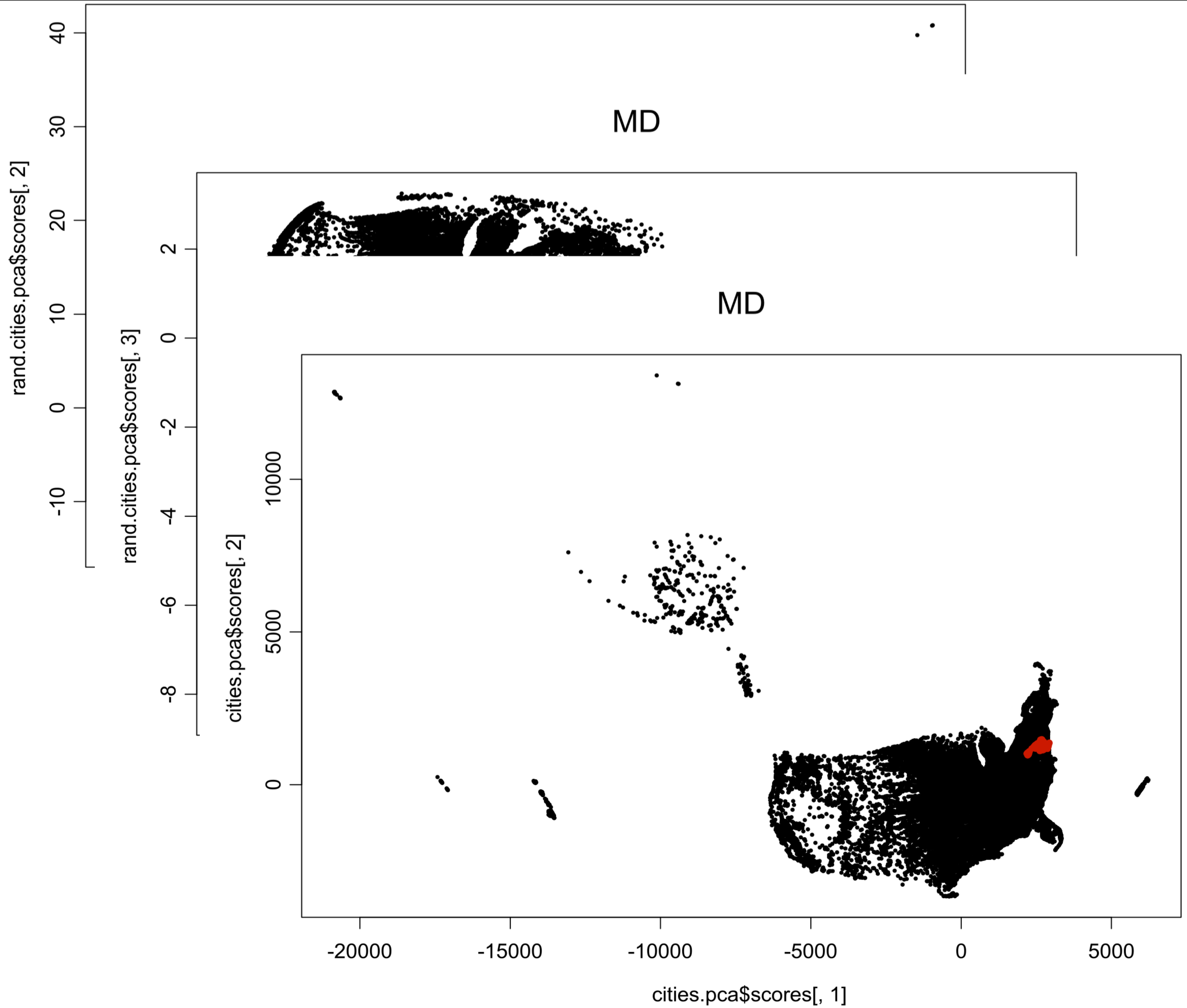
Recovering shapes from data



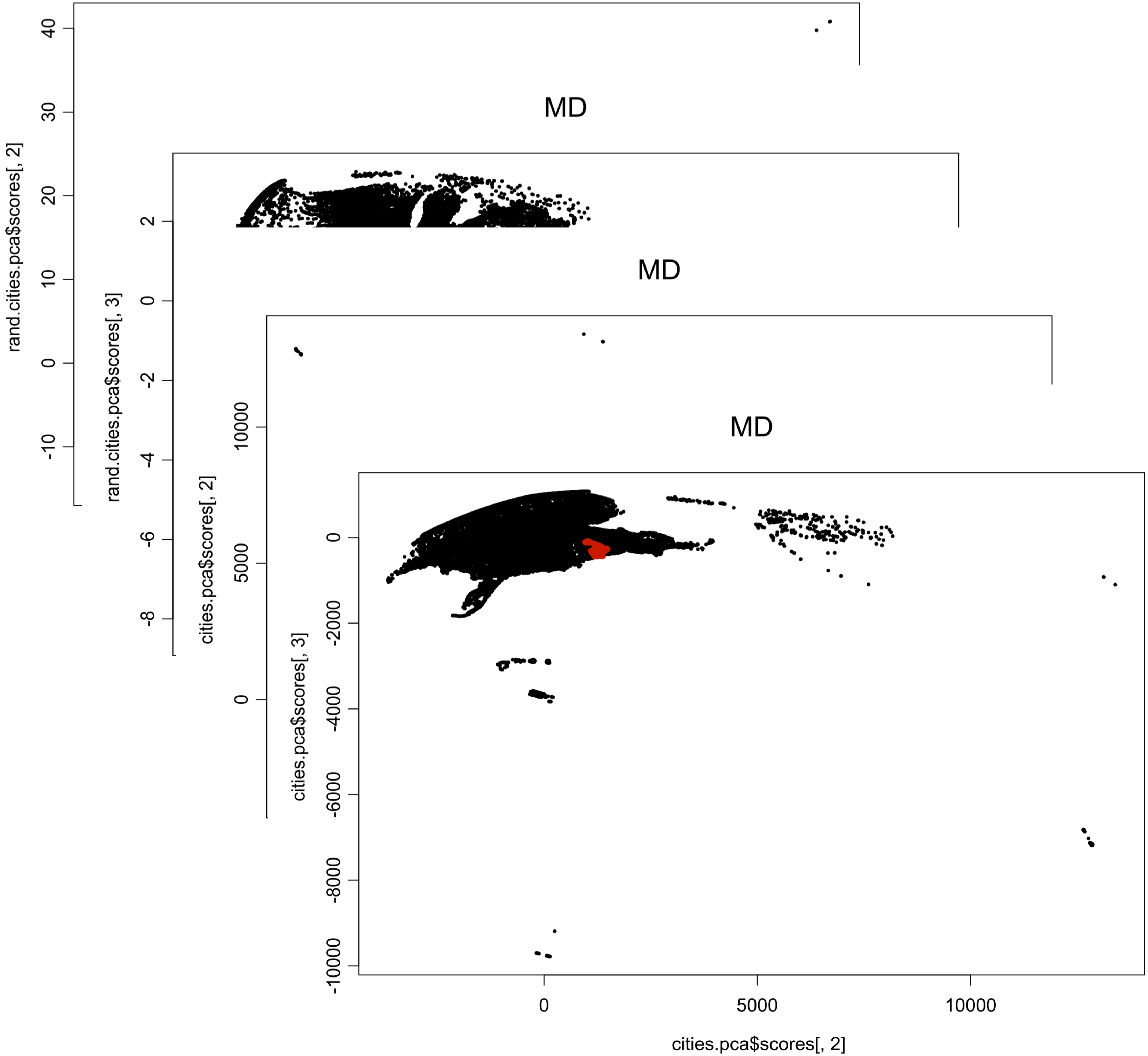


rand.cities.pca$scores[, 2]
40
30
20
10
0
-10
0
20
40
60
80
rand.cities.pca$scores[, 1]

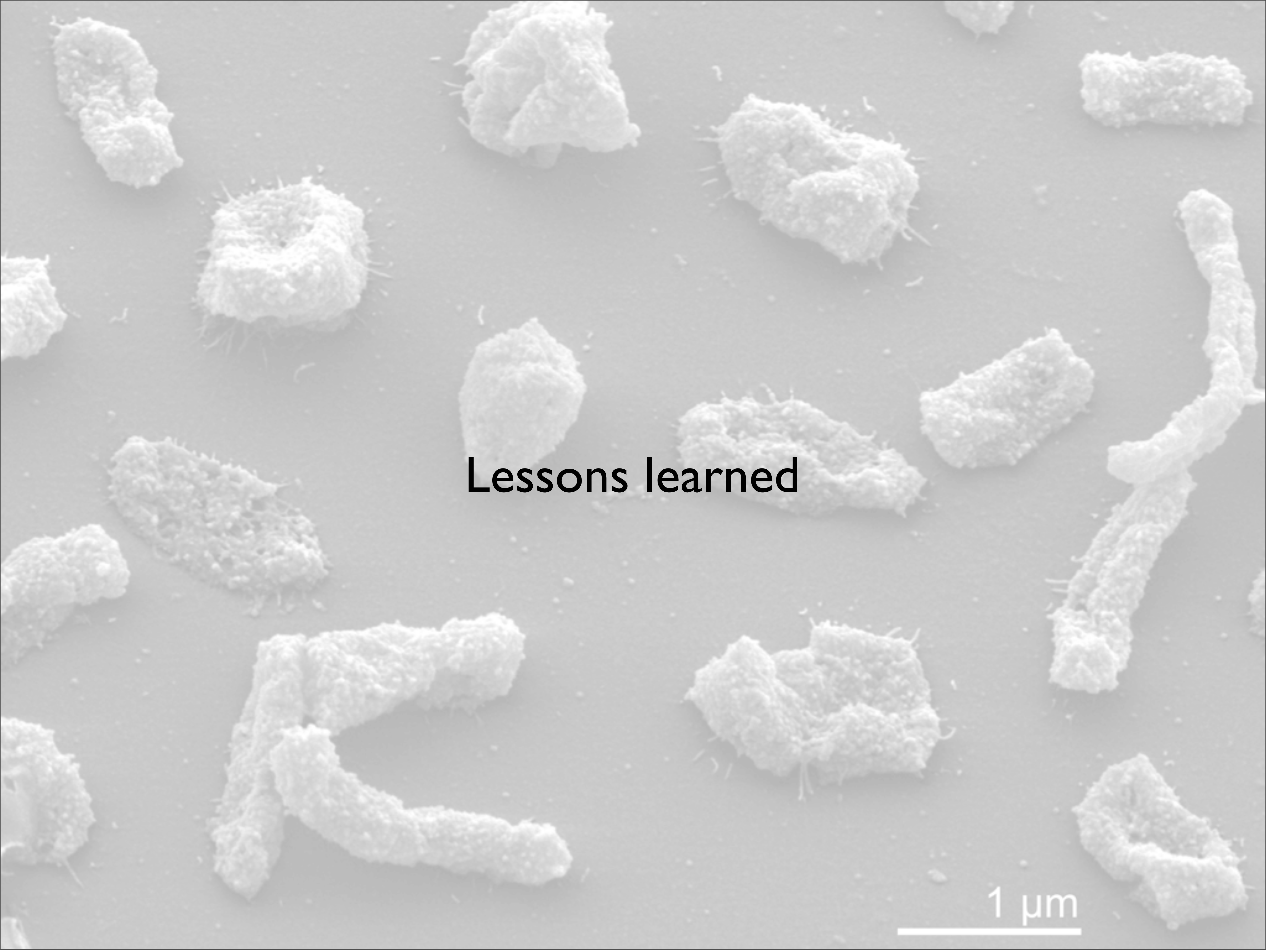

MD
rand.cities.pca$scores[, 2]
rand.cities.pca$scores[, 3]
rand.cities.pca$scores[, 2]
40
30
20
10
0
-10
2
0
-2
-4
-6
-8
-10
0
10
20
30
40



MD
MD
rand.cities.pca$scores[, 2]
rand.cities.pca$scores[, 3]
cities.pca$scores[, 2]
cities.pca$scores[, 1]
40
30
20
10
0
-10
2
0
-2
-4
-6
-8
10000
5000
0
-20000
-15000
-10000
-5000
0
5000


MD
MD
MD
rand.cities.pca$scores[, 2]
rand.cities.pca$scores[, 3]
cities.pca$scores[, 2]
cities.pca$scores[, 3]
cities.pca$scores[, 2]
0
5000
10000
-10000
-8000
-6000
-4000
-2000
0

Lessons learned



Untitled — Edited

/Volumes/Drive 2/khalid2/ASM NGS Lecture - RStudio

ASM NGS Lecture

aai.dist | kmer.dist | exemplar.16S.dist | exemplar.subset.dist | ani.dist2 | lsa.50.words | lsa.10.words.ordered | N4L.16S.dist.values.20150901

Filter

	ex.23951	ex.23950	ex.23955	ex.23954	ex.23952	ex.23958	ex.23956	ex.23942	ex.23941	ex.23948	ex.23949	ex.23977	ex.23975	ex.23973	ex.23972
ex.23951	0.00000	0.23080	0.19270	0.37520	0.23210	0.24350	0.18670	0.22590	0.2916	0.2231	0.23060	0.23490	0.25210	0.24480	0.23490
ex.23950	0.23080	0.00000	0.24450	0.40430	0.25600	0.12620	0.22950	0.21780	0.3092	0.2101	0.03252	0.11290	0.14340	0.10040	0.08921
ex.23955	0.19270	0.24450	0.00000	0.38580	0.21000	0.24910	0.04436	0.23530	0.2803	0.2424	0.23980	0.23830	0.25600	0.24090	0.23480
ex.23954	0.37520	0.40430	0.38580	0.00000	0.39740	0.40540	0.37360	0.39770	0.4142	0.3965	0.40110	0.39980	0.40260	0.39200	0.39250
ex.23952	0.23210	0.25600	0.21000	0.39740	0.00000	0.26840	0.20090	0.26200	0.3188	0.2624	0.25420	0.25750	0.27280	0.26210	0.24970
ex.23958	0.24350	0.12620	0.24910	0.40540	0.26840	0.00000	0.23730	0.22740	0.3128	0.2224	0.12480	0.13770	0.16860	0.13780	0.13740
ex.23956	0.18670	0.22950	0.04436	0.37360	0.20090	0.23730	0.00000	0.23340	0.2751	0.2298	0.22090	0.22610	0.24630	0.22970	0.22240
ex.23942	0.22590	0.21780	0.23530	0.39770	0.26200	0.22740	0.23340	0.00000	0.3105	0.1135	0.21530	0.22500	0.23500	0.22940	0.22180
ex.23941	0.29160	0.30920	0.28030	0.41420	0.31880	0.31280	0.27510	0.31050	0.0000	0.3238	0.30880	0.30470	0.30670	0.30800	0.29930
ex.23948	0.22310	0.21010	0.24240	0.39650	0.26240	0.22240	0.22980	0.11350	0.3238	0.0000	0.21010	0.21490	0.23140	0.21920	0.21530
ex.23949	0.23060	0.03252	0.23980	0.40110	0.25420	0.12480	0.22090	0.21530	0.3088	0.2101	0.00000	0.11410	0.14940	0.09621	0.08075
ex.23977	0.23490	0.11290	0.23830	0.39980	0.25750	0.13770	0.22610	0.22500	0.3047	0.2149	0.11410	0.00000	0.16150	0.11860	0.12180
ex.23975	0.25210	0.14340	0.25600	0.40260	0.27280	0.16860	0.24630	0.23500	0.3067	0.2314	0.14940	0.16150	0.00000	0.14400	0.14460
ex.23973	0.24480	0.10040	0.24090	0.39200	0.26210	0.13780	0.22970	0.22940	0.3080	0.2192	0.09621	0.11860	0.14400	0.00000	0.09775
ex.23972	0.23490	0.08921	0.23480	0.39250	0.24970	0.13740	0.22240	0.22180	0.2993	0.2153	0.08075	0.12180	0.14460	0.09775	0.00000
ex.23970	0.23930	0.12770	0.24930	0.39490	0.27130	0.14660	0.22720	0.22450	0.3095	0.2336	0.12470	0.12030	0.15240	0.13700	0.12700
ex.23978	0.26380	0.14950	0.26680	0.41030	0.27380	0.17180	0.24700	0.25390	0.3279	0.2525	0.14520	0.15460	0.17280	0.13020	0.14680
ex.23963	0.23310	0.21570	0.25540	0.38970	0.26170	0.23280	0.24060	0.24440	0.3103	0.2315	0.20980	0.22670	0.23160	0.21830	0.20950
ex.23966	0.23910	0.11270	0.24080	0.38830	0.25230	0.13940	0.22250	0.22150	0.3167	0.2126	0.10640	0.12380	0.12330	0.11740	0.10730
ex.23965	0.23860	0.11590	0.24070	0.38900	0.25440	0.13590	0.22260	0.22620	0.3104	0.2150	0.11000	0.12380	0.11690	0.12240	0.10340
ex.23968	0.23640	0.12410	0.24200	0.39660	0.25820	0.13770	0.22470	0.22330	0.3098	0.2151	0.11750	0.13540	0.12320	0.12100	0.11100
ex.23967	0.24300	0.11550	0.24090	0.39410	0.25240	0.14200	0.22350	0.22700	0.3137	0.2177	0.11140	0.13000	0.12650	0.11720	0.10960
ex.23994	0.23640	0.23730	0.26230	0.40450	0.27350	0.24250	0.25360	0.21710	0.3192	0.2154	0.23110	0.22980	0.24250	0.23880	0.23490
ex.23995	0.23860	0.24790	0.25460	0.39950	0.27470	0.25740	0.24130	0.16880	0.3214	0.1895	0.25070	0.24190	0.25440	0.24540	0.25380
ex.23993	0.23820	0.24350	0.24730	0.40750	0.25980	0.24790	0.24360	0.22500	0.3092	0.2209	0.24690	0.24440	0.25600	0.25050	0.25450
ex.23998	0.16820	0.24340	0.21560	0.38700	0.24320	0.25820	0.21010	0.22790	0.3070	0.2255	0.24370	0.24530	0.26720	0.24790	0.24000
ex.23999	0.20430	0.22770	0.16040	0.39280	0.17130	0.24240	0.15430	0.22060	0.2875	0.2356	0.22550	0.23140	0.24720	0.23830	0.22020
ex.23996	0.21400	0.25170	0.18490	0.37910	0.17360	0.26350	0.17580	0.24330	0.2889	0.2543	0.25000	0.25030	0.26040	0.25210	0.24140
ex.23997	0.21760	0.24910	0.18910	0.37940	0.17570	0.26580	0.18030	0.24980	0.2926	0.2617	0.24680	0.24910	0.26130	0.24810	0.23820
ex.23991	0.23550	0.23640	0.24100	0.40290	0.25870	0.24830	0.23740	0.19100	0.2952	0.2011	0.23200	0.22640	0.24040	0.23840	0.23440
ex.23989	0.22650	0.21900	0.25000	0.39350	0.27370	0.23060	0.23760	0.14130	0.3177	0.1475	0.21980	0.22190	0.24220	0.22310	0.22210
ex.23981	0.24440	0.12400	0.23860	0.39250	0.25860	0.14520	0.22370	0.22720	0.3038	0.2255	0.11820	0.14100	0.12540	0.12770	0.11130
ex.23982	0.21900	0.23550	0.24010	0.39820	0.26620	0.24050	0.22900	0.14570	0.3107	0.1698	0.23480	0.23600	0.24700	0.24190	0.23330
ex.23983	0.22950	0.22570	0.24710	0.40180	0.27570	0.22850	0.23270	0.13760	0.3215	0.1343	0.22420	0.23120	0.23420	0.22560	0.21820
ex.23984	0.23530	0.22950	0.23900	0.39730	0.26940	0.23420	0.22940	0.11950	0.3169	0.1398	0.22330	0.23260	0.23320	0.23180	0.22380
ex.23985	0.22600	0.21650	0.22740	0.39690	0.24390	0.23290	0.21670	0.18110	0.2968	0.1803	0.21730	0.22340	0.24390	0.21770	0.20720
ex.23986	0.23150	0.24090	0.25100	0.40810	0.27630	0.25290	0.23280	0.13010	0.3317	0.1717	0.23800	0.24490	0.25710	0.24690	0.24630
ex.23987	0.23070	0.24380	0.23770	0.39830	0.25020	0.25730	0.22720	0.16080	0.3129	0.1782	0.24320	0.24530	0.25360	0.23280	0.24200
ex.23988	0.22960	0.23530	0.23660	0.39440	0.26660	0.24060	0.22070	0.10400	0.3211	0.1477	0.23000	0.23320	0.24060	0.24030	0.23050
ex.23980	0.24000	0.11390	0.25440	0.39670	0.26540	0.14120	0.24250	0.23340	0.3057	0.2288	0.12110	0.12680	0.14240	0.13090	0.13580
ex.11460	0.22490	0.22270	0.21720	0.40000	0.23960	0.23480	0.20690	0.24070	0.2988	0.2318	0.21550	0.22550	0.26030	0.23300	0.21820
ex.11465	0.15710	0.22370	0.20850	0.37640	0.23350	0.22820	0.20710	0.21090	0.2912	0.2017	0.21920	0.22860	0.24090	0.22630	0.22450
ex.11466	0.14810	0.22360	0.20360	0.37380	0.22730	0.23240	0.20180	0.20580	0.2922	0.2066	0.21980	0.22630	0.23890	0.22550	0.22370

Showing 1 to 44 of 13,125 entries

Console

Environment | History

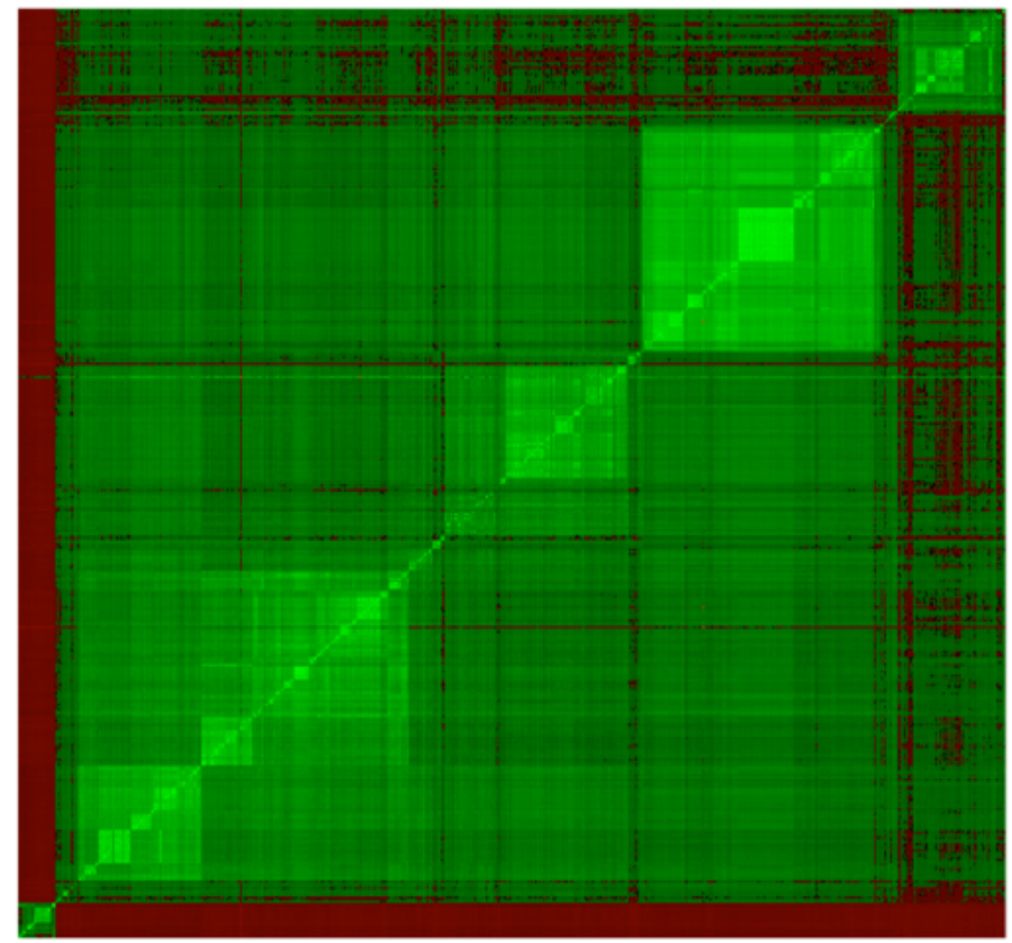
Files | Plots | Packages | Help | Viewer

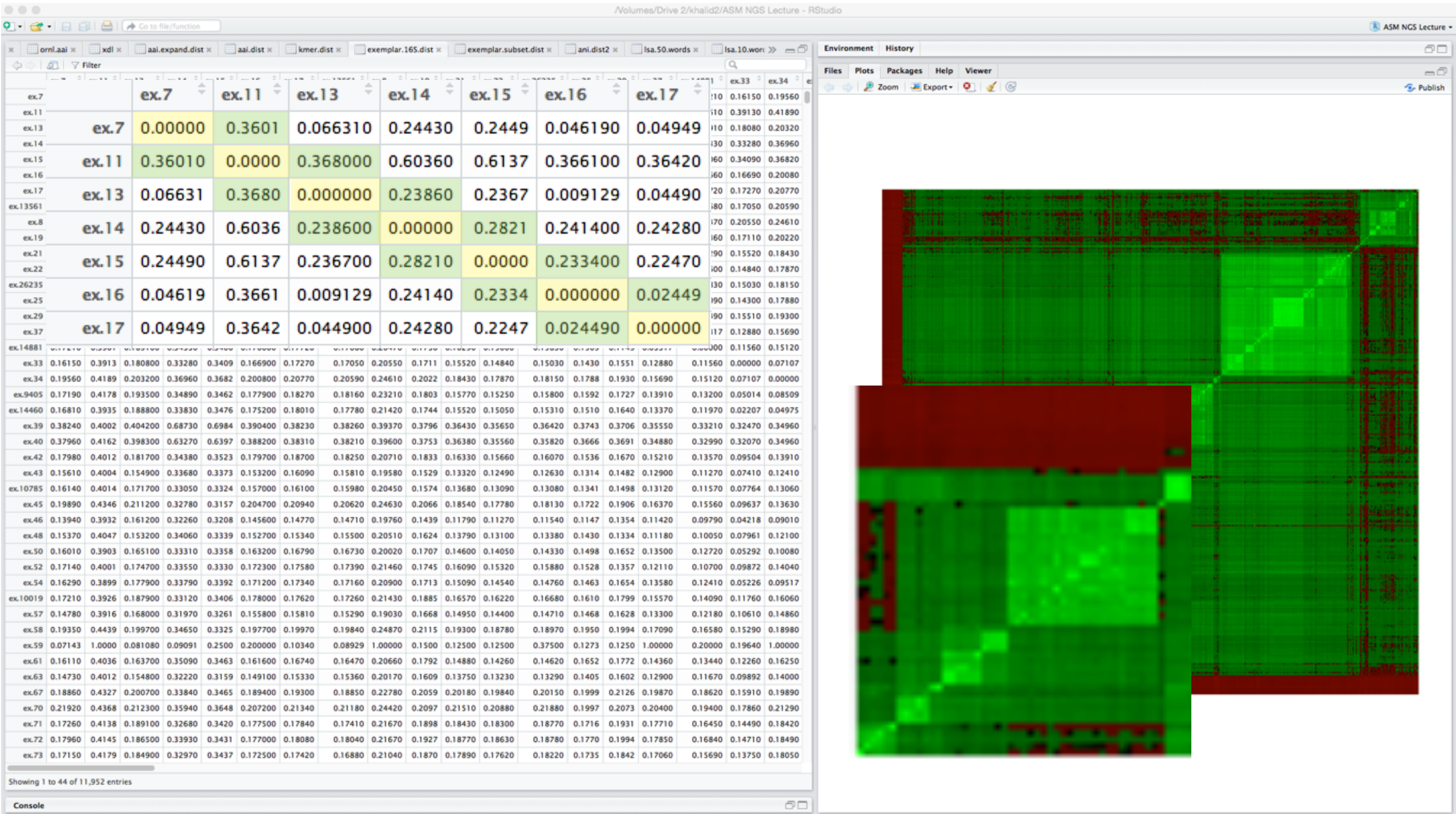
Zoom | Export | Publish

/Volumes/Drive 2/khalid2/ASM NGS Lecture - RStudio

	ex.7	ex.11	ex.13	ex.14	ex.15	ex.16	ex.17	ex.13561	ex.8	ex.19	ex.21	ex.22	ex.26235	ex.25	ex.29	ex.37	ex.14881	ex.33	ex.34
ex.7	0.00000	0.3601	0.066310	0.24430	0.2449	0.046190	0.04949	0.04136	0.08813	0.1451	0.11740	0.11200	0.11930	0.1580	0.2046	0.18330	0.17210	0.16150	0.19560
ex.11	0.36010	0.0000	0.368000	0.60360	0.6137	0.366100	0.36420	0.36030	0.34580	0.3829	0.37350	0.37100	0.36900	0.4012	0.4311	0.40010	0.39610	0.39130	0.41890
ex.13	0.06631	0.3680	0.000000	0.23860	0.2367	0.009129	0.04490	0.04268	0.06567	0.1514	0.11710	0.11300	0.11810	0.1717	0.2079	0.18480	0.18910	0.18080	0.20320
ex.14	0.24430	0.6036	0.238600	0.00000	0.2821	0.241400	0.24280	0.23330	0.36070	0.3102	0.29680	0.29370	0.29960	0.3031	0.3531	0.35080	0.34330	0.33280	0.36960
ex.15	0.24490	0.6137	0.236700	0.28210	0.0000	0.233400	0.22470	0.23420	0.35110	0.3103	0.29680	0.29350	0.29200	0.3141	0.3548	0.34540	0.34060	0.34090	0.36820
ex.16	0.04619	0.3661	0.009129	0.24140	0.2334	0.000000	0.02449	0.02172	0.05536	0.1477	0.11340	0.10850	0.11220	0.1596	0.2071	0.18340	0.17660	0.16690	0.20080
ex.17	0.04949	0.3642	0.044900	0.24280	0.2247	0.024490	0.00000	0.02510	0.06185	0.1487	0.11930	0.11310	0.11690	0.1616	0.2050	0.18800	0.17720	0.17270	0.20770
ex.13561	0.04136	0.3603	0.042680	0.23330	0.2342	0.021720	0.02510	0.00000	0.06283	0.1464	0.11590	0.11180	0.11570	0.1545	0.2058	0.18680	0.17680	0.17050	0.20590
ex.8	0.08813	0.3458	0.065670	0.36070	0.3511	0.055360	0.06185	0.06283	0.00000	0.1881	0.15320	0.14880	0.15030	0.2097	0.2453	0.21530	0.20470	0.20550	0.24610
ex.19	0.14510	0.3829	0.151400	0.31020	0.3103	0.147700	0.14870	0.14640	0.18810	0.0000	0.10870	0.10330	0.10900	0.1659	0.1976	0.18560	0.17360	0.17110	0.20220
ex.21	0.11740	0.3735	0.117100	0.29680	0.2968	0.113400	0.11930	0.11590	0.15320	0.1087	0.00000	0.01691	0.03320	0.1440	0.1965	0.17460	0.16290	0.15520	0.18430
ex.22	0.11200	0.3710	0.113000	0.29370	0.2935	0.108500	0.11310	0.11180	0.14880	0.1033	0.01691	0.00000	0.02282	0.1406	0.1924	0.16830	0.15600	0.14840	0.17870
ex.26235	0.11930	0.3690	0.118100	0.29960	0.2920	0.112200	0.11690	0.11570	0.15030	0.1090	0.03320	0.02282	0.00000	0.1405	0.1981	0.16980	0.15830	0.15030	0.18150
ex.25	0.15800	0.4012	0.171700	0.30310	0.3141	0.159600	0.16160	0.15450	0.20970	0.1659	0.14400	0.14060	0.14050	0.0000	0.1880	0.16640	0.15090	0.14300	0.17880
ex.29	0.20460	0.4311	0.207900	0.35310	0.3548	0.207100	0.20500	0.20580	0.24530	0.1976	0.19650	0.19240	0.19810	0.1880	0.0000	0.13550	0.11490	0.15510	0.19300
ex.37	0.18330	0.4001	0.184800	0.35080	0.3454	0.183400	0.18800	0.18680	0.21530	0.1856	0.17460	0.16830	0.16980	0.1664	0.1355	0.00000	0.03317	0.12880	0.15690
ex.14881	0.17210	0.3961	0.189100	0.34330	0.3406	0.176600	0.17720	0.17680	0.20470	0.1736	0.16290	0.15600	0.15830	0.1509	0.1149	0.03317	0.00000	0.11560	0.15120
ex.33	0.16150	0.3913	0.180800	0.33280	0.3409	0.166900	0.17270	0.17050	0.20550	0.1711	0.15520	0.14840	0.15030	0.1430	0.1551	0.12880	0.11560	0.00000	0.07107
ex.34	0.19560	0.4189	0.203200	0.36960	0.3682	0.200800	0.20770	0.20590	0.24610	0.2022	0.18430	0.17870	0.18150	0.1788	0.1930	0.15690	0.15120	0.07107	0.00000
ex.9405	0.17190	0.4178	0.193500	0.34890	0.3462	0.177900	0.18270	0.18160	0.23210	0.1803	0.15770	0.15250	0.15800	0.1592	0.1727	0.13910	0.13200	0.05014	0.08509
ex.14460	0.16810	0.3935	0.188800	0.33830	0.3476	0.175200	0.18010	0.17780	0.21420	0.1744	0.15520	0.15050	0.15310	0.1510	0.1640	0.13370	0.11970	0.02207	0.04975
ex.39	0.38240	0.4002	0.404200	0.68730	0.6984	0.390400	0.38230	0.38260	0.39370	0.3796	0.36430	0.35650	0.36420	0.3743	0.3706	0.35550	0.33210	0.32470	0.34960
ex.40	0.37960	0.4162	0.398300	0.63270	0.6397	0.388200	0.38310	0.38210	0.39600	0.3753	0.36380	0.35560	0.35820	0.3666	0.3691	0.34880	0.32990	0.32070	0.34960
ex.42	0.17980	0.4012	0.181700	0.34380	0.3523	0.179700	0.18700	0.18250	0.20710	0.1833	0.16330	0.15660	0.16070	0.1536	0.1670	0.15210	0.13570	0.09504	0.13910
ex.43	0.15610	0.4004	0.154900	0.33680	0.3373	0.153200	0.16090	0.15810	0.19580	0.1529	0.13320	0.12490	0.12630	0.1314	0.1482	0.12900	0.11270	0.07410	0.12410
ex.10785	0.16140	0.4014	0.171700	0.33050	0.3324	0.157000	0.16100	0.15980	0.20450	0.1574	0.13680	0.13090	0.13080	0.1341	0.1498	0.13120	0.11570	0.07764	0.13060
ex.45	0.19890	0.4346	0.211200	0.32780	0.3157	0.204700	0.20940	0.20620	0.24630	0.2066	0.18540	0.17780	0.18130	0.1722	0.1906	0.16370	0.15560	0.09637	0.13630
ex.46	0.13940	0.3932	0.161200	0.32260	0.3208	0.145600	0.14770	0.14710	0.19760	0.1439	0.11790	0.11270	0.11540	0.1147	0.1354	0.11420	0.09790	0.04218	0.09010
ex.48	0.15370	0.4047	0.153200	0.34060	0.3339	0.152700	0.15340	0.15500	0.20510	0.1624	0.13790	0.13100	0.13380	0.1430	0.1334	0.11180	0.10050	0.07961	0.12100
ex.50	0.16010	0.3903	0.165100	0.33310	0.3358	0.163200	0.16790	0.16730	0.20020	0.1707	0.14600	0.14050	0.14330	0.1498	0.1652	0.13500	0.12720	0.05292	0.10080
ex.52	0.17140	0.4001	0.174700	0.33550	0.3330	0.172300	0.17580	0.17390	0.21460	0.1745	0.16090	0.15320	0.15880	0.1528	0.1357	0.12110	0.10700	0.09872	0.14040
ex.54	0.16290	0.3899	0.177900	0.33790	0.3392	0.171200	0.17340	0.17160	0.20900	0.1713	0.15090	0.14540	0.14760	0.1463	0.1654	0.13580	0.12410	0.05226	0.09517
ex.10019	0.17210	0.3926	0.187900	0.33120	0.3406	0.178000	0.17620	0.17260	0.21430	0.1885	0.16570	0.16220	0.16680	0.1610	0.1799	0.15570	0.14090	0.11760	0.16060
ex.57	0.14780	0.3916	0.168000	0.31970	0.3261	0.155800	0.15810	0.15290	0.19030	0.1668	0.14950	0.14400	0.14710	0.1468	0.1628	0.13300	0.12180	0.10610	0.14860
ex.58	0.19350	0.4439	0.199700	0.34650	0.3325	0.197700	0.19970	0.19840	0.24870	0.2115	0.19300	0.18780	0.18970	0.1950	0.1994	0.17090	0.16580	0.15290	0.18980
ex.59	0.07143	1.0000	0.081080	0.09091	0.2500	0.200000	0.10340	0.08929	1.00000	0.1500	0.12500	0.12500	0.37500	0.1273	0.1250	1.00000	0.20000	0.19640	1.00000
ex.61	0.16110	0.4036	0.163700	0.35090	0.3463	0.161600	0.16740	0.16470	0.20660	0.1792	0.14880	0.14260	0.14620	0.1652	0.1772	0.14360	0.13440	0.12260	0.16250
ex.63	0.14730	0.4012	0.154800	0.32220	0.3159	0.149100	0.15330	0.15360	0.20170	0.1609	0.13750	0.13230	0.13290	0.1405	0.1602	0.12900	0.11670	0.09892	0.14000
ex.67	0.18860	0.4327	0.200700	0.33840	0.3465	0.189400	0.19300	0.18850	0.22780	0.2059	0.20180	0.19840	0.20150	0.1999	0.2126	0.19870	0.18620	0.15910	0.19890
ex.70	0.21920	0.4368	0.212300	0.35940	0.3648	0.207200	0.21340	0.21180	0.24420	0.2097	0.21510	0.20880	0.21880	0.1997	0.2073	0.20400	0.19400	0.17860	0.21290
ex.71	0.17260	0.4138	0.189100	0.32680	0.3420	0.177500	0.17840	0.17410	0.21670	0.1898	0.18430	0.18300	0.18770	0.1716	0.1931	0.17710	0.16450	0.14490	0.18420
ex.72	0.17960	0.4145	0.186500	0.33930	0.3431	0.177000	0.18080	0.18040	0.21670	0.1927	0.18770	0.18630	0.18780	0.1770	0.1994	0.17850	0.16840	0.14710	0.18490
ex.73	0.17150	0.4179	0.184900	0.32970	0.3437	0.172500	0.17420	0.16880	0.21040	0.1870	0.17890	0.17620	0.18220	0.1735	0.1842	0.17060	0.15690	0.13750	0.18050

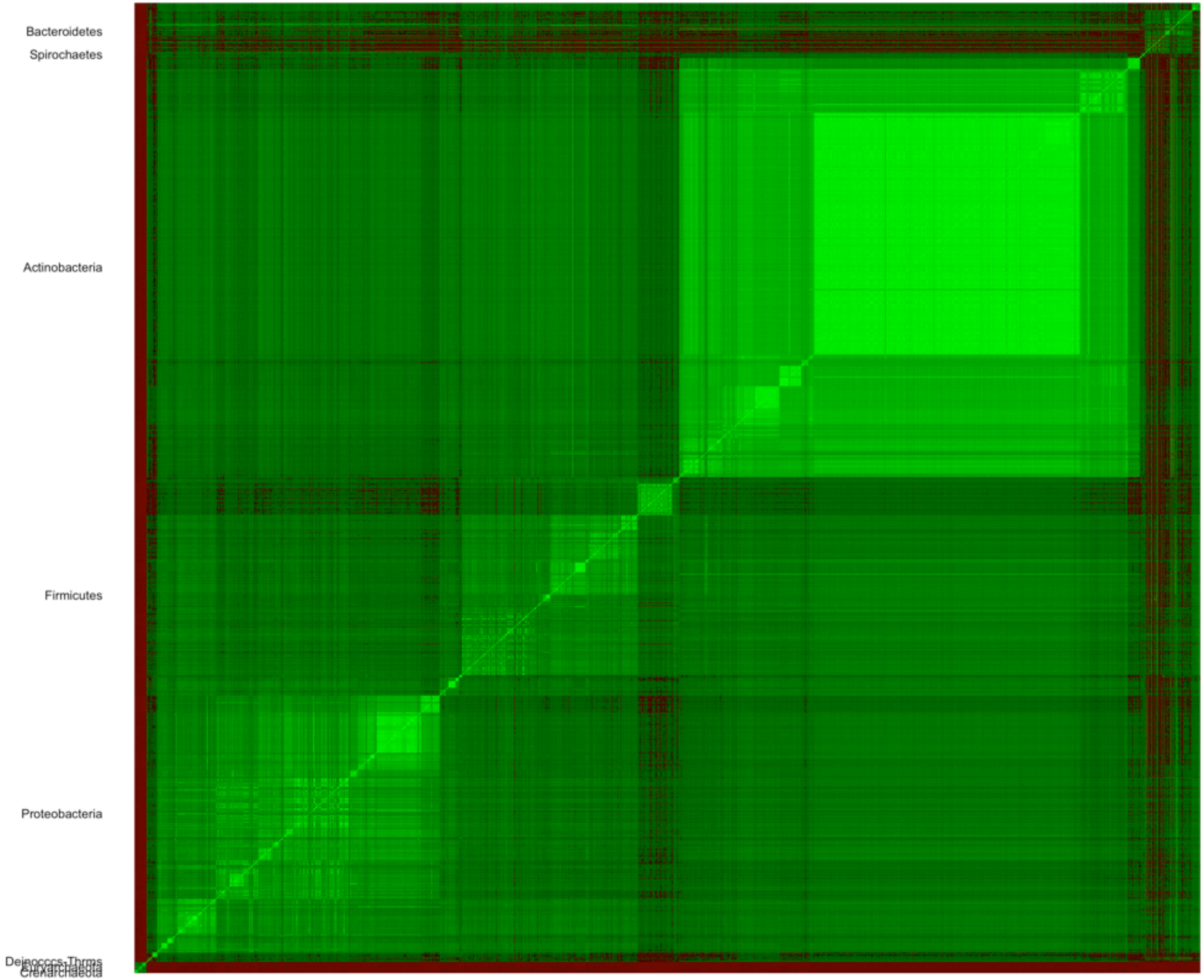
Showing 1 to 44 of 11,952 entries



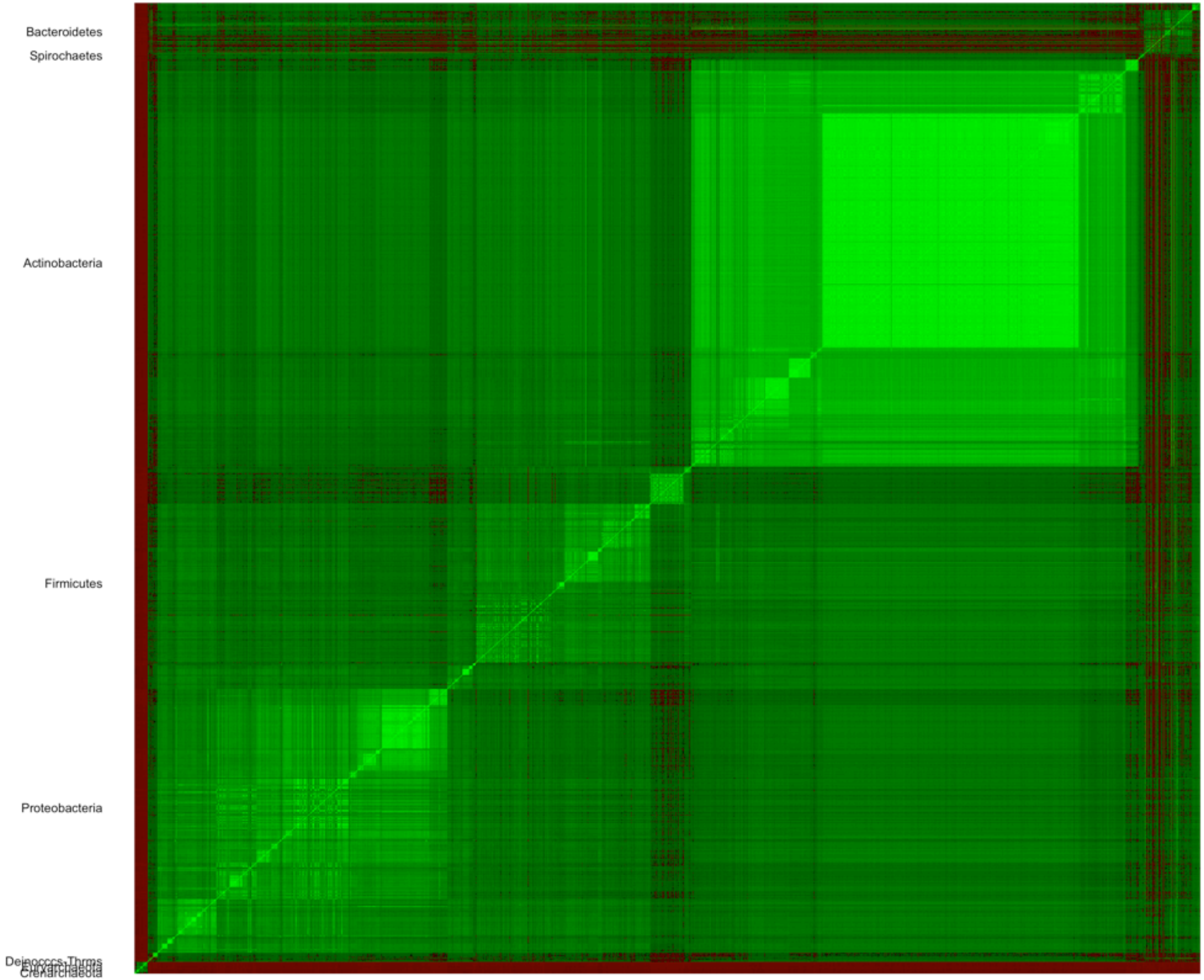


	ex.7	ex.11	ex.13	ex.14	ex.15	ex.16	ex.17
ex.7	0.00000	0.3601	0.066310	0.24430	0.2449	0.046190	0.04949
ex.11	0.36010	0.0000	0.368000	0.60360	0.6137	0.366100	0.36420
ex.13	0.06631	0.3680	0.000000	0.23860	0.2367	0.009129	0.04490
ex.14	0.24430	0.6036	0.238600	0.00000	0.2821	0.241400	0.24280
ex.15	0.24490	0.6137	0.236700	0.28210	0.0000	0.233400	0.22470
ex.16	0.04619	0.3661	0.009129	0.24140	0.2334	0.000000	0.02449
ex.17	0.04949	0.3642	0.044900	0.24280	0.2247	0.024490	0.00000

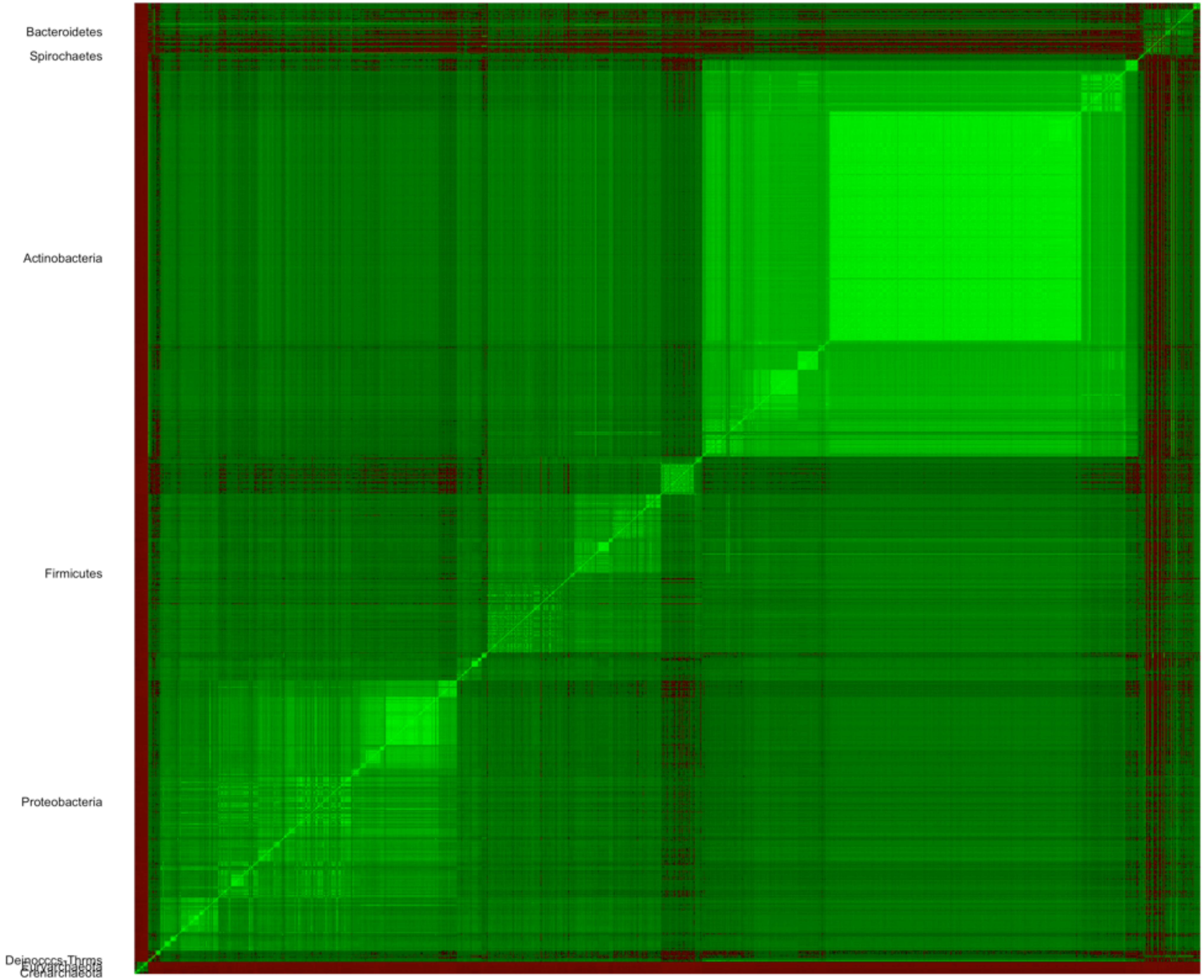
16S rRNA similarity of type strains of Bacteria and Archaea 1980



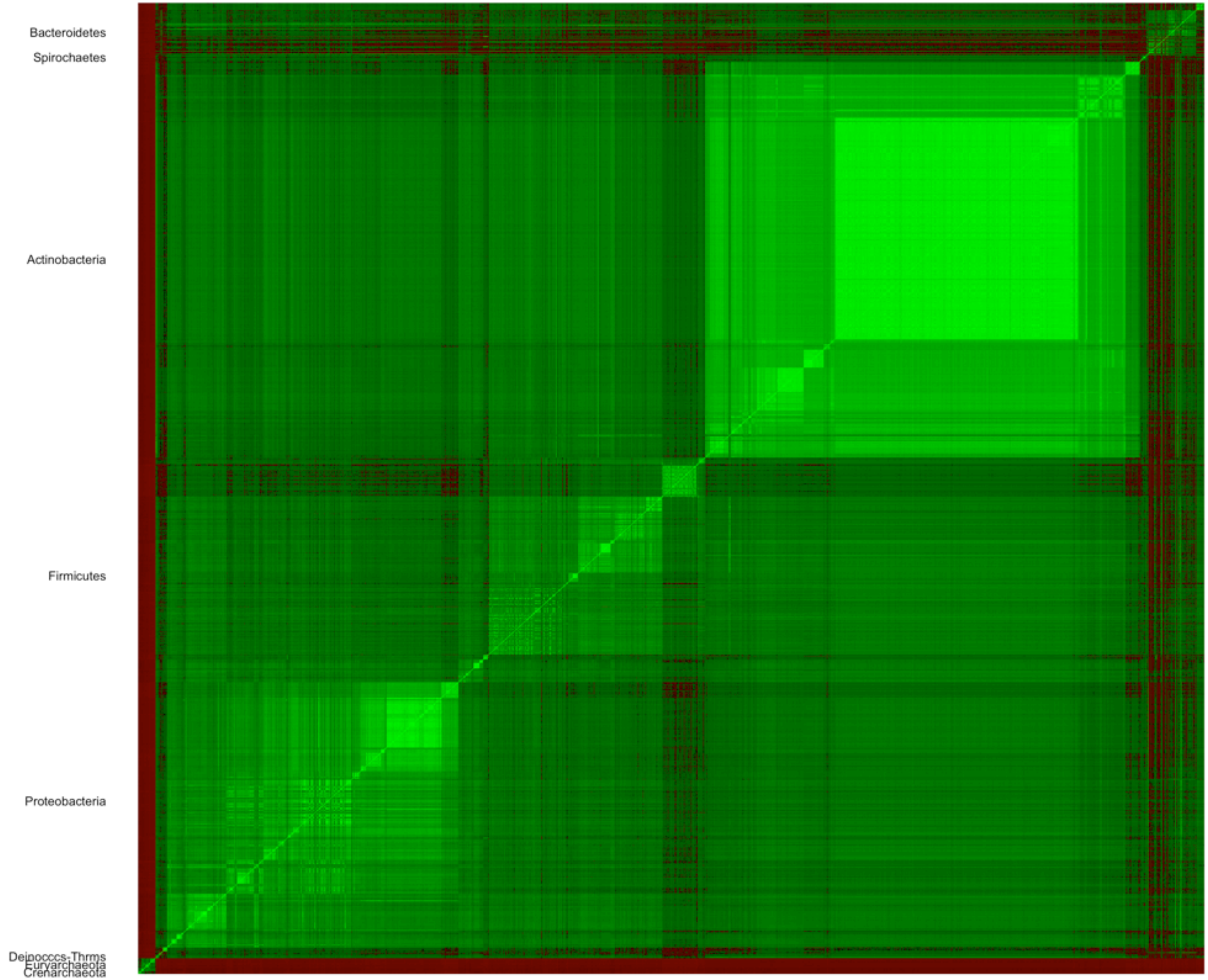
16S rRNA similarity of type strains of Bacteria and Archaea 1981



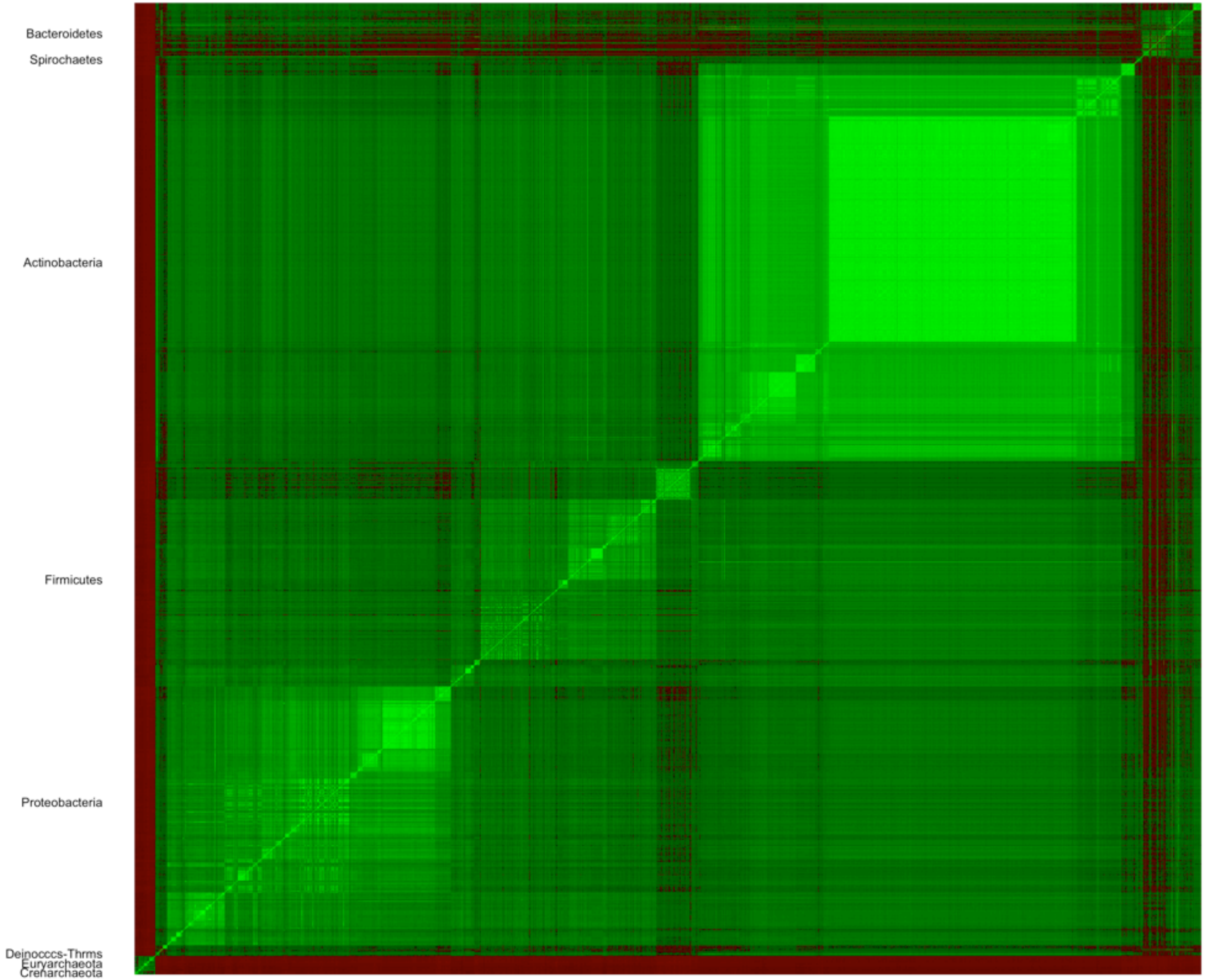
16S rRNA similarity of type strains of Bacteria and Archaea 1982



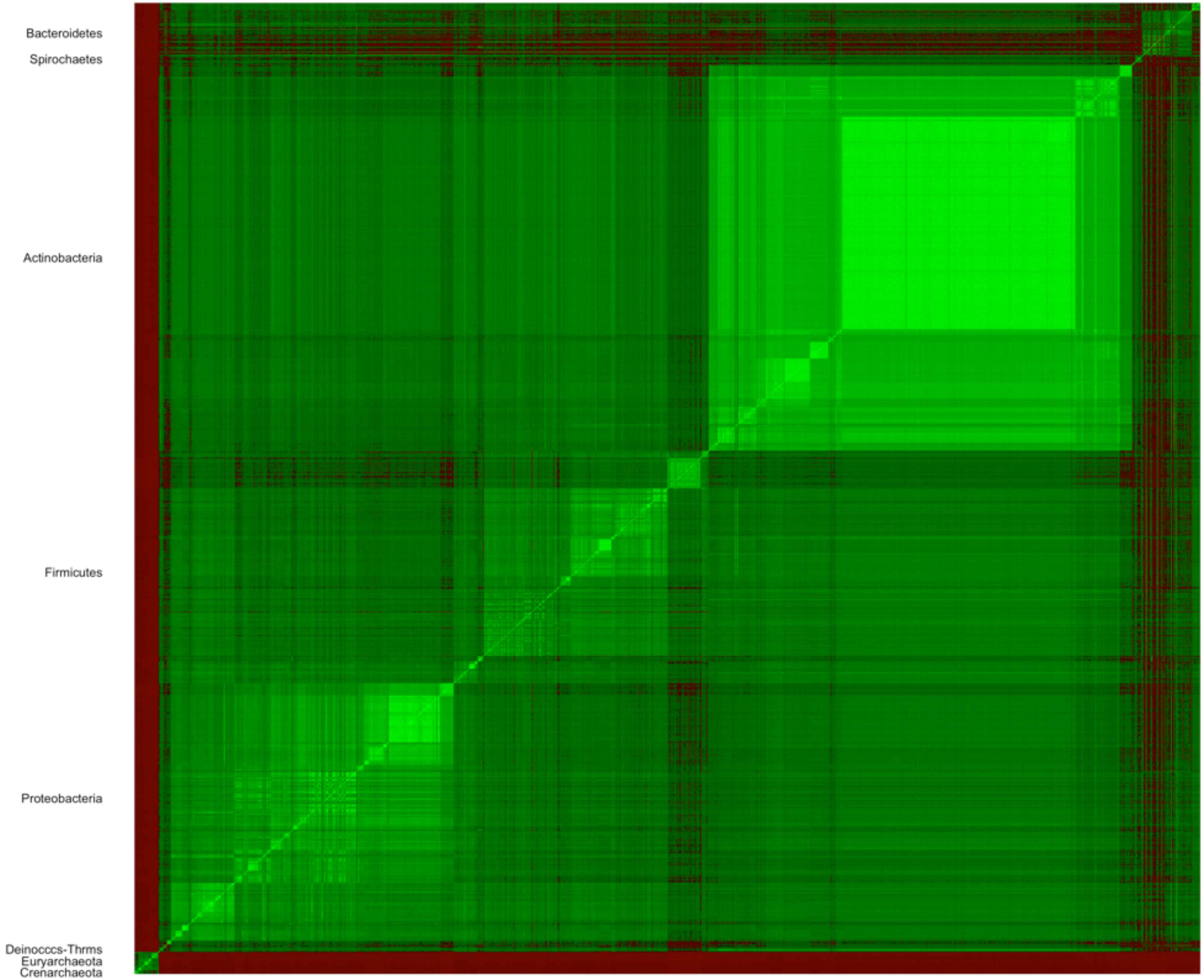
16S rRNA similarity of type strains of Bacteria and Archaea 1983



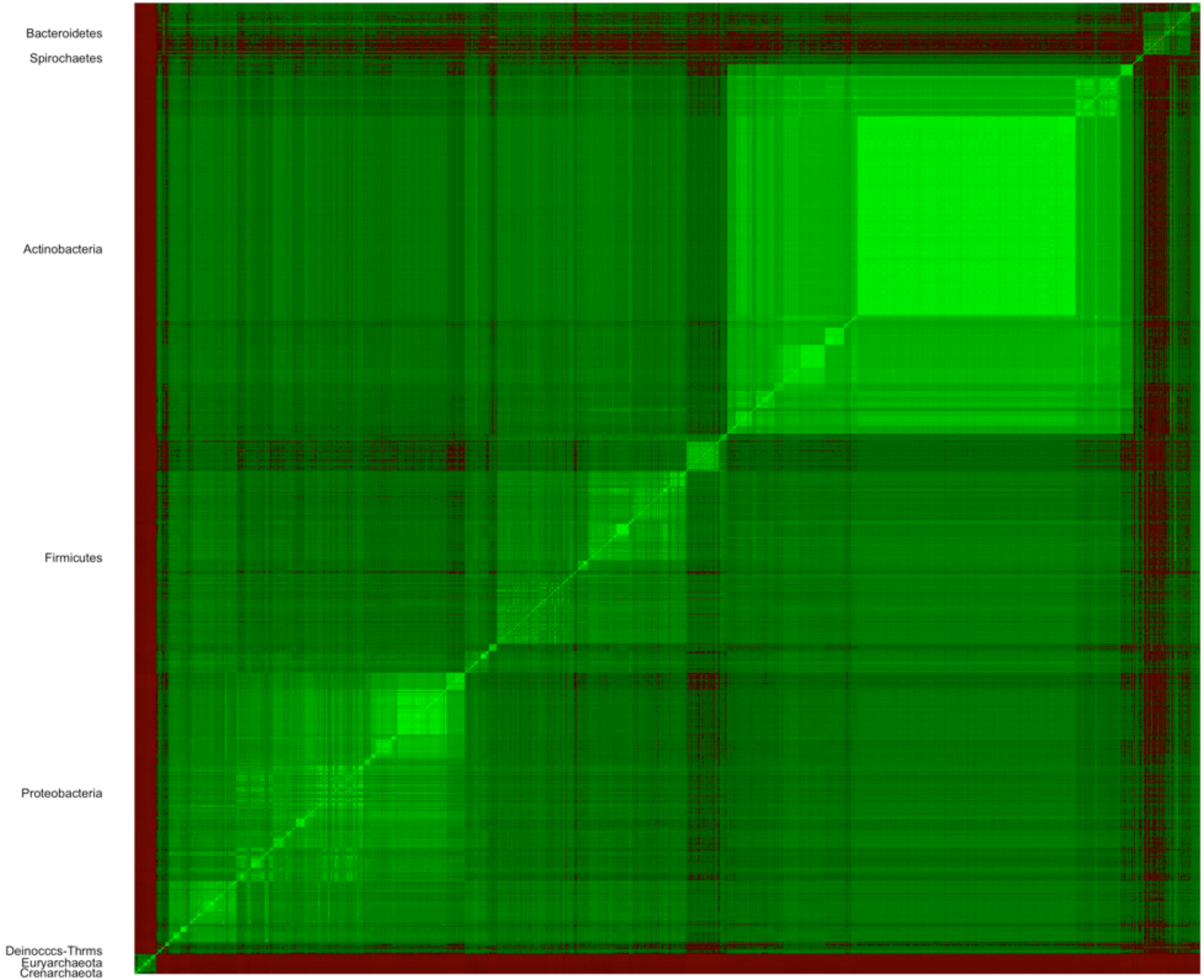
16S rRNA similarity of type strains of Bacteria and Archaea 1984



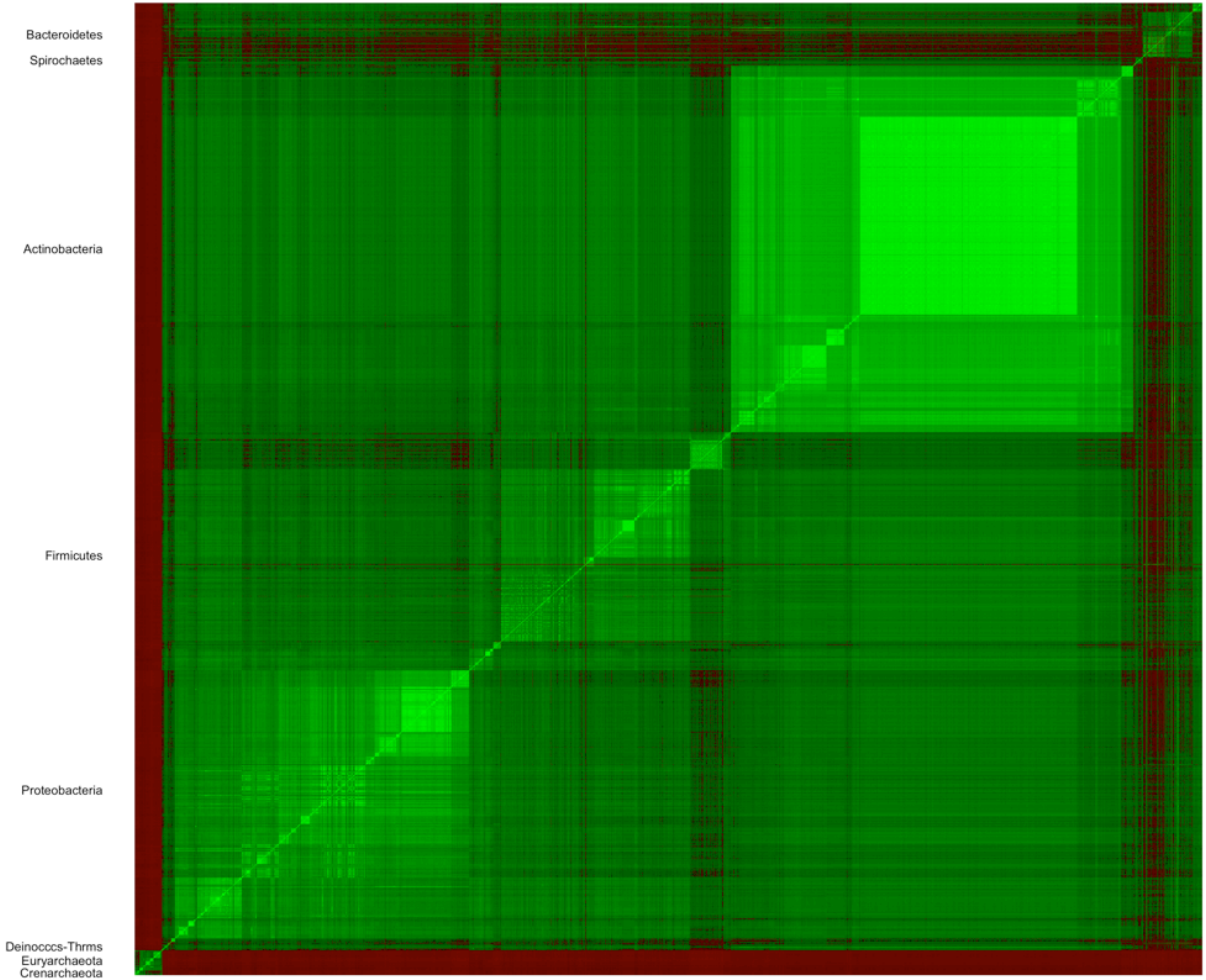
16S rRNA similarity of type strains of Bacteria and Archaea 1985



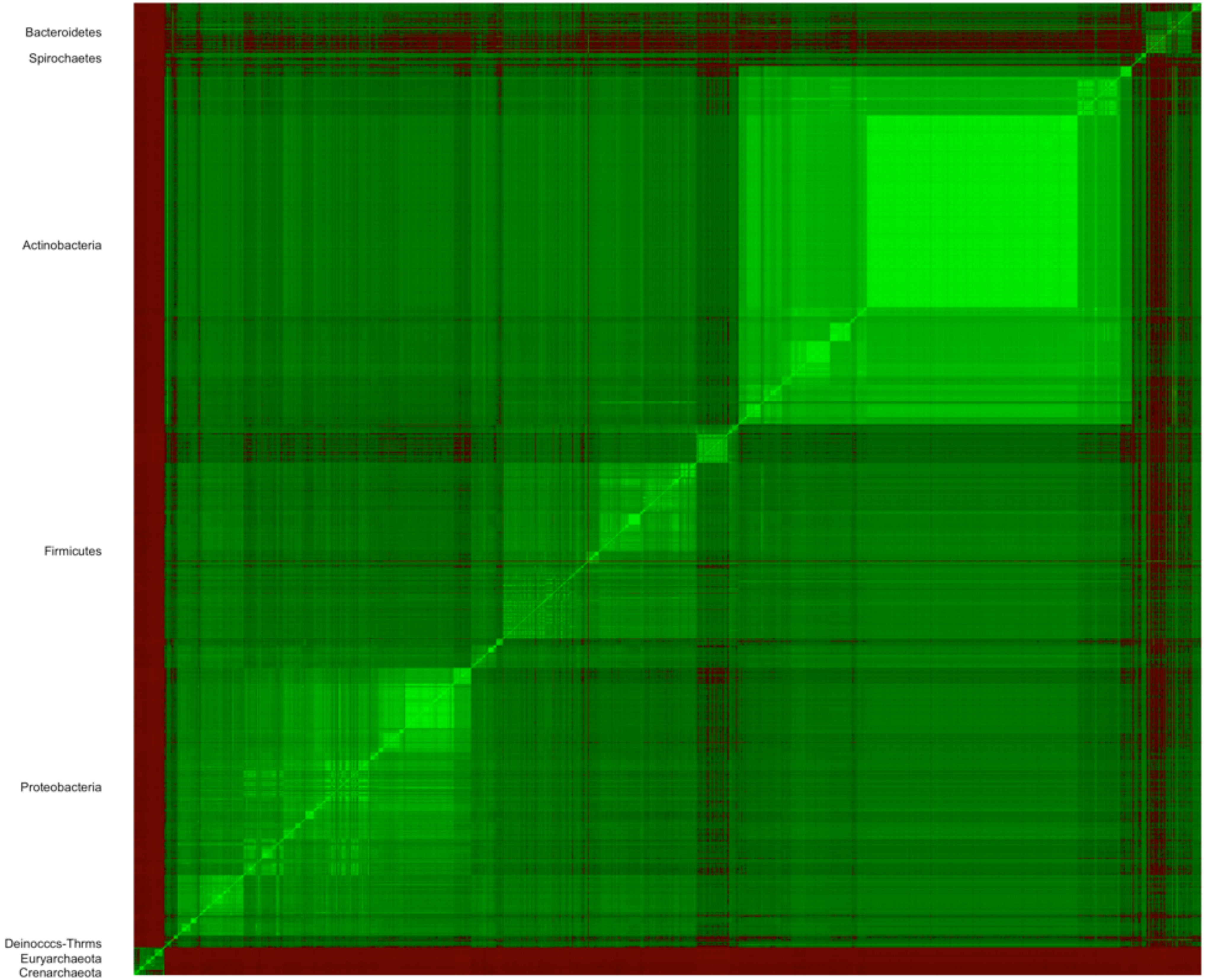
16S rRNA similarity of type strains of Bacteria and Archaea 1986



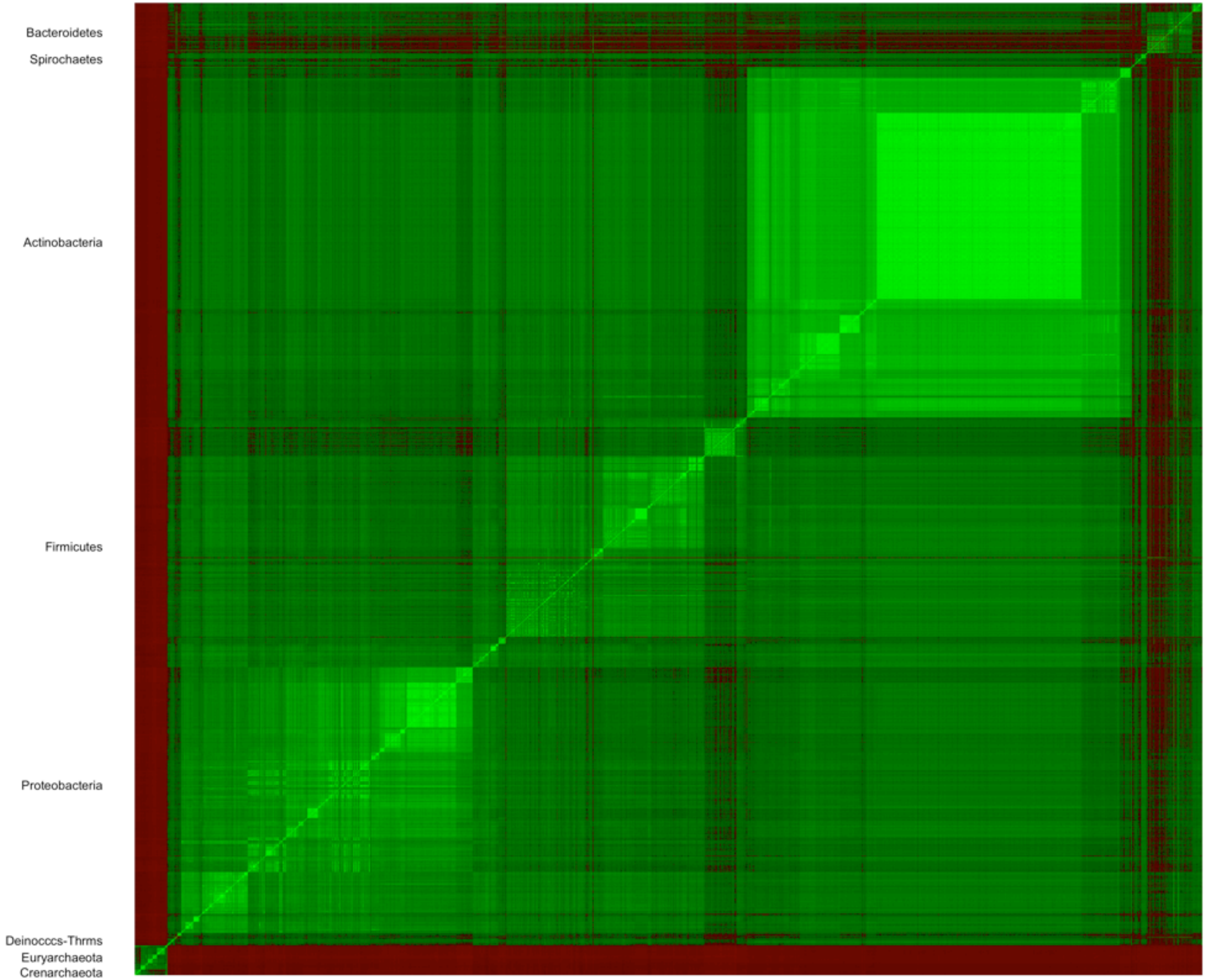
16S rRNA similarity of type strains of Bacteria and Archaea 1987



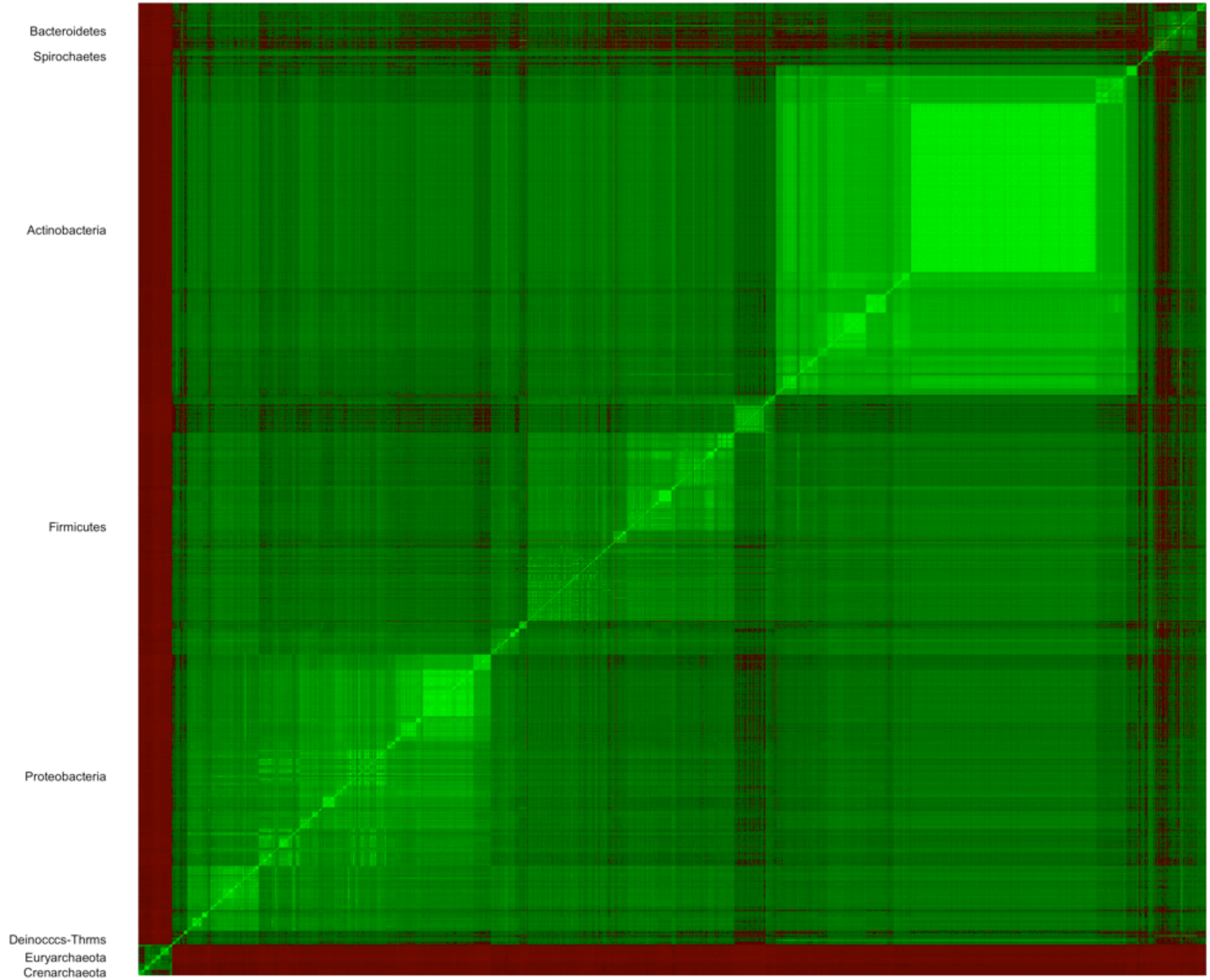
16S rRNA similarity of type strains of Bacteria and Archaea 1988



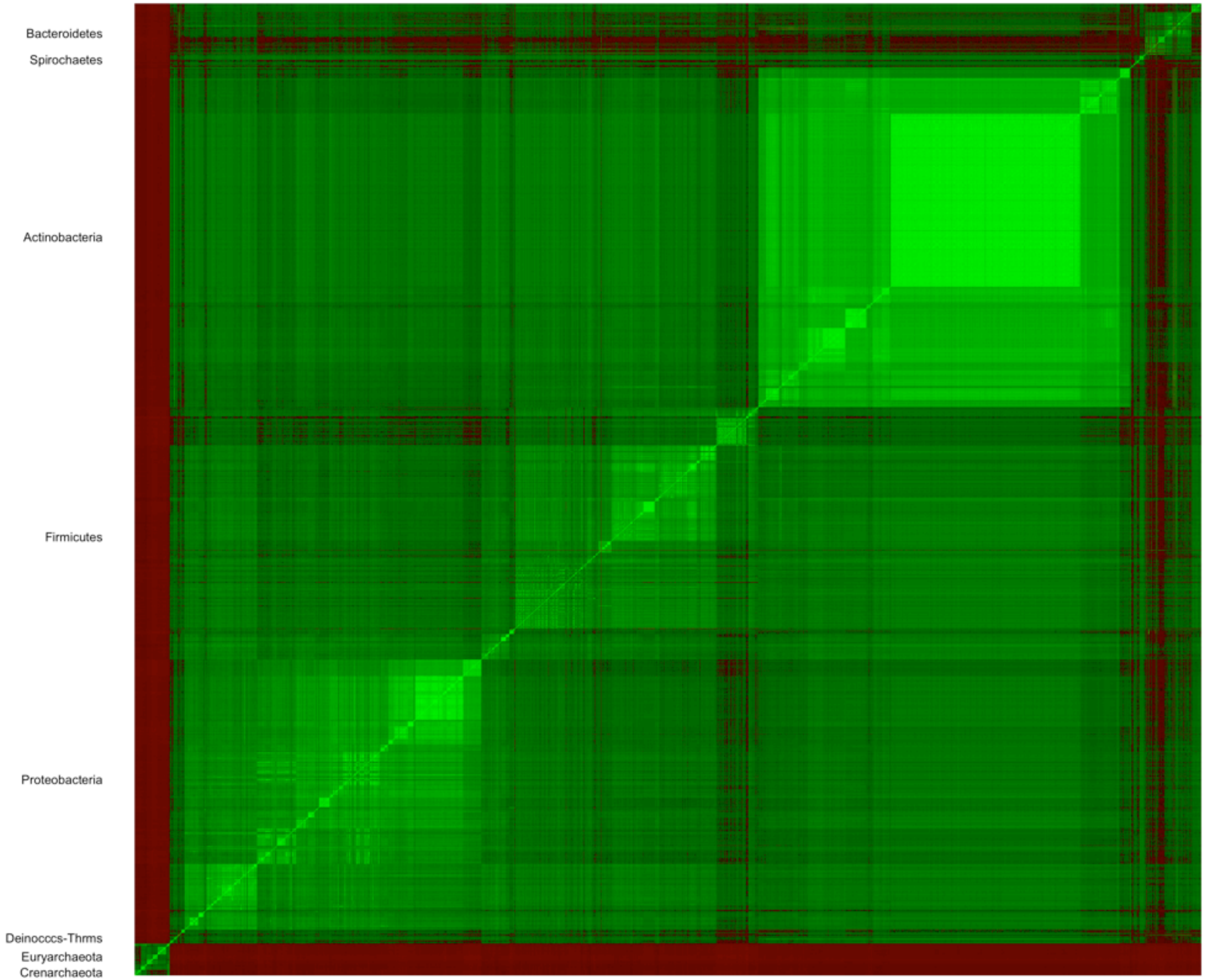
16S rRNA similarity of type strains of Bacteria and Archaea 1989



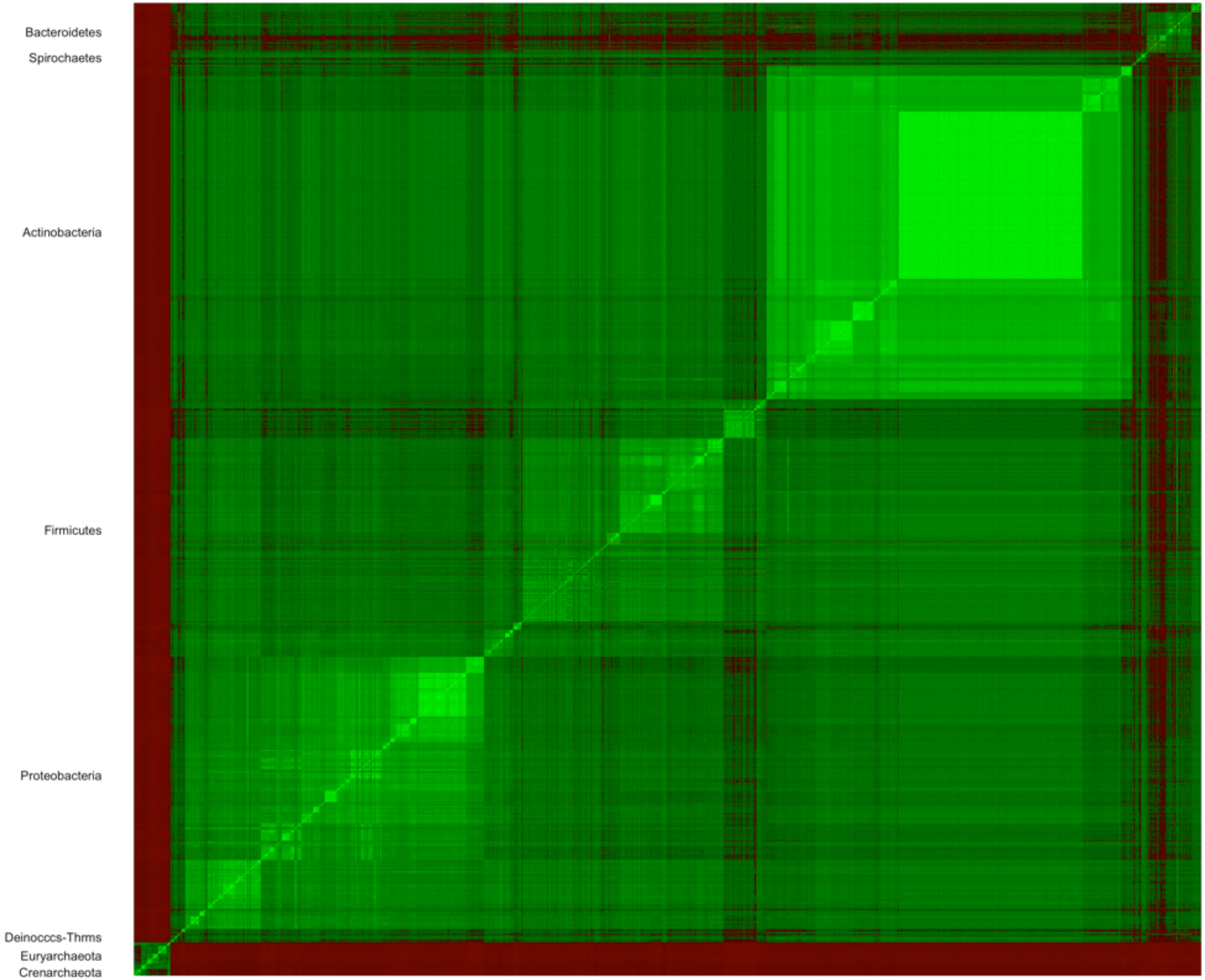
16S rRNA similarity of type strains of Bacteria and Archaea 1990



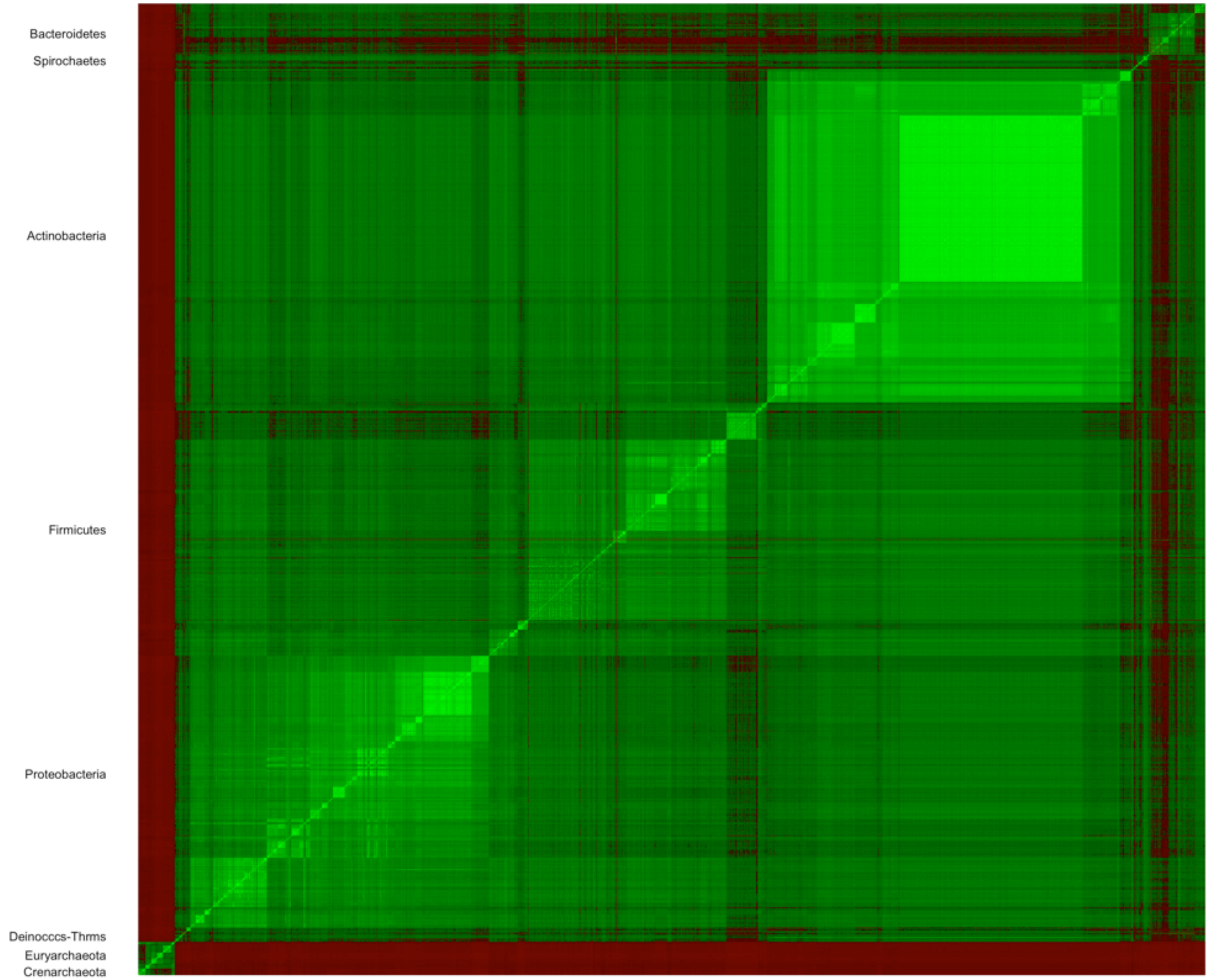
16S rRNA similarity of type strains of Bacteria and Archaea 1991



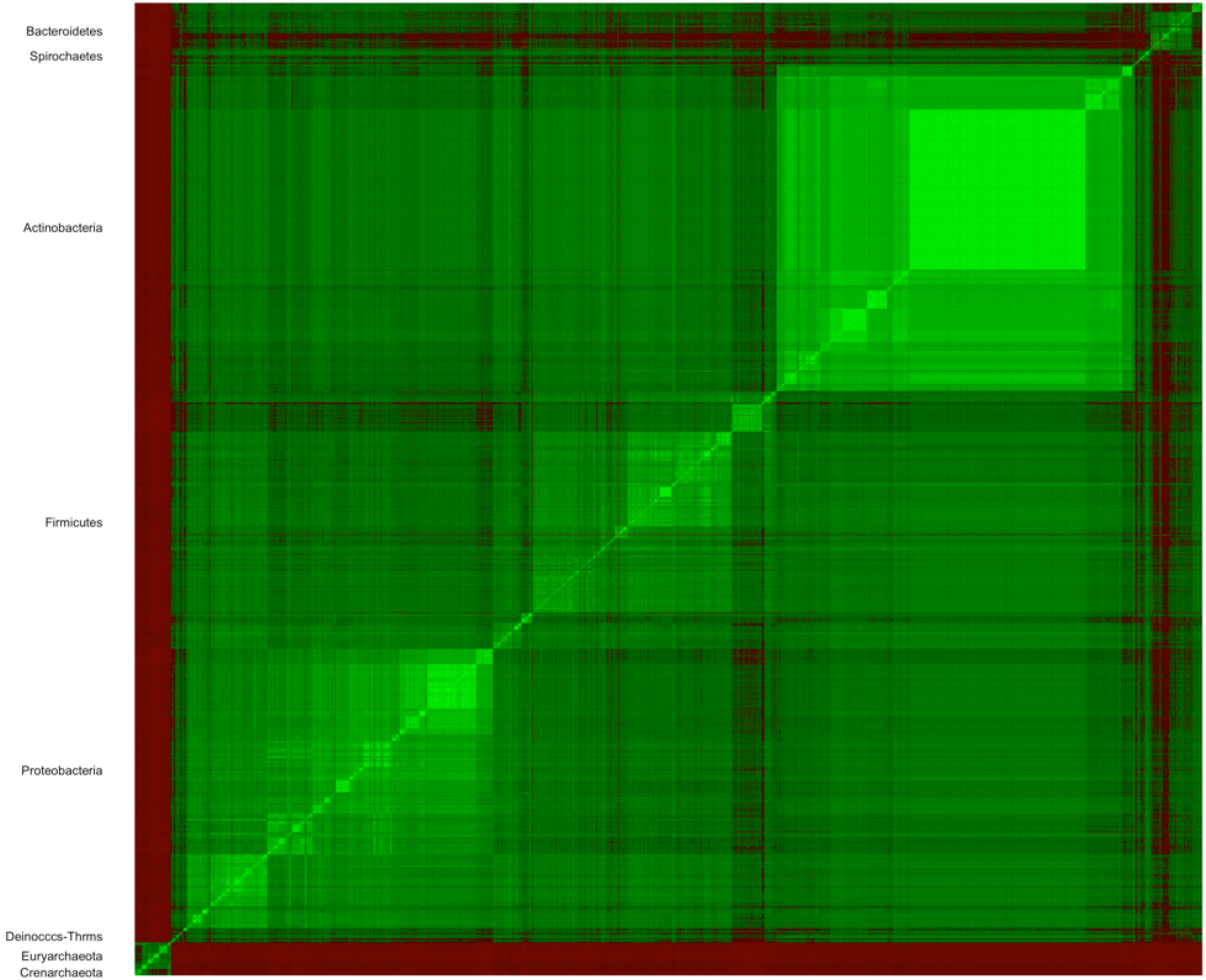
16S rRNA similarity of type strains of Bacteria and Archaea 1992



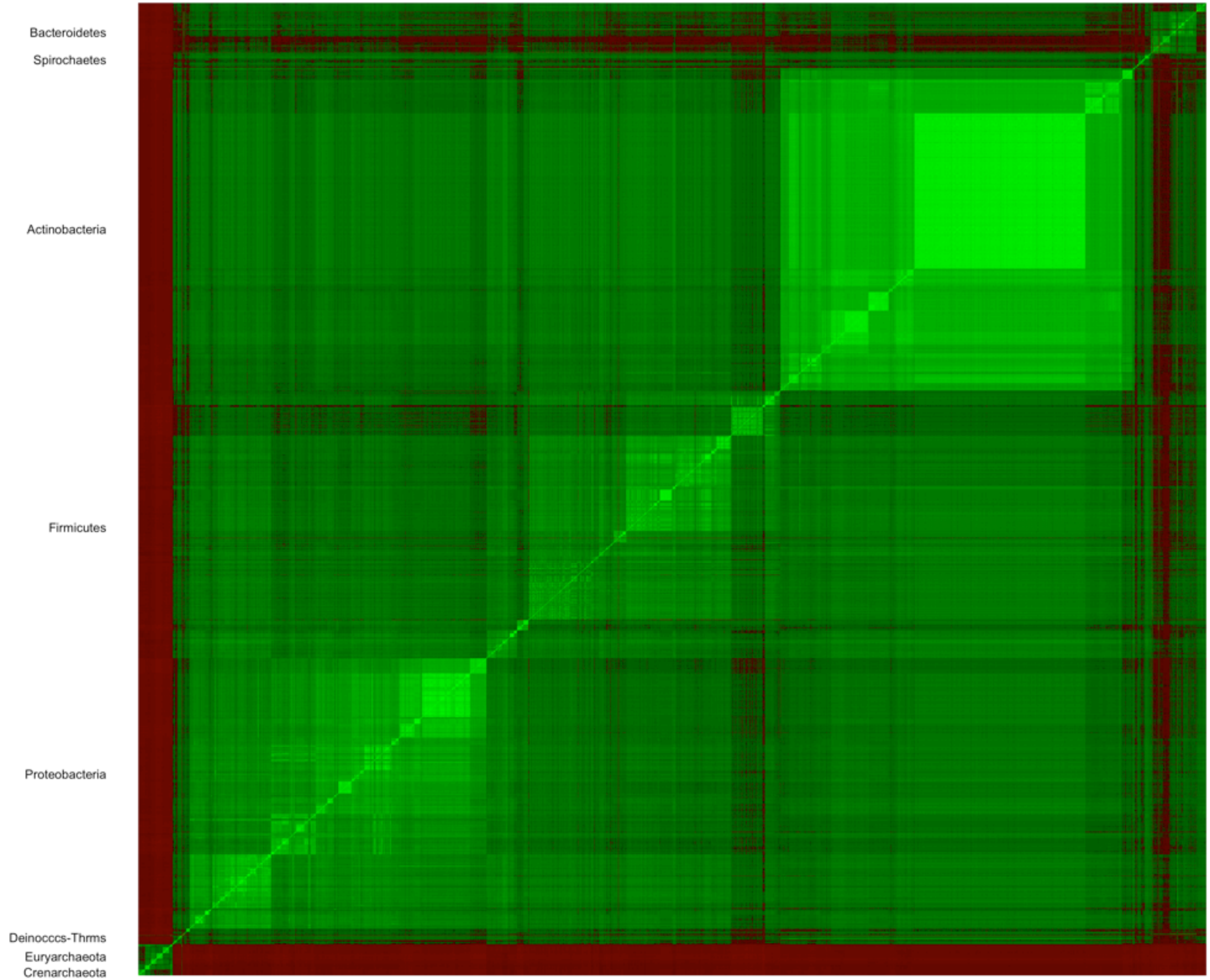
16S rRNA similarity of type strains of Bacteria and Archaea 1993



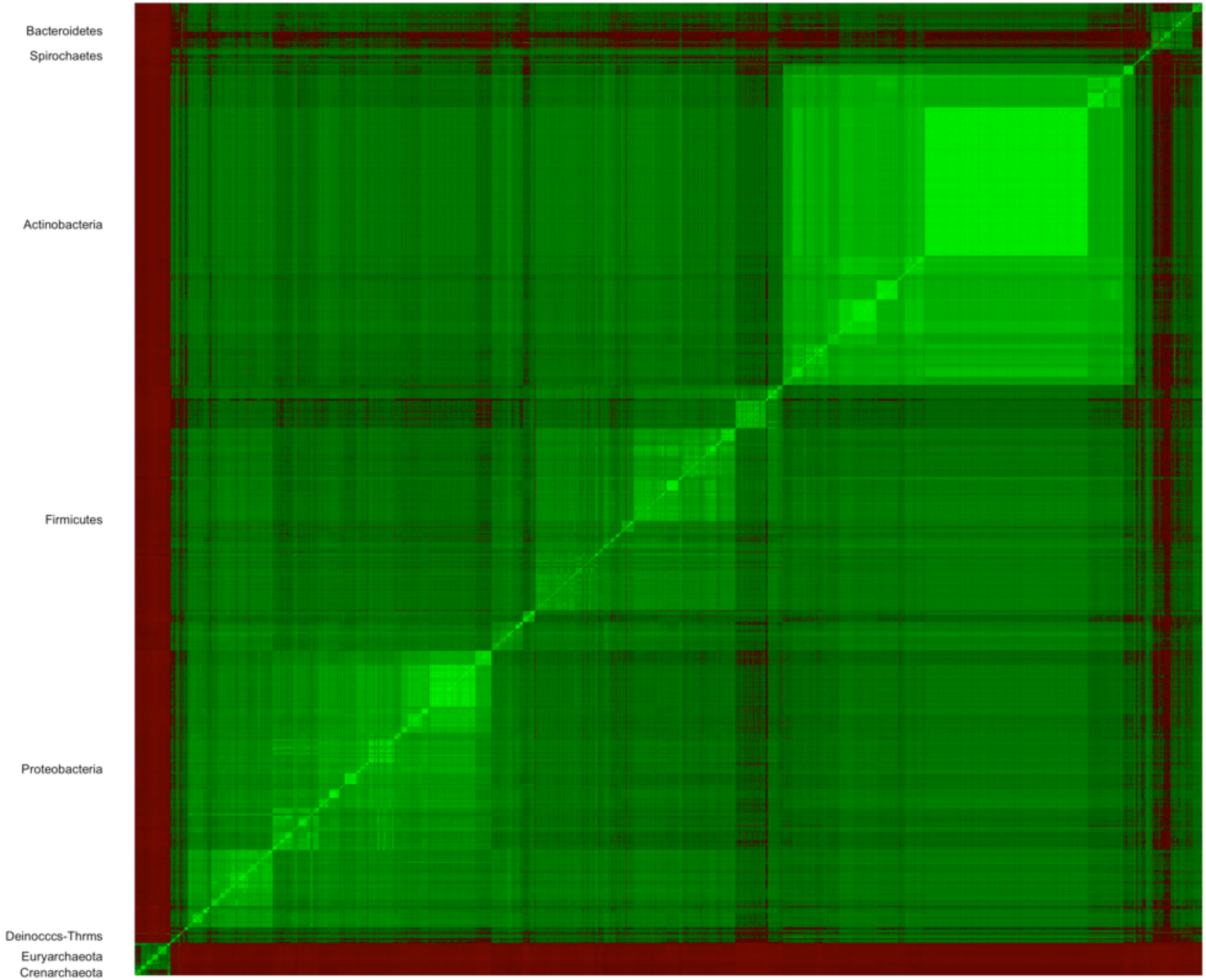
16S rRNA similarity of type strains of Bacteria and Archaea 1994



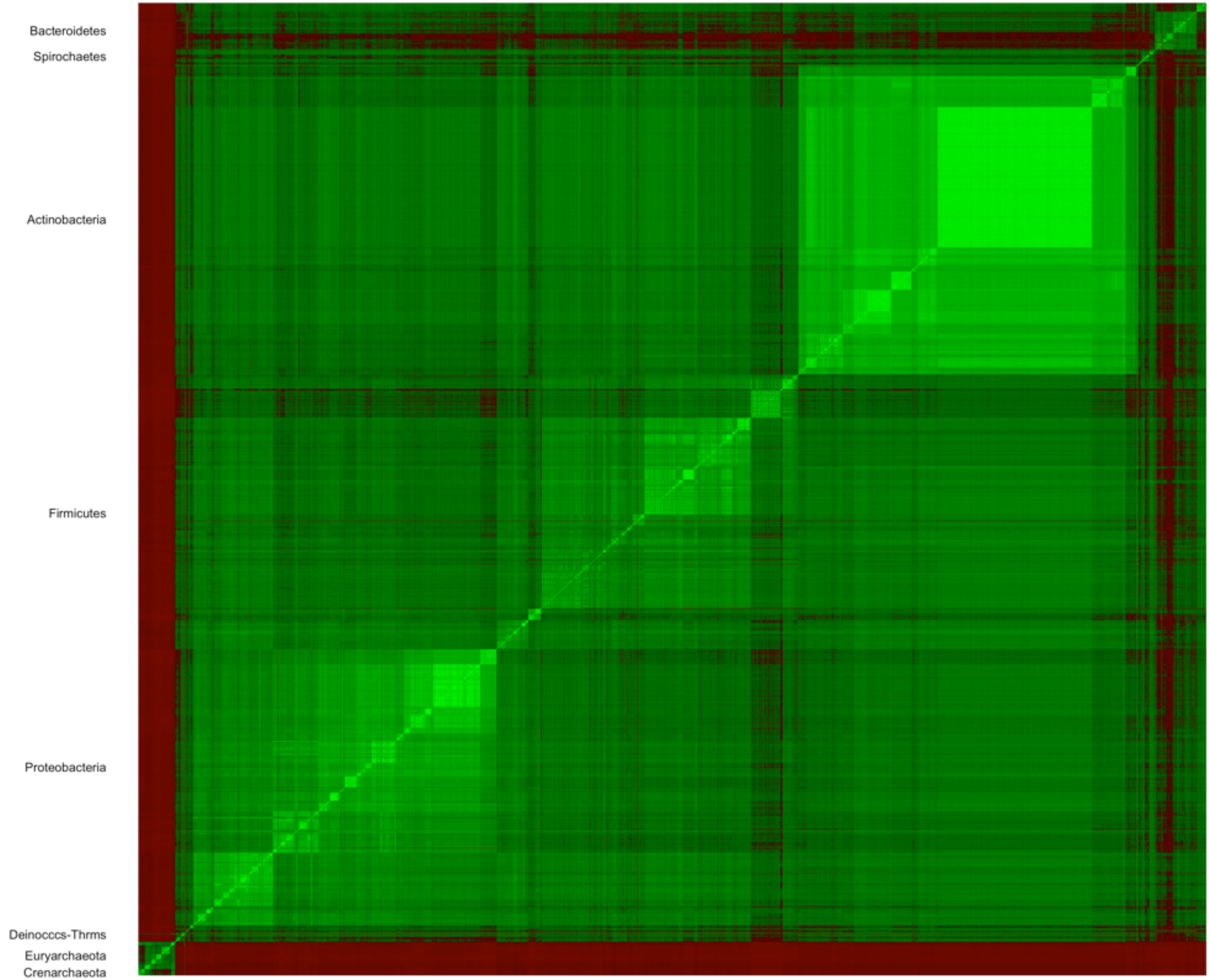
16S rRNA similarity of type strains of Bacteria and Archaea 1995



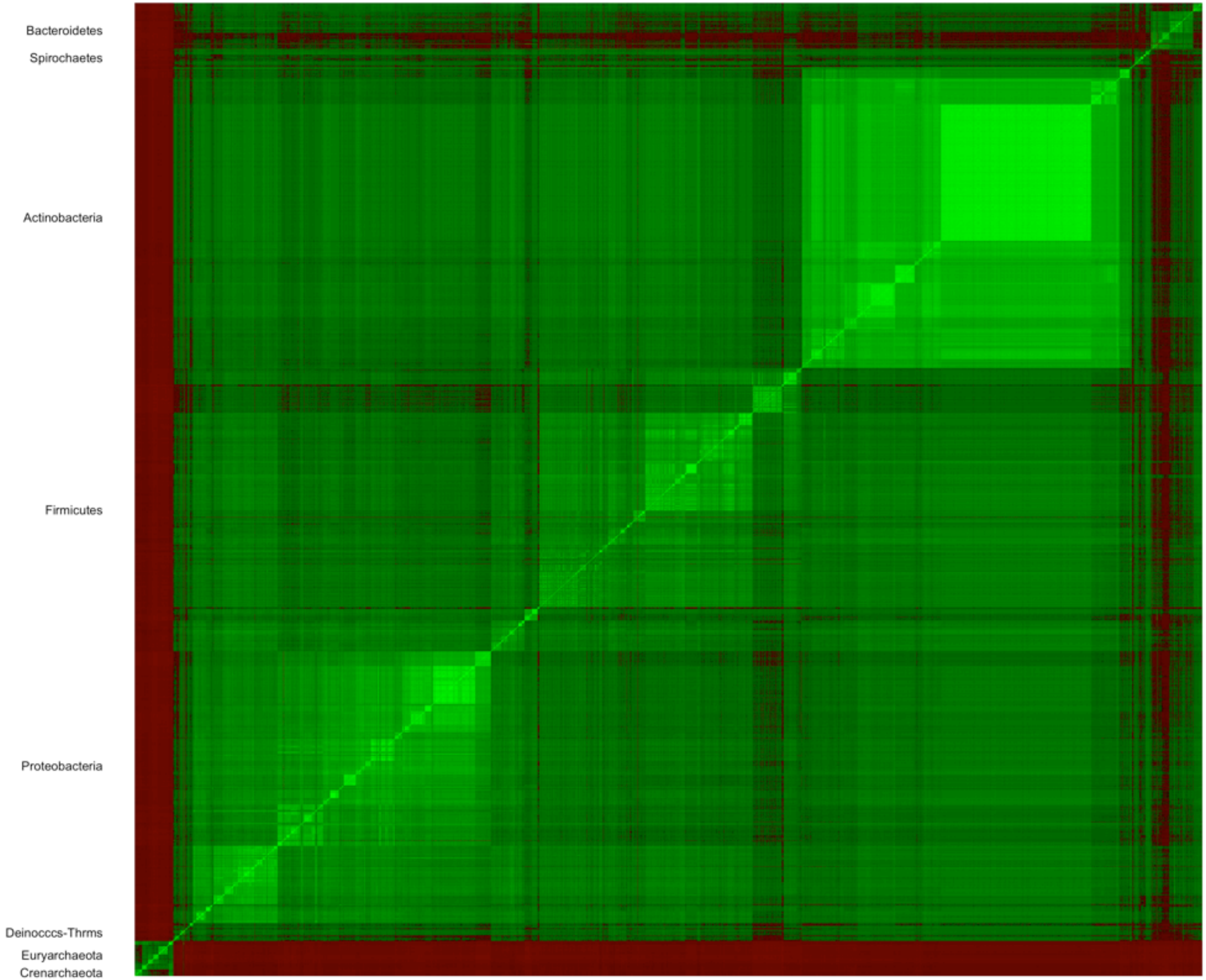
16S rRNA similarity of type strains of Bacteria and Archaea 1996



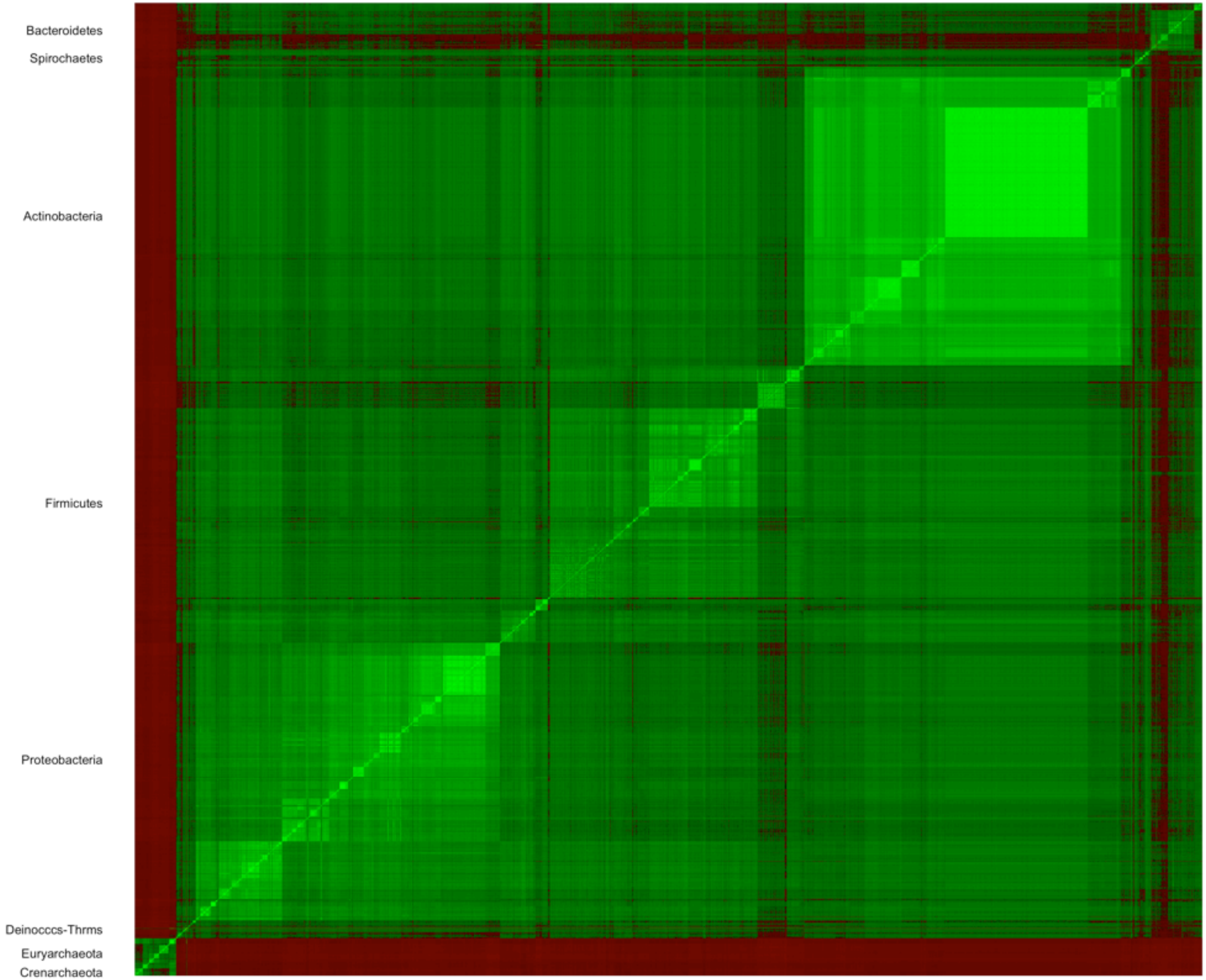
16S rRNA similarity of type strains of Bacteria and Archaea 1997



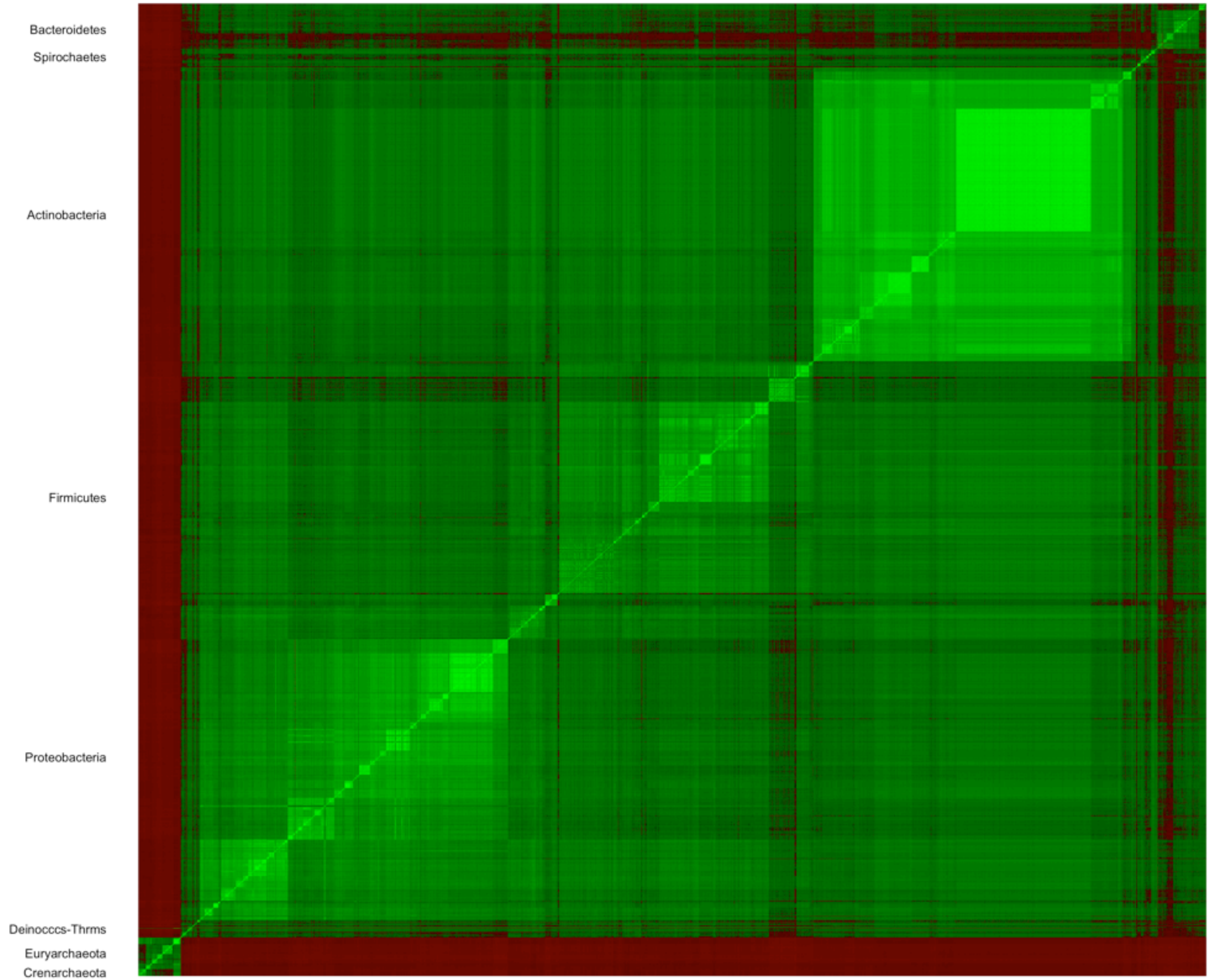
16S rRNA similarity of type strains of Bacteria and Archaea 1998



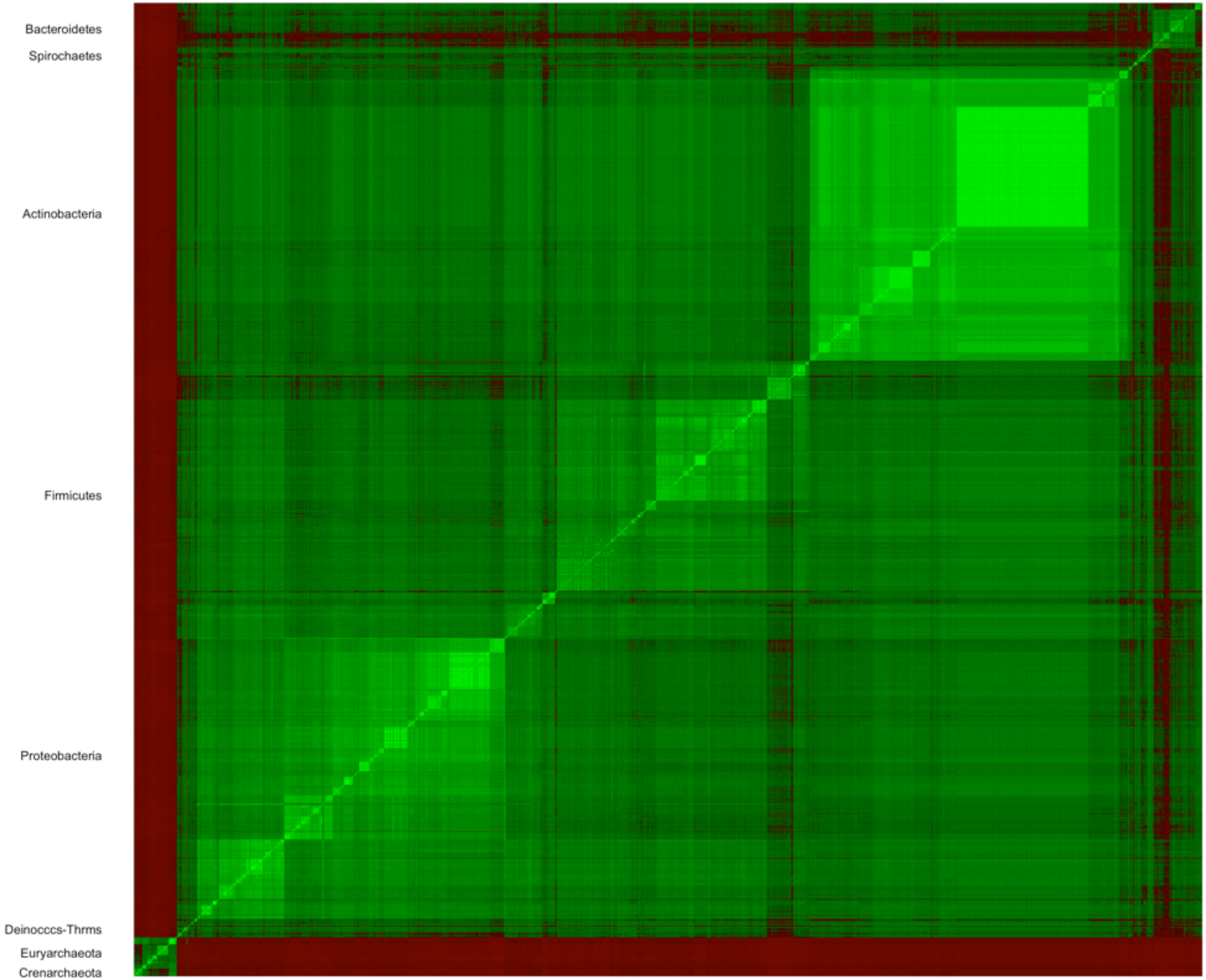
16S rRNA similarity of type strains of Bacteria and Archaea 1999



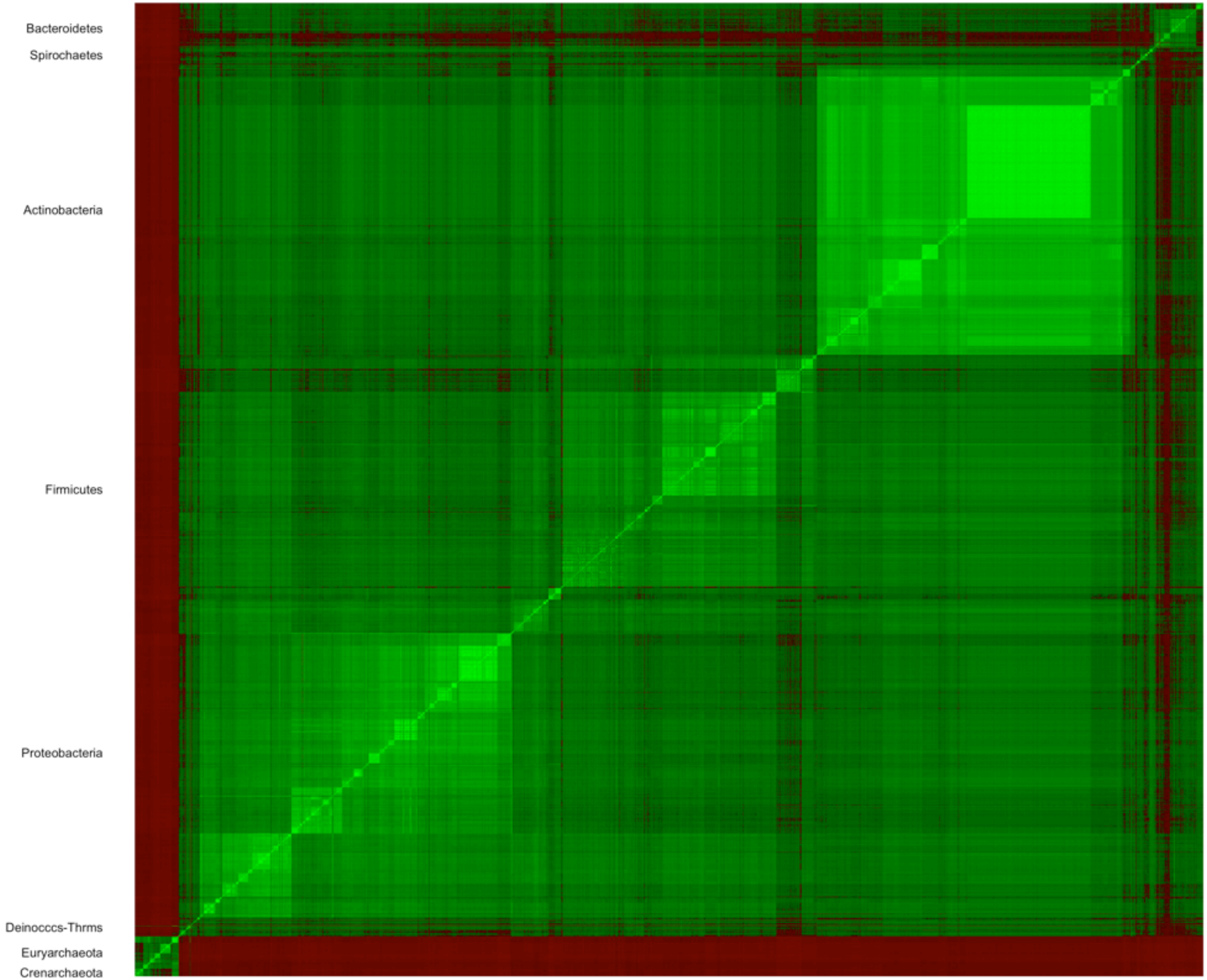
16S rRNA similarity of type strains of Bacteria and Archaea 2000



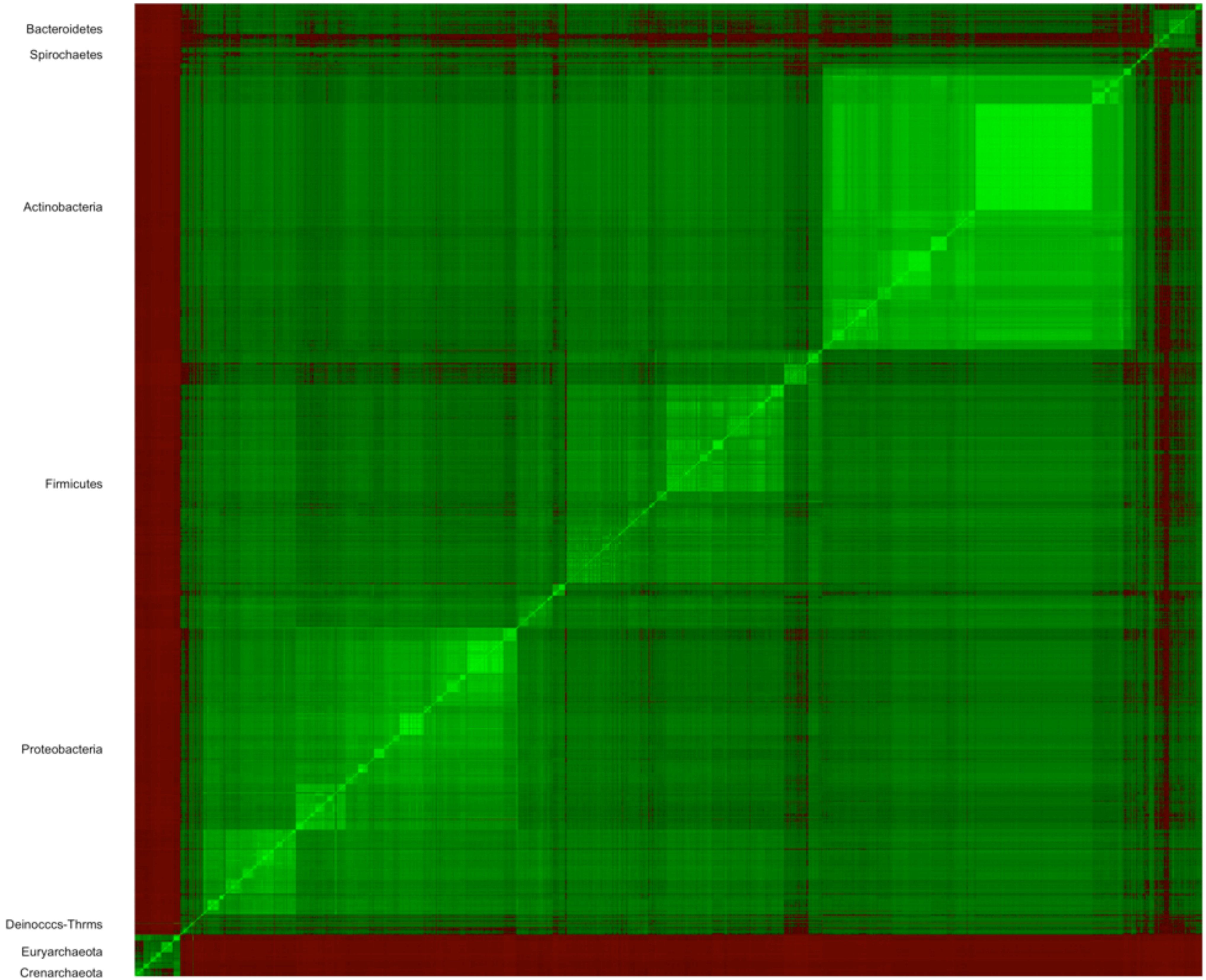
16S rRNA similarity of type strains of Bacteria and Archaea 2001



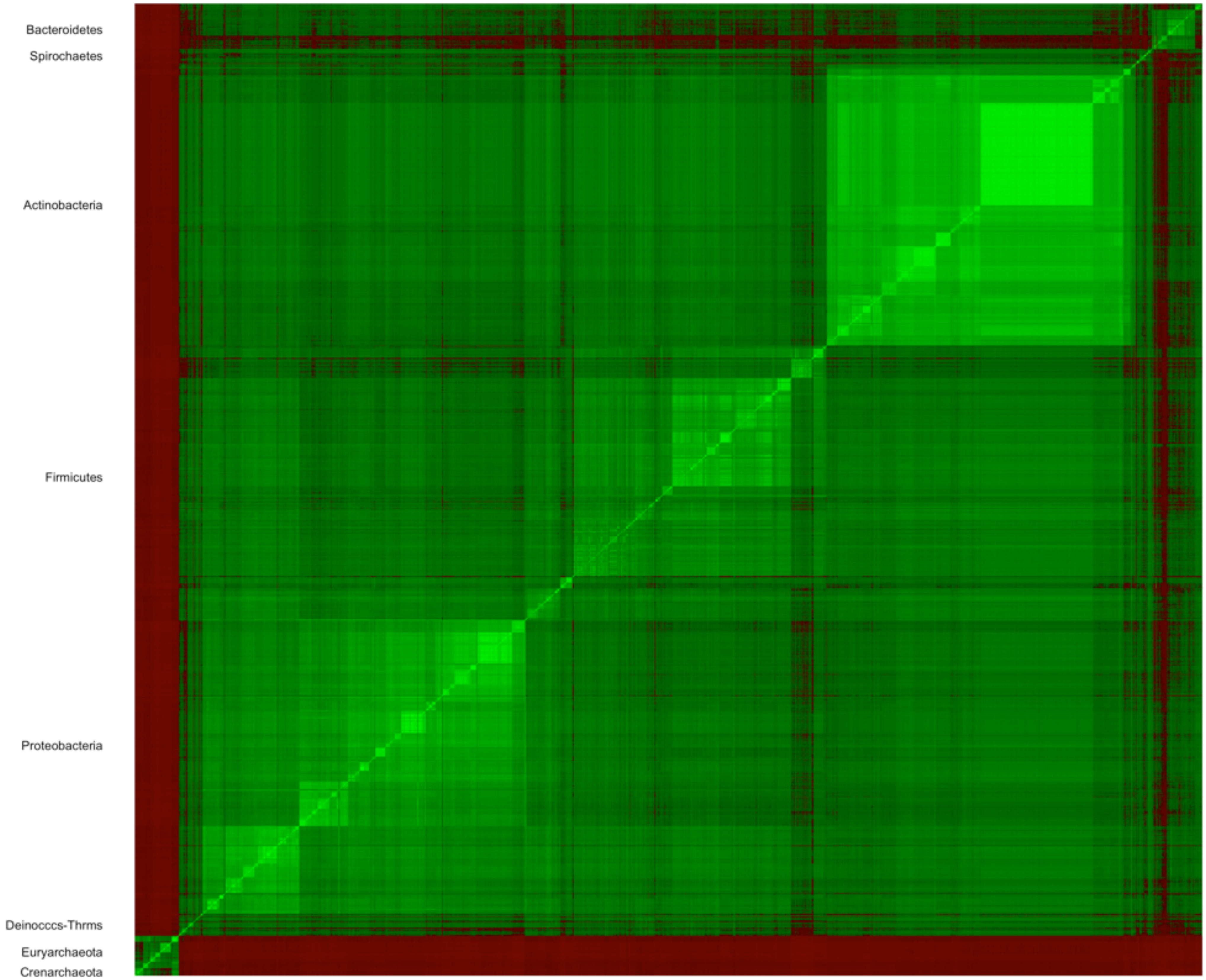
16S rRNA similarity of type strains of Bacteria and Archaea 2002



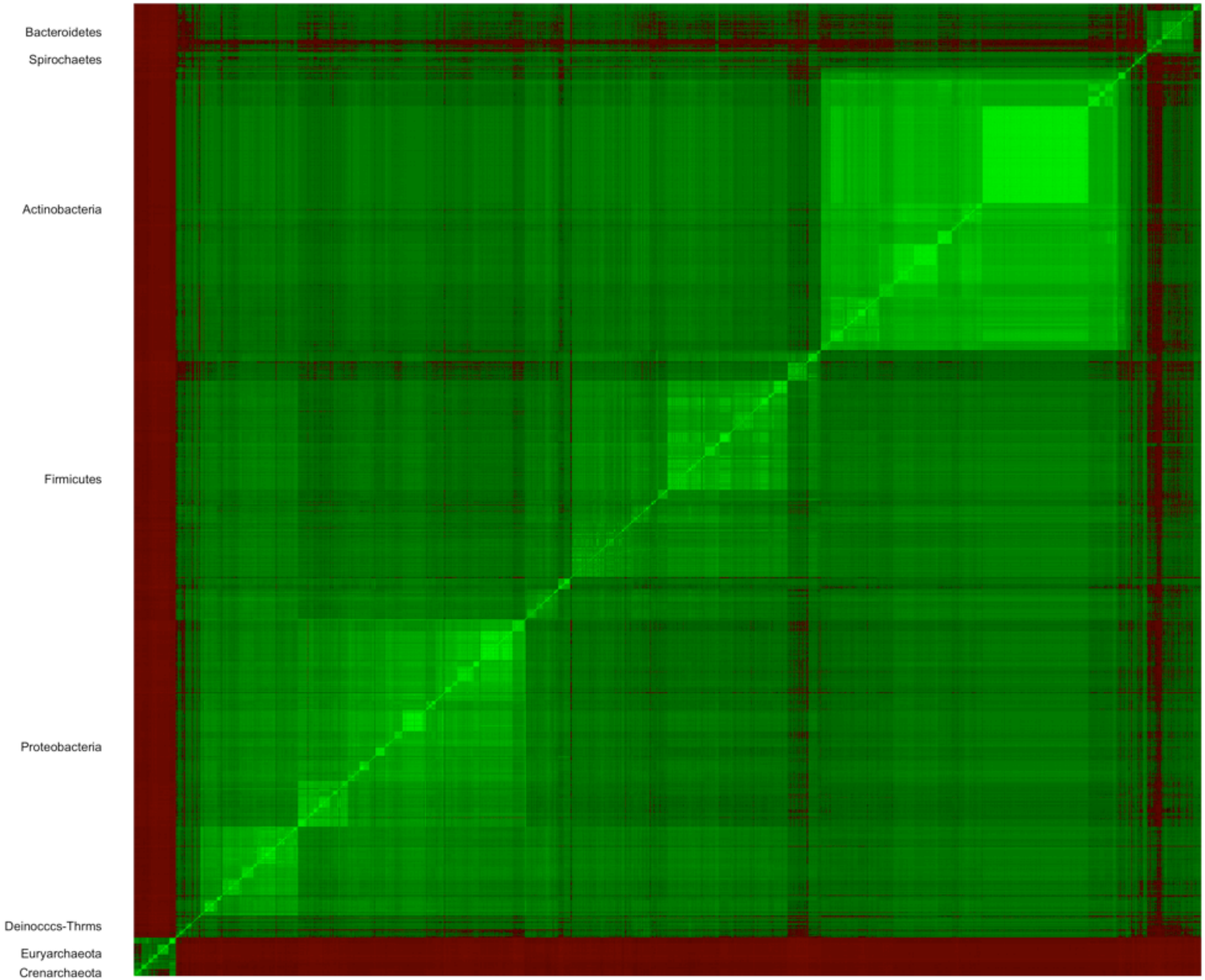
16S rRNA similarity of type strains of Bacteria and Archaea 2003



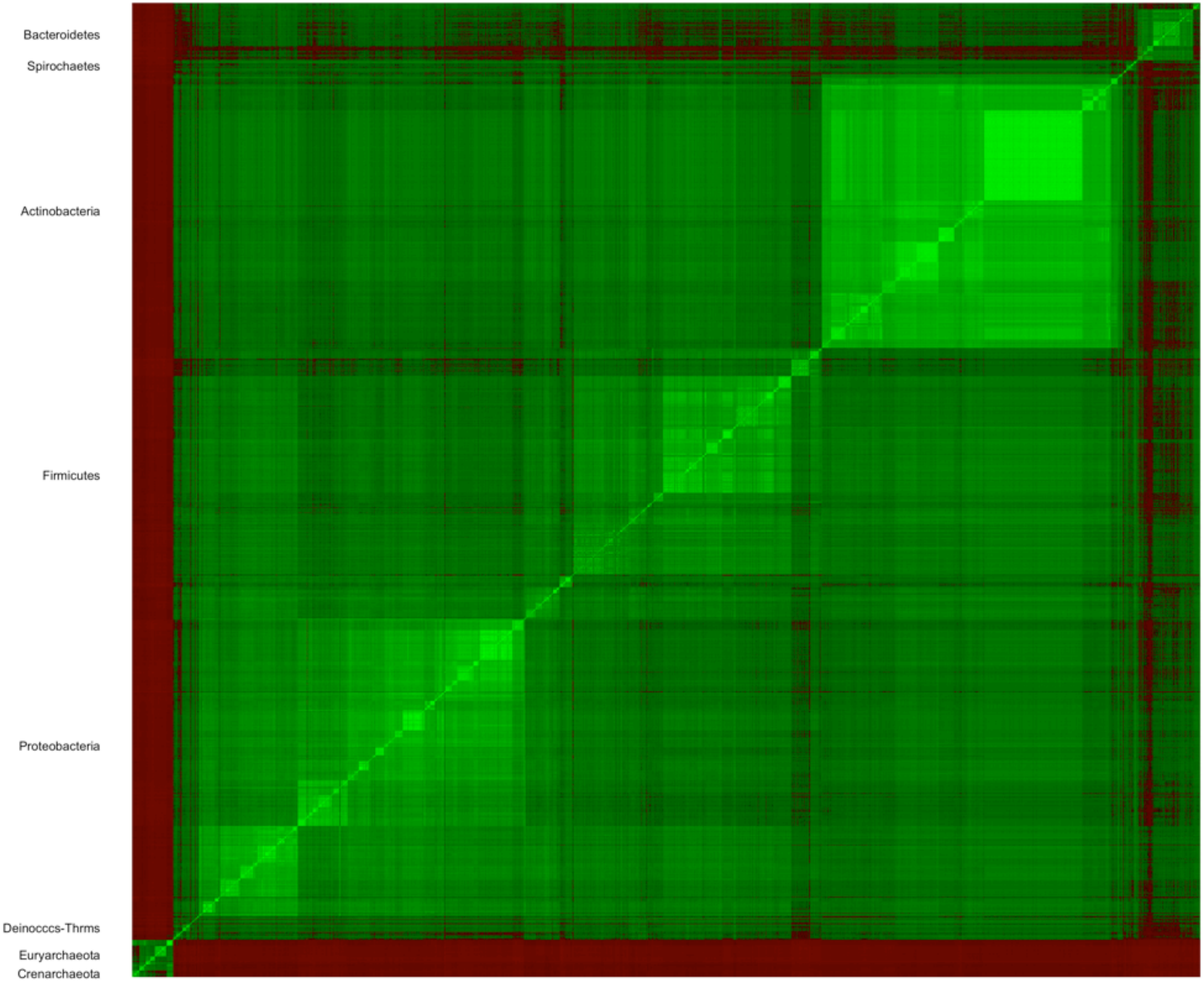
16S rRNA similarity of type strains of Bacteria and Archaea 2004



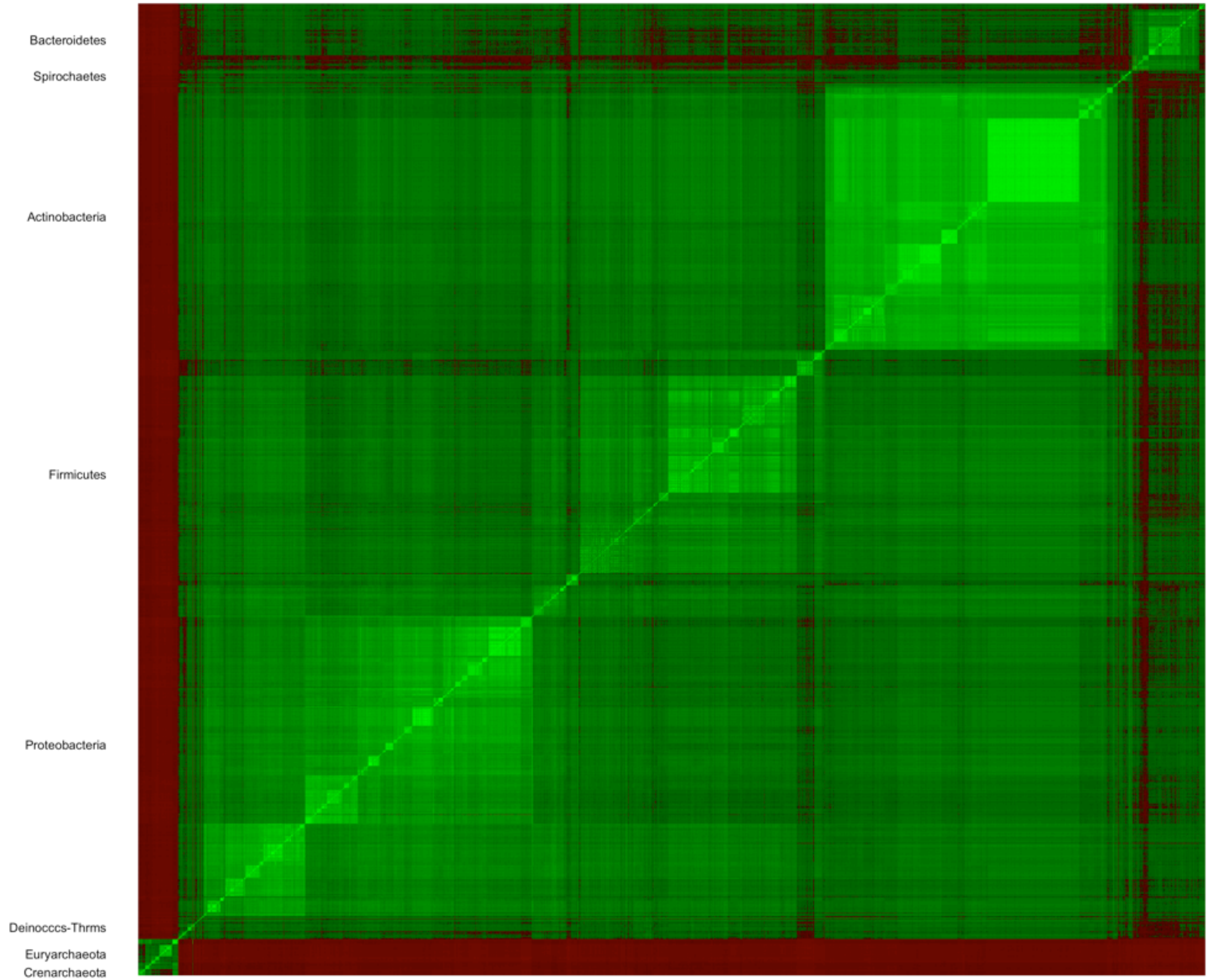
16S rRNA similarity of type strains of Bacteria and Archaea 2005



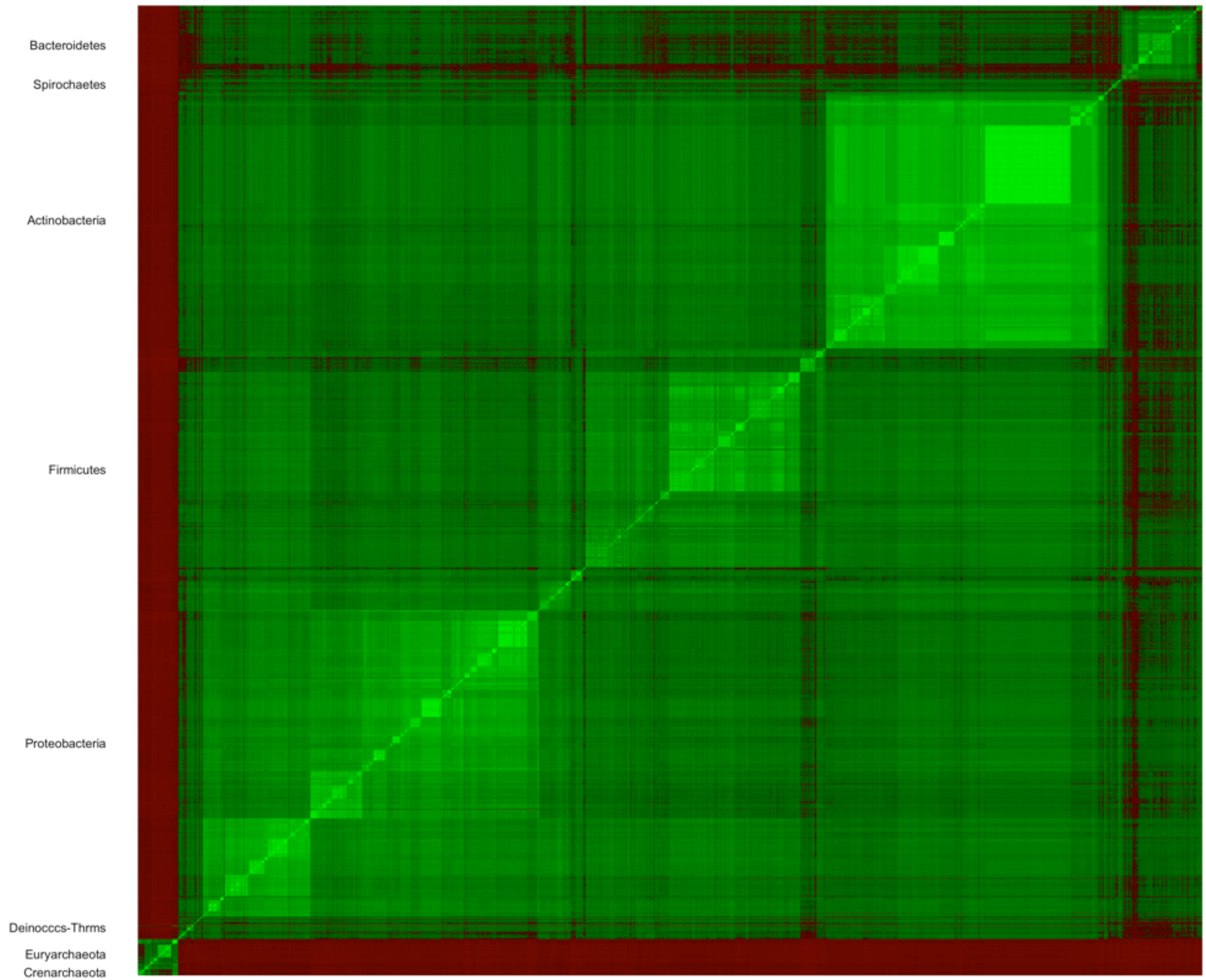
16S rRNA similarity of type strains of Bacteria and Archaea 2006



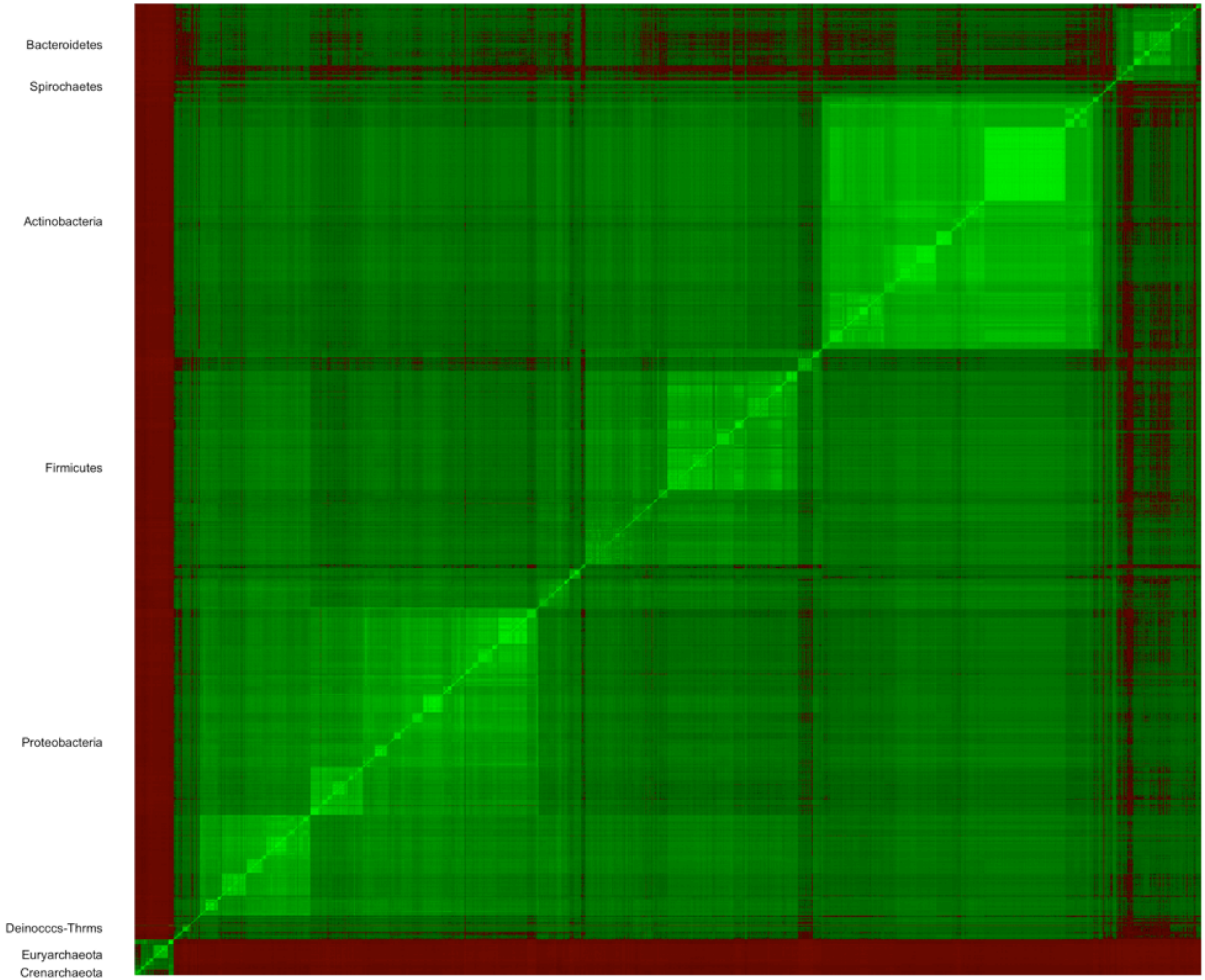
16S rRNA similarity of type strains of Bacteria and Archaea 2007



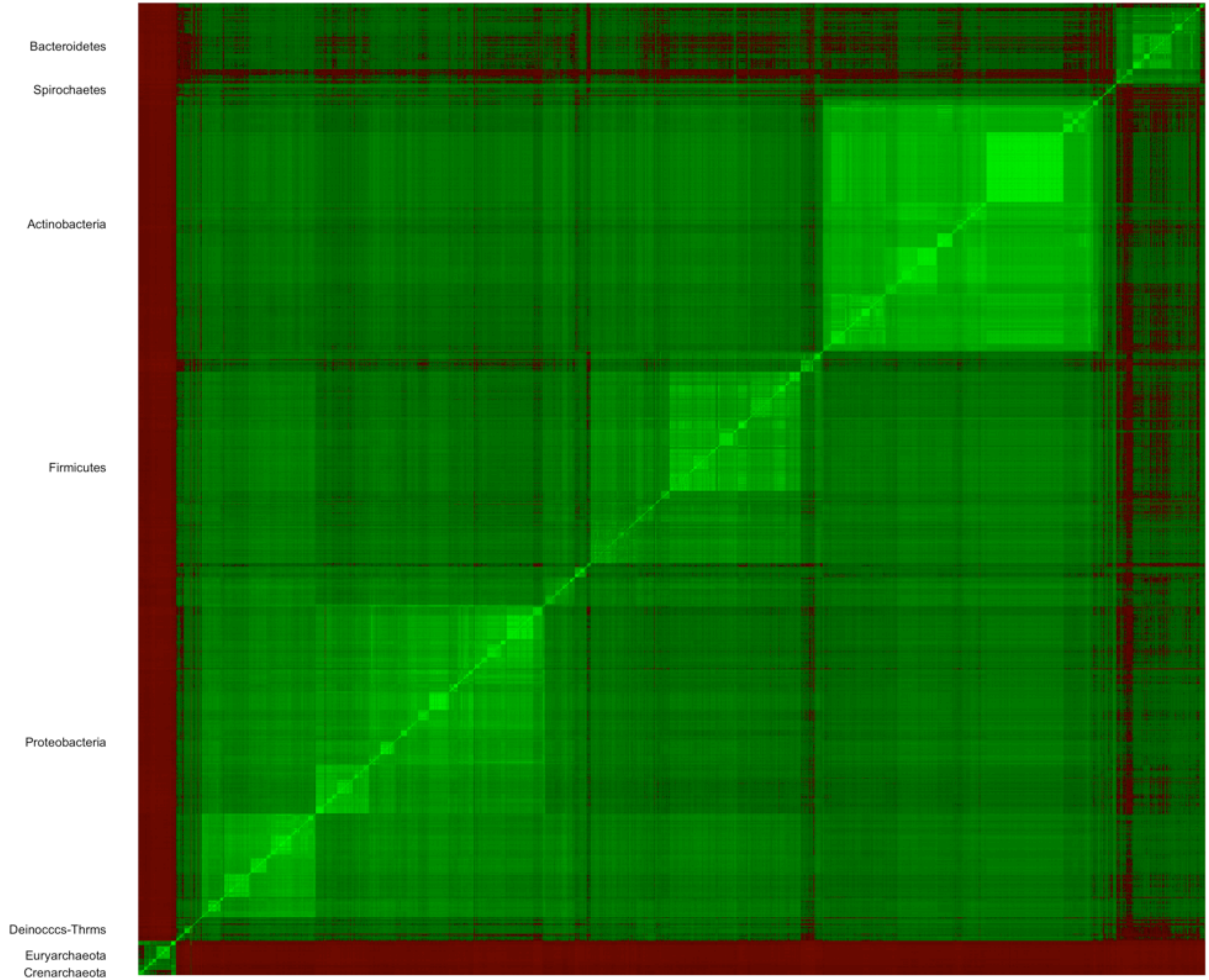
16S rRNA similarity of type strains of Bacteria and Archaea 2008



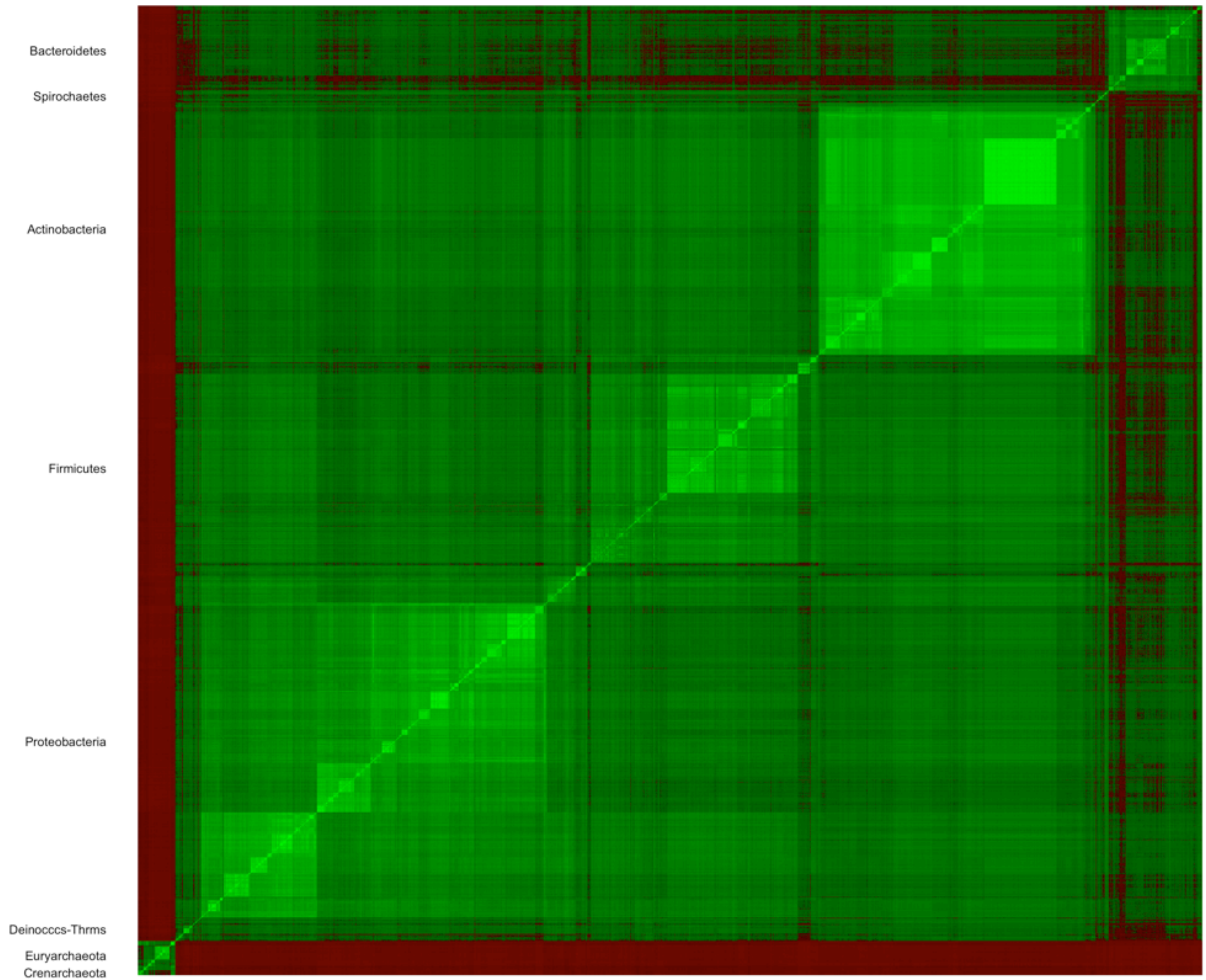
16S rRNA similarity of type strains of Bacteria and Archaea 2009



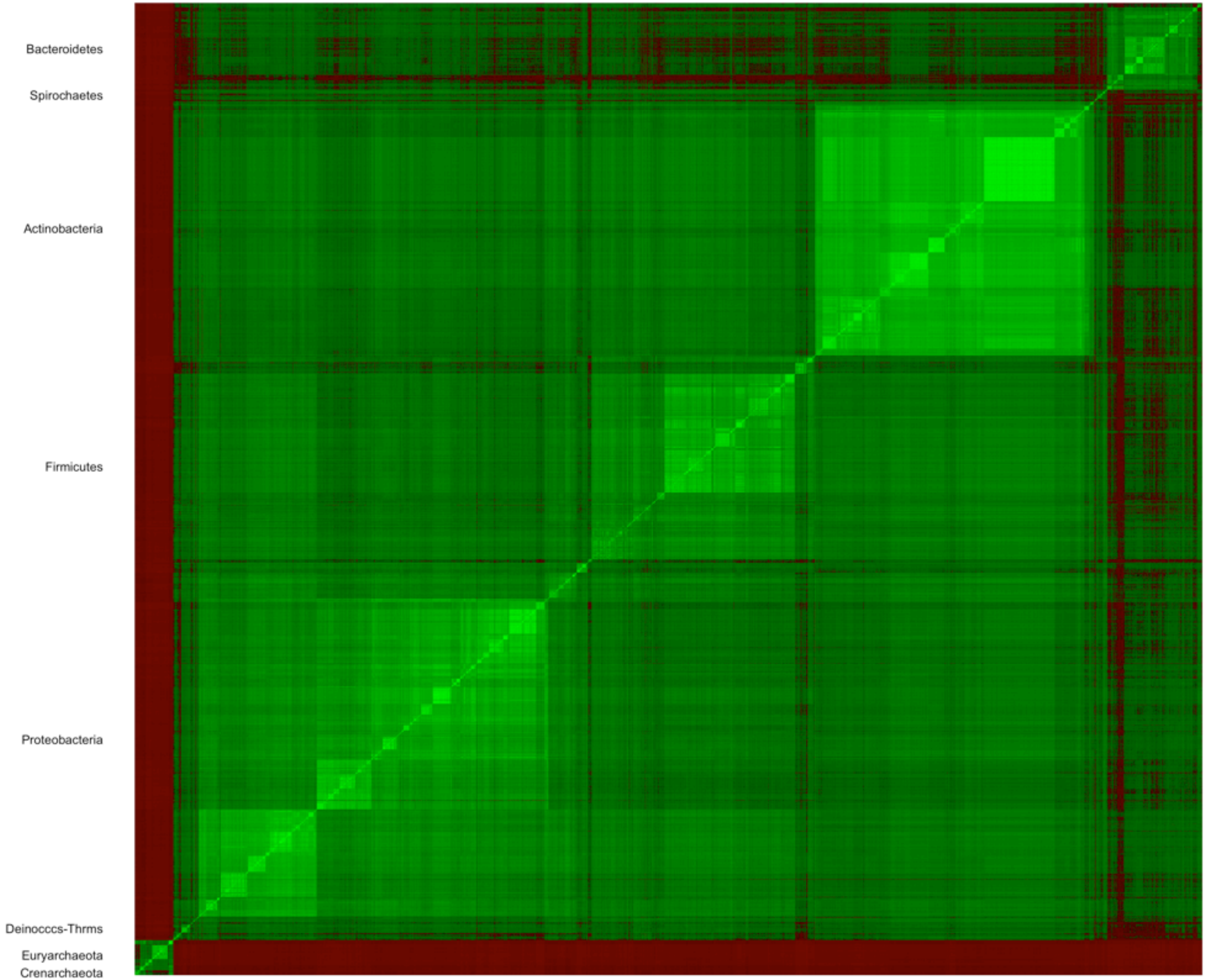
16S rRNA similarity of type strains of Bacteria and Archaea 2010



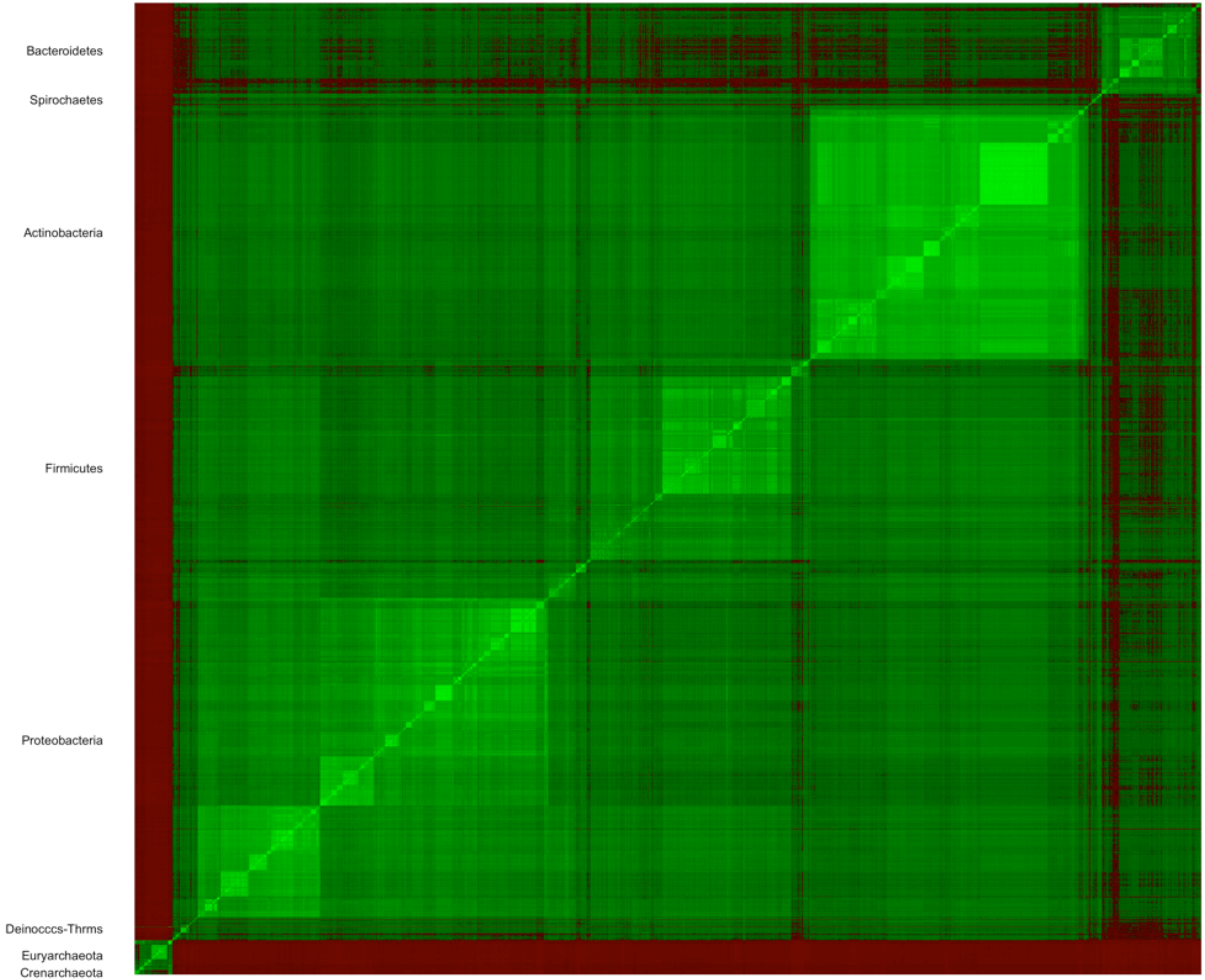
16S rRNA similarity of type strains of Bacteria and Archaea 2011



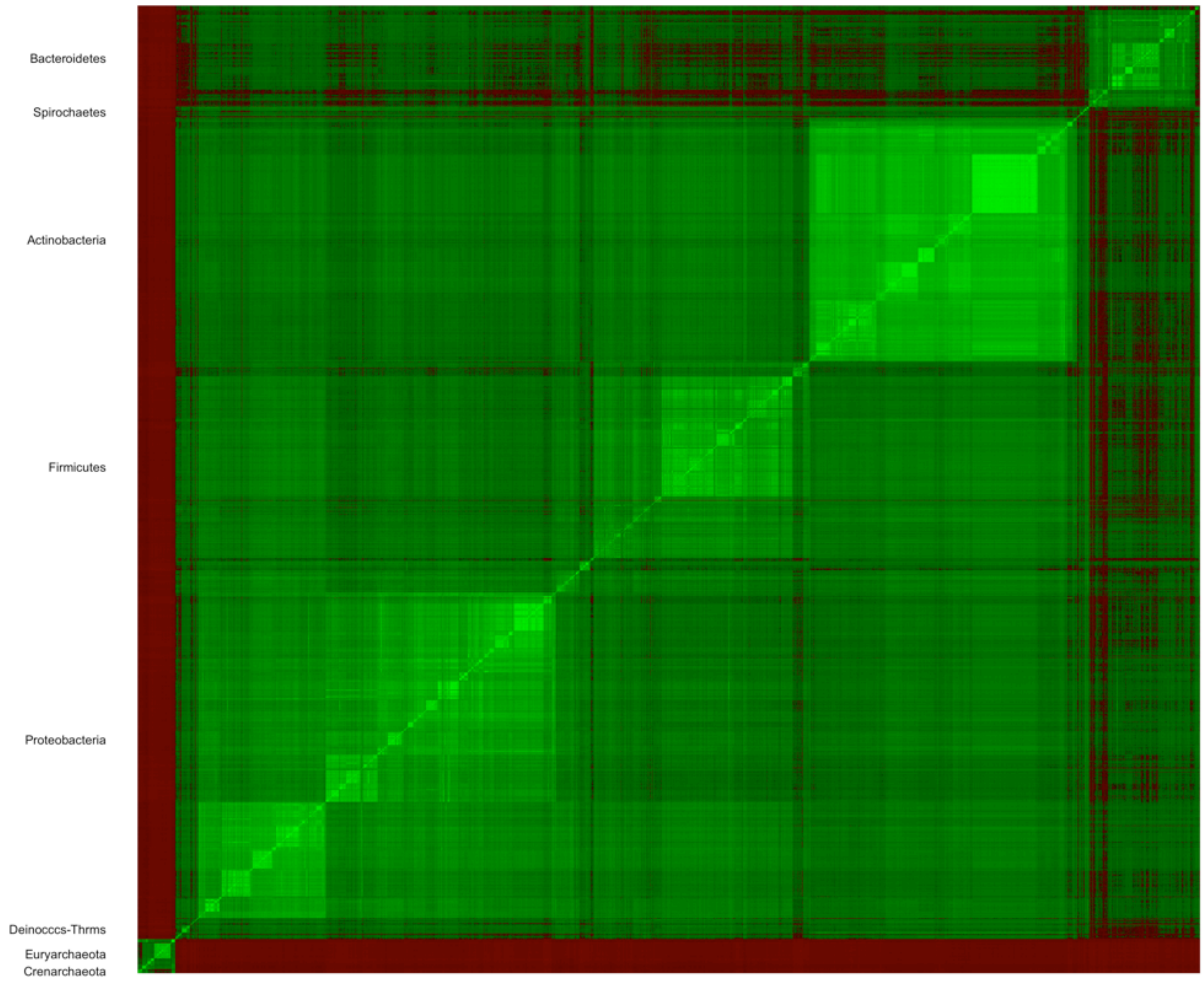
16S rRNA similarity of type strains of Bacteria and Archaea 2012



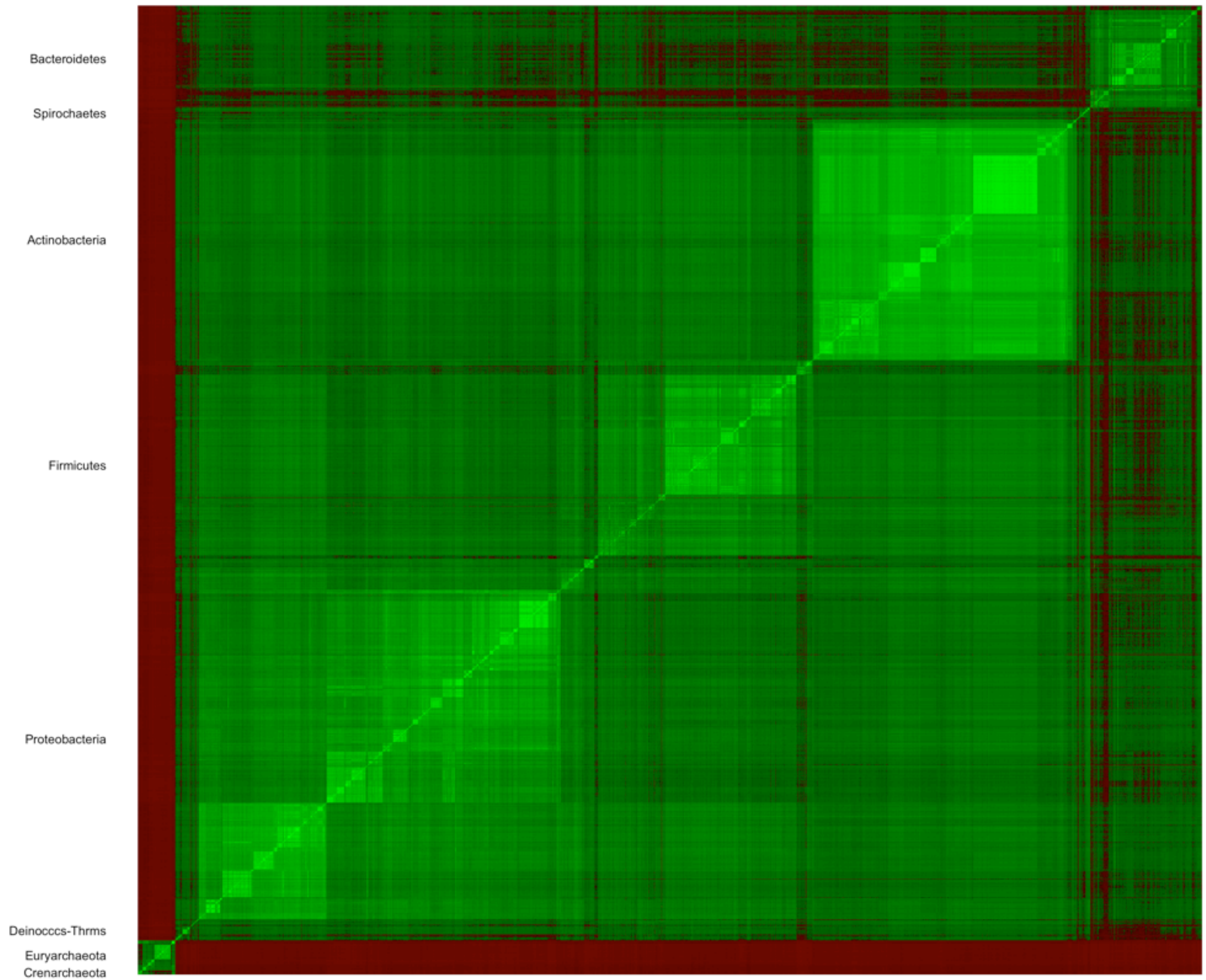
16S rRNA similarity of type strains of Bacteria and Archaea 2013



16S rRNA similarity of type strains of Bacteria and Archaea 2014



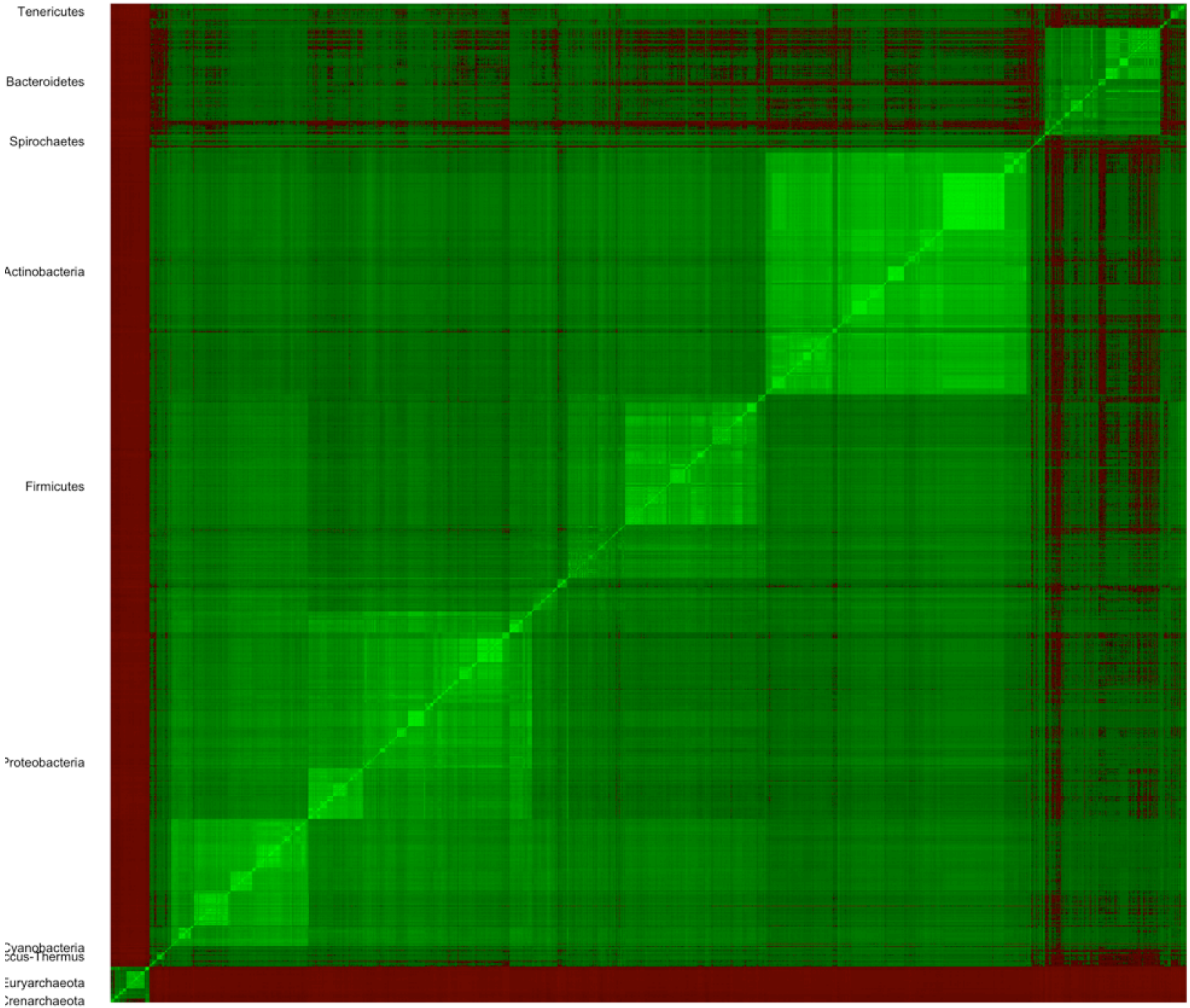
16S rRNA similarity of type strains of Bacteria and Archaea 2015



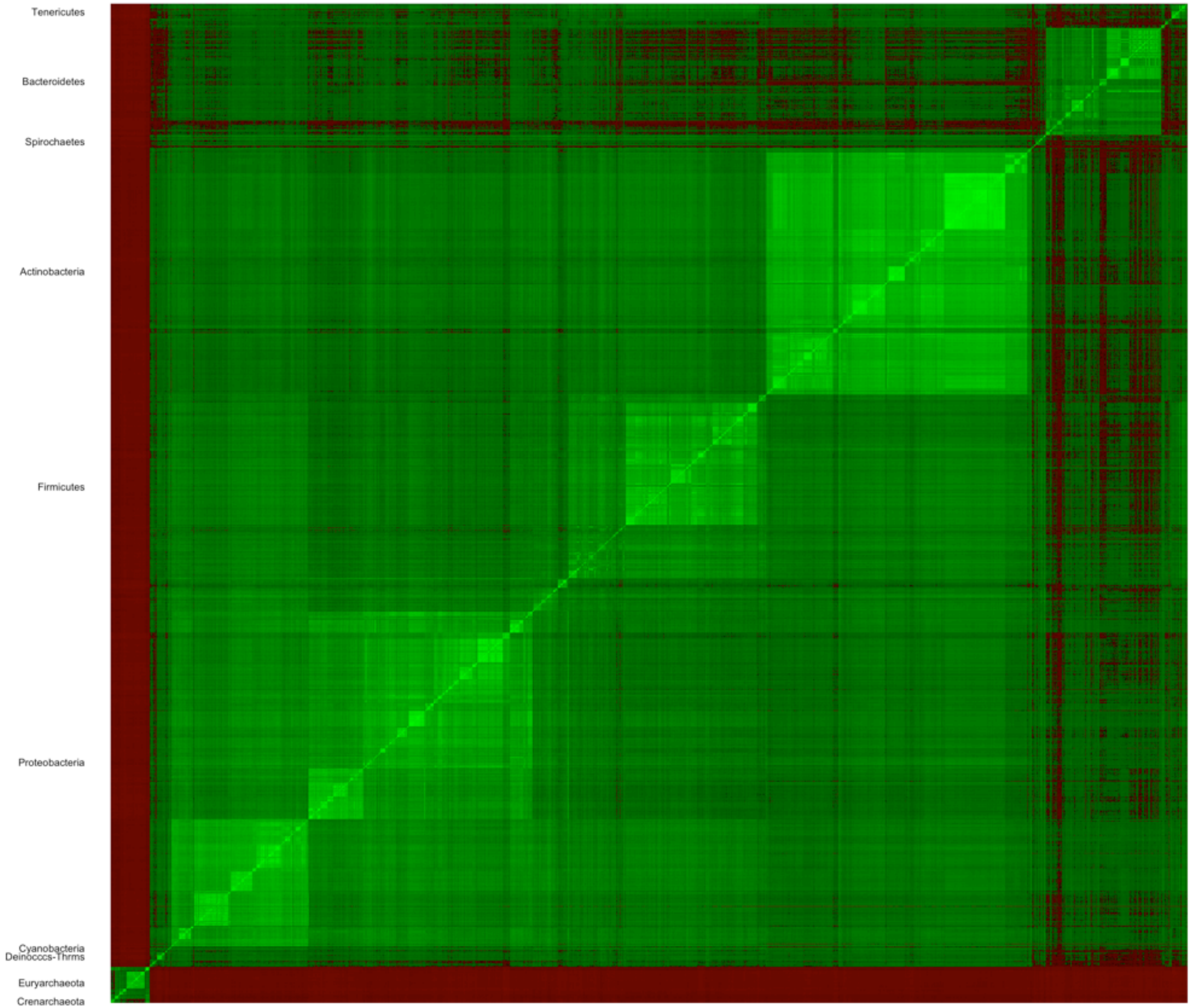
Refinement of the underlying taxonomy

- Improved version of the Self-Organizing Self Correcting Classifier
  - Goal - smooth classification while protecting taxonomic structure
- Input
  - 13,125 HQ full-length 16S rRNA sequences
    - Aligned to Greengenes alignment (Release 13.5)
    - Jukes-Cantor model, computed using Mothur
    - Output filtered to N4L-exemplar DOIs to select for validly published named species/subspecies
  - Outliers are identified at species level for members with four or more members based on within group distribution as represented in the lower triangle of each corresponding sub-matrix (2 σ), output visualized as heatmaps, phylograms, (APE) and histograms.
- 207/11,800 HQ 16S sequences were were flagged as outliers and excluded from further analysis

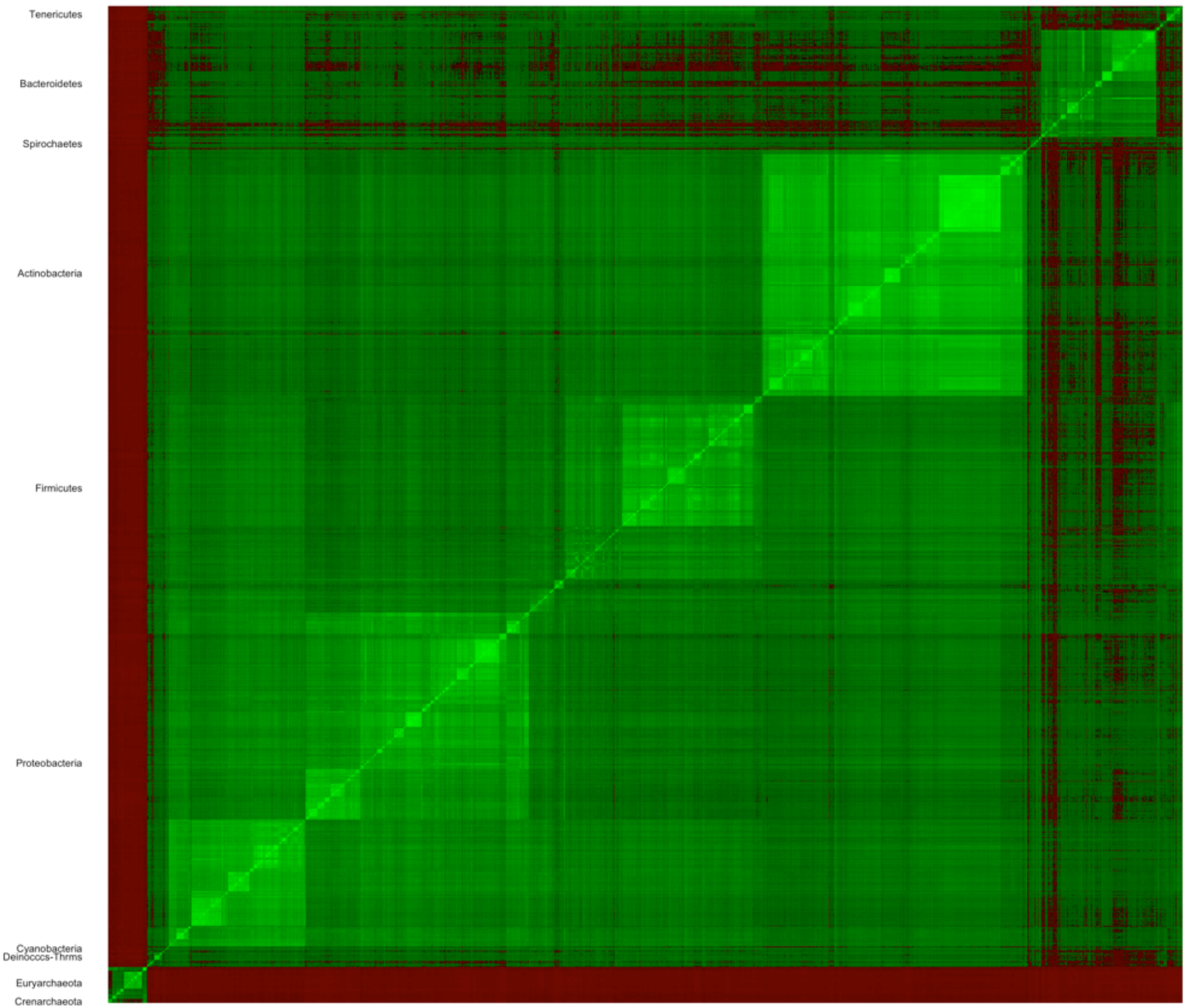
Species level rearrangement



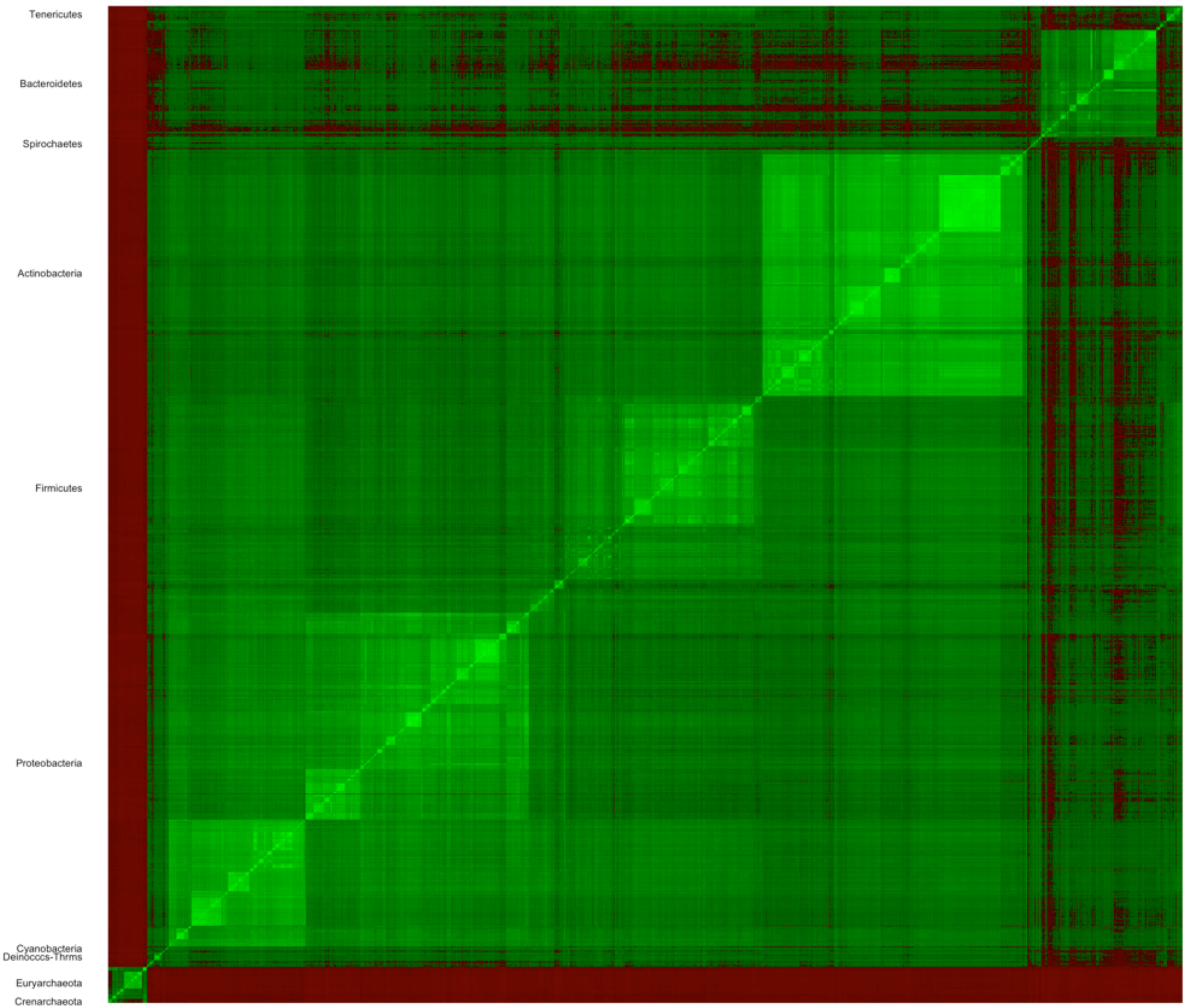
Genus level rearrangement



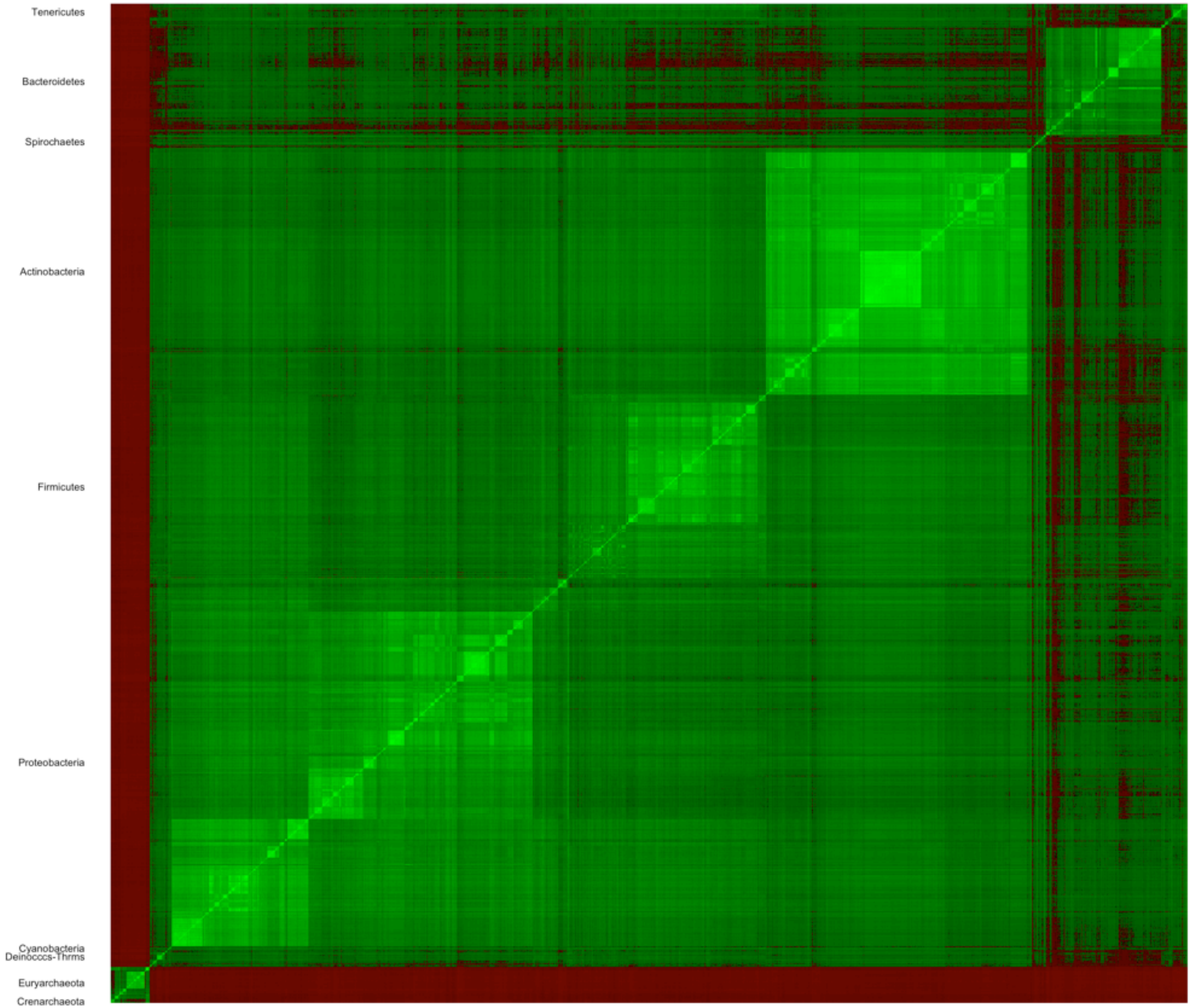
Family level rearrangement



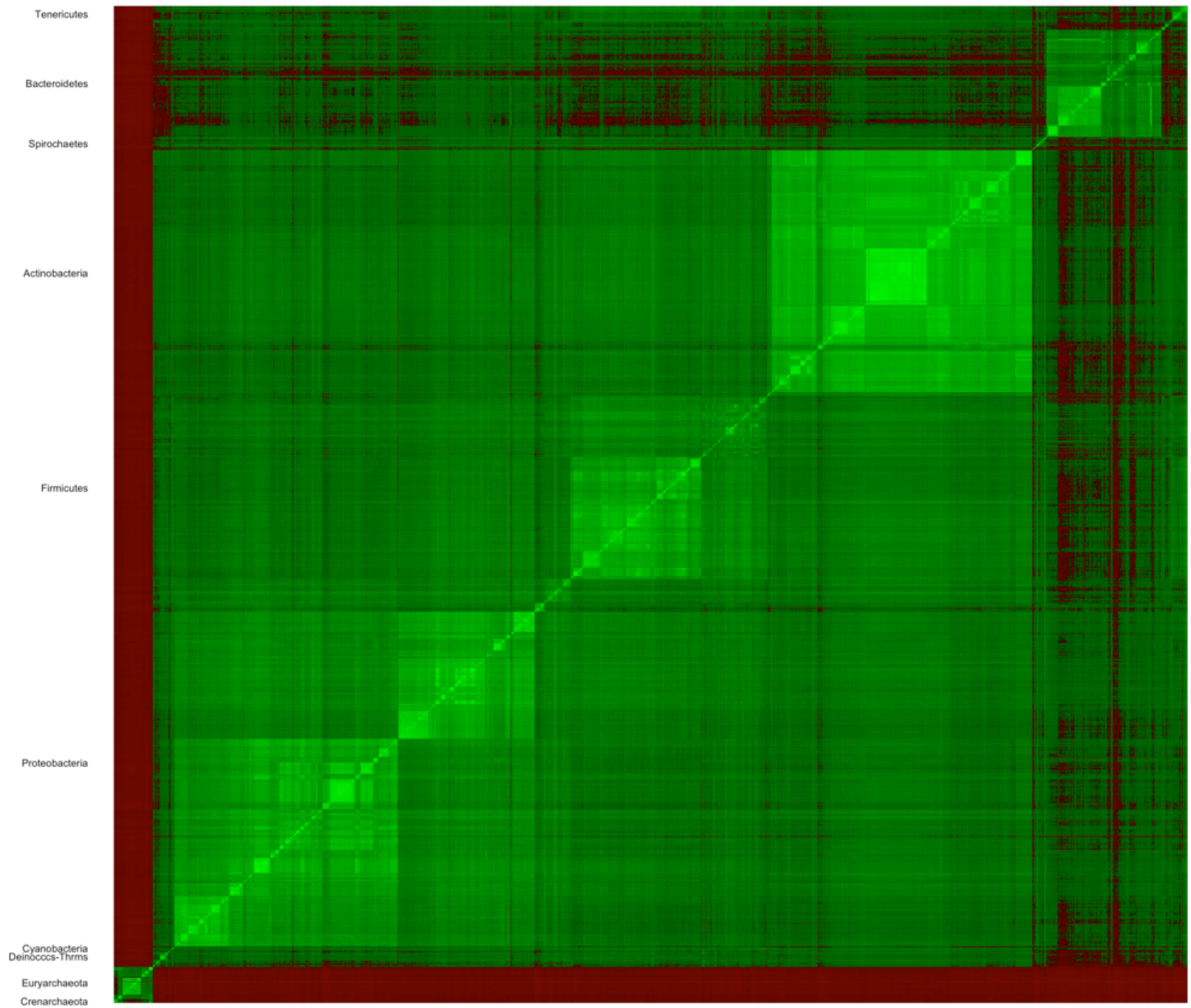
Order level rearrangement



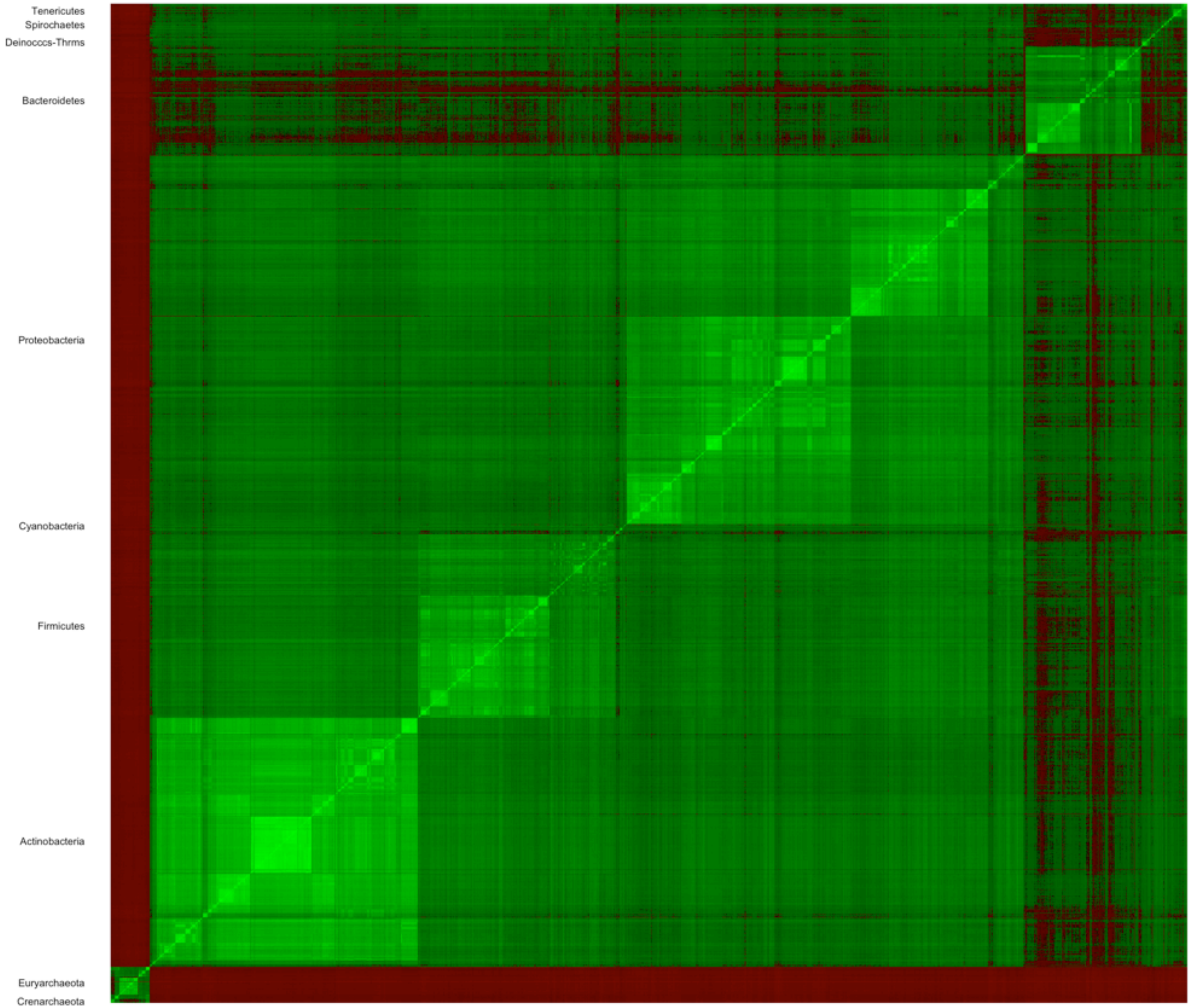
Class level rearrangement

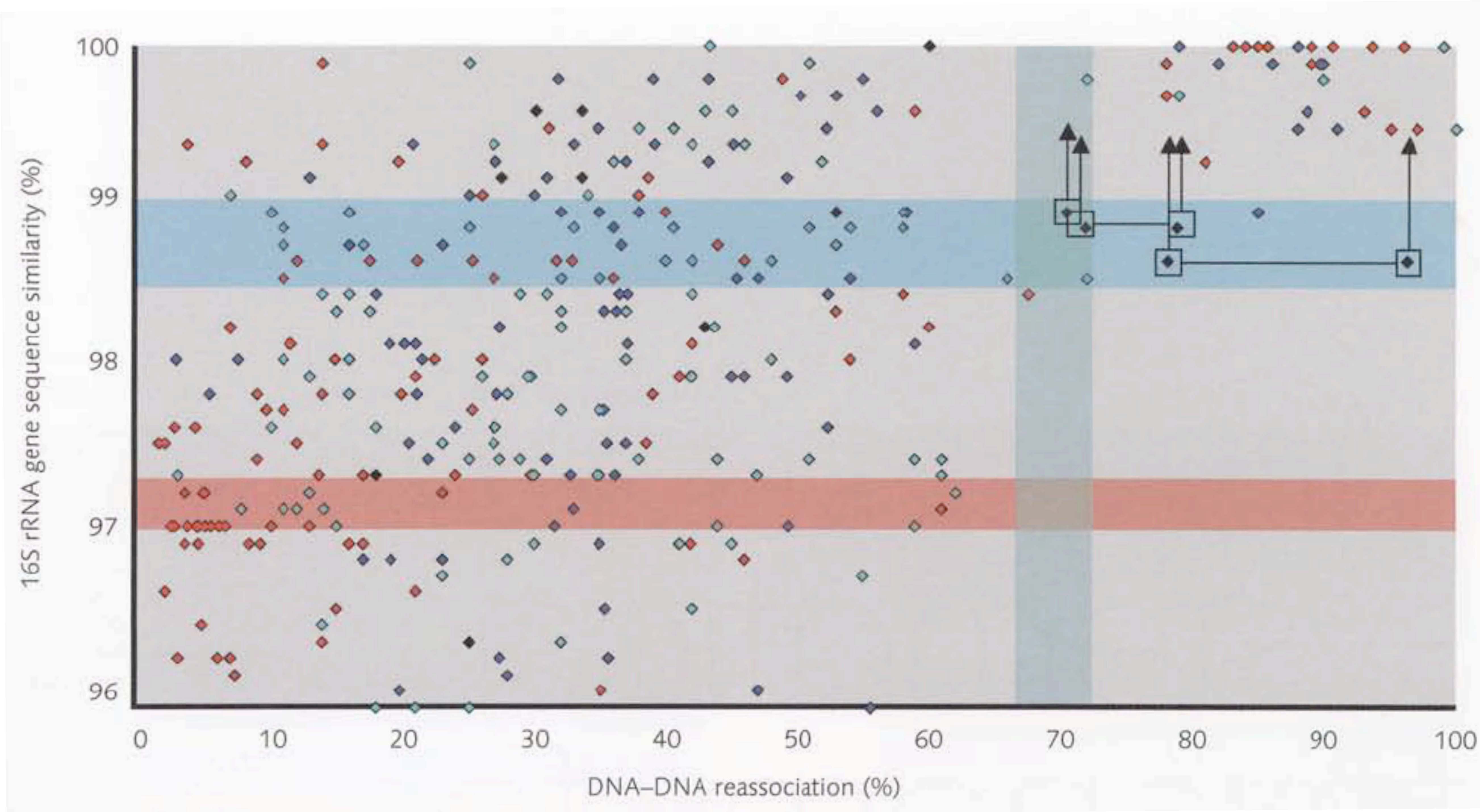


Phylum level rearrangement

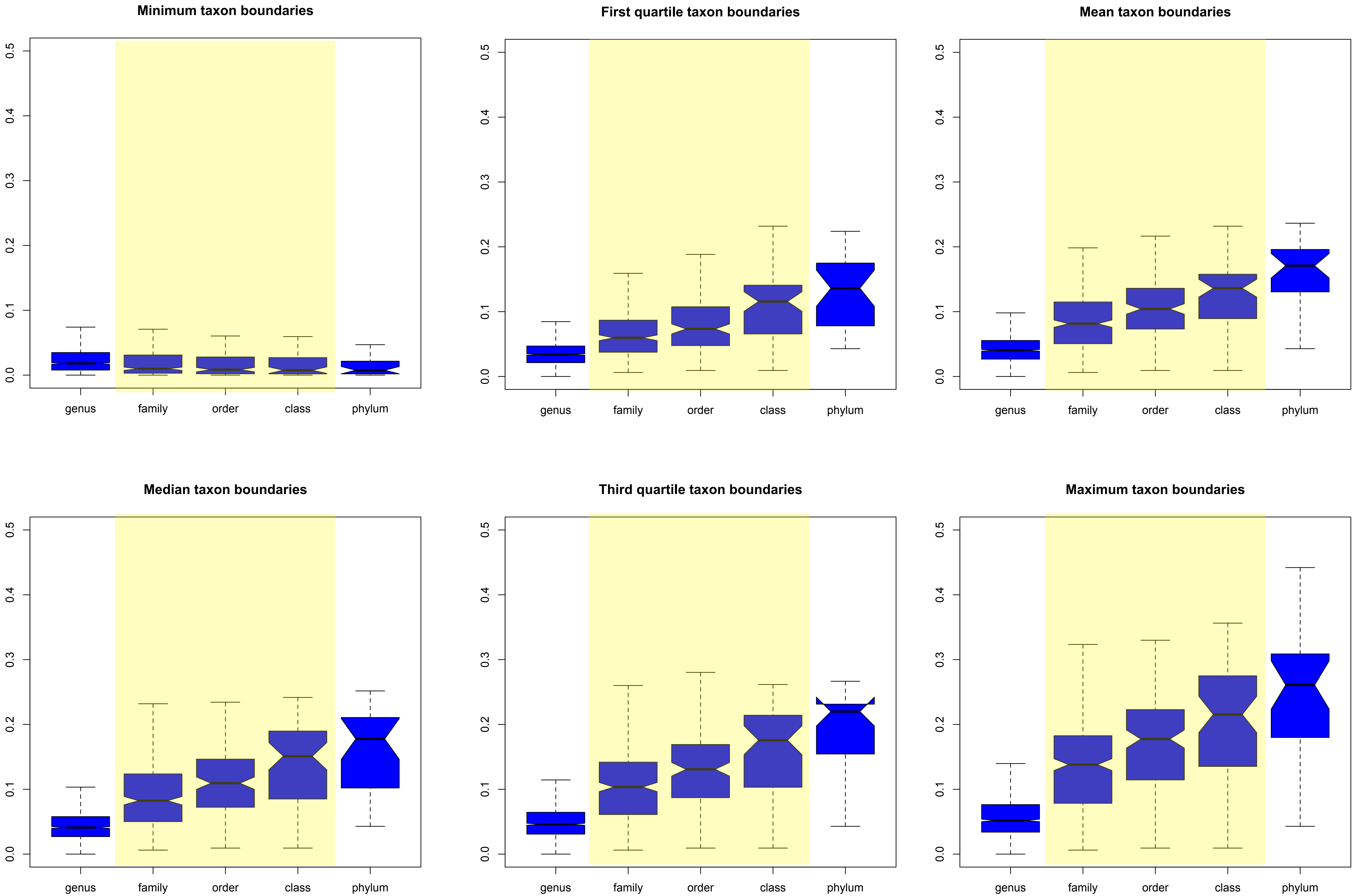


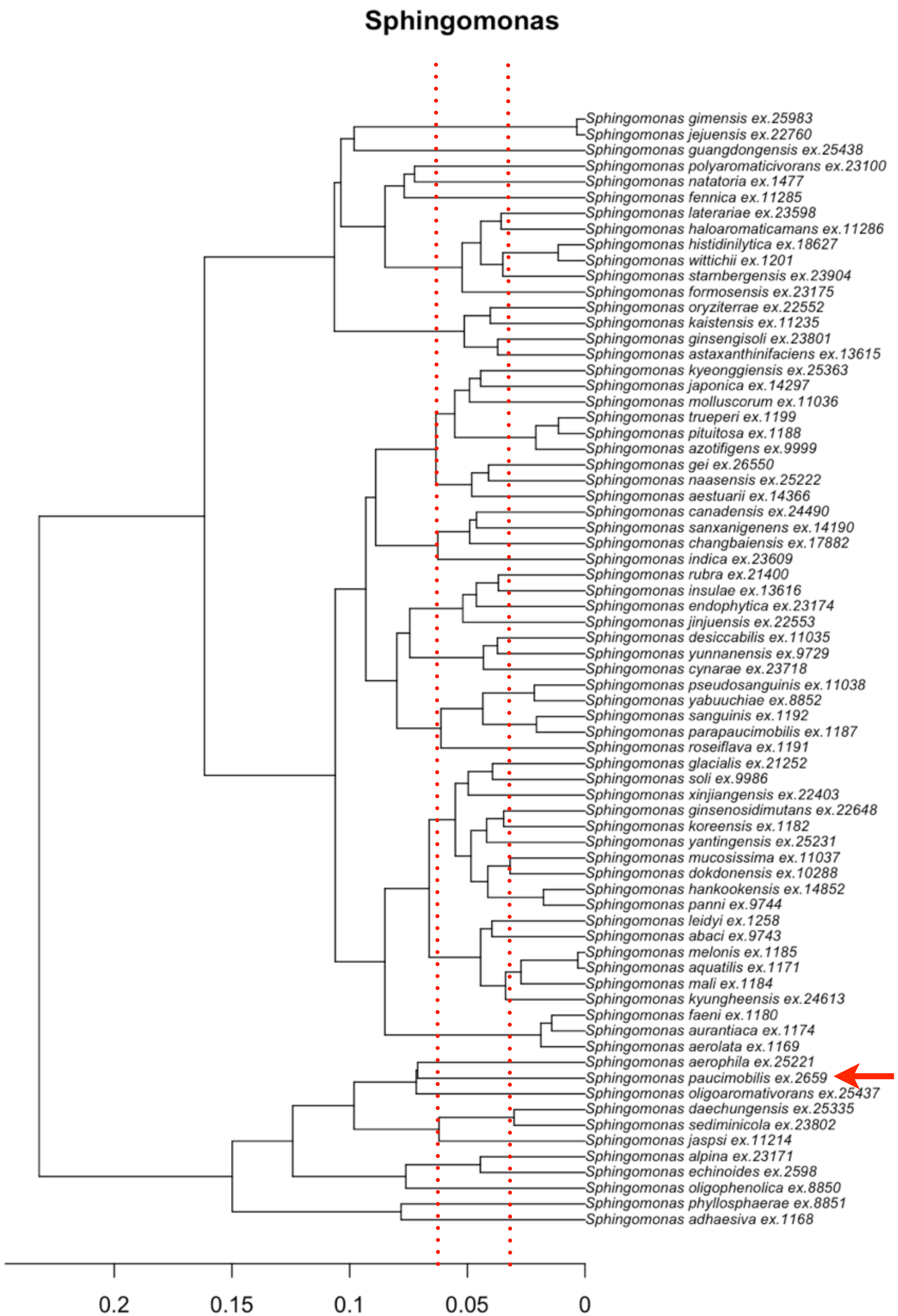
Domain level rearrangement

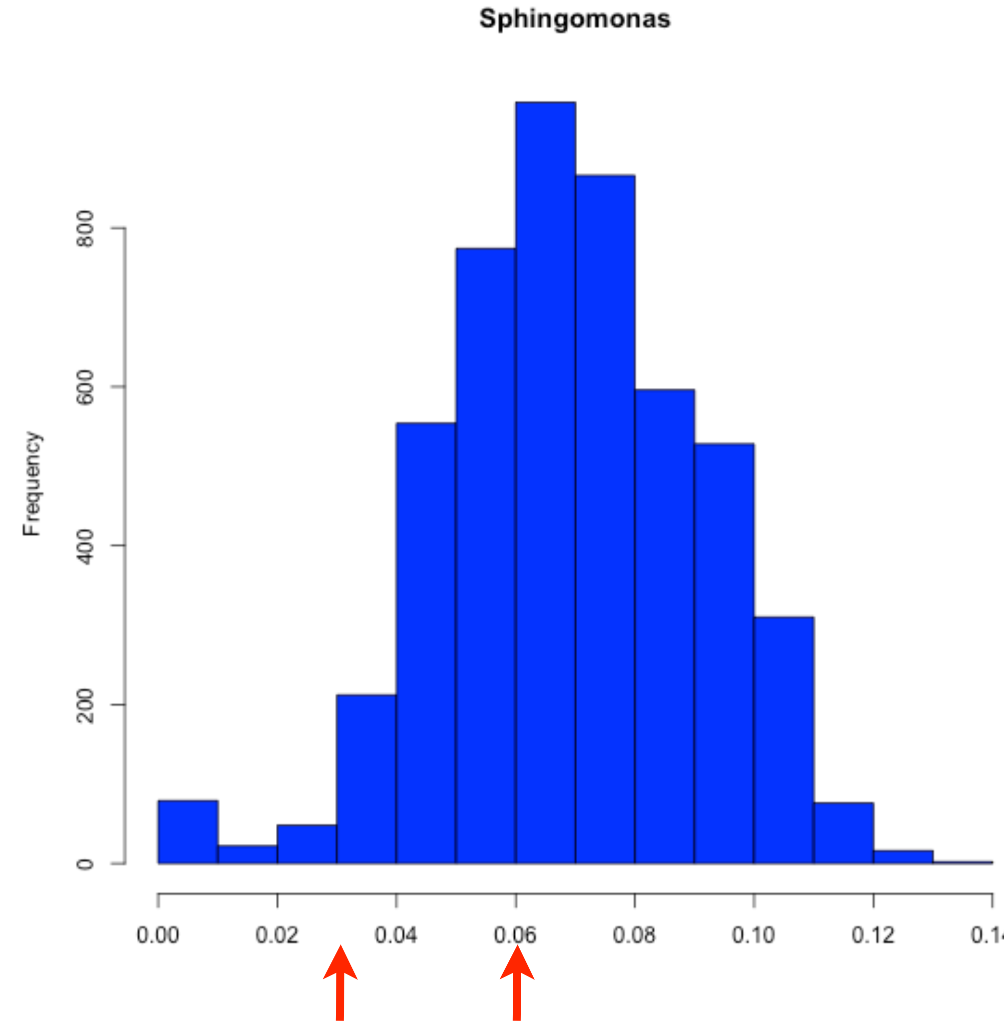


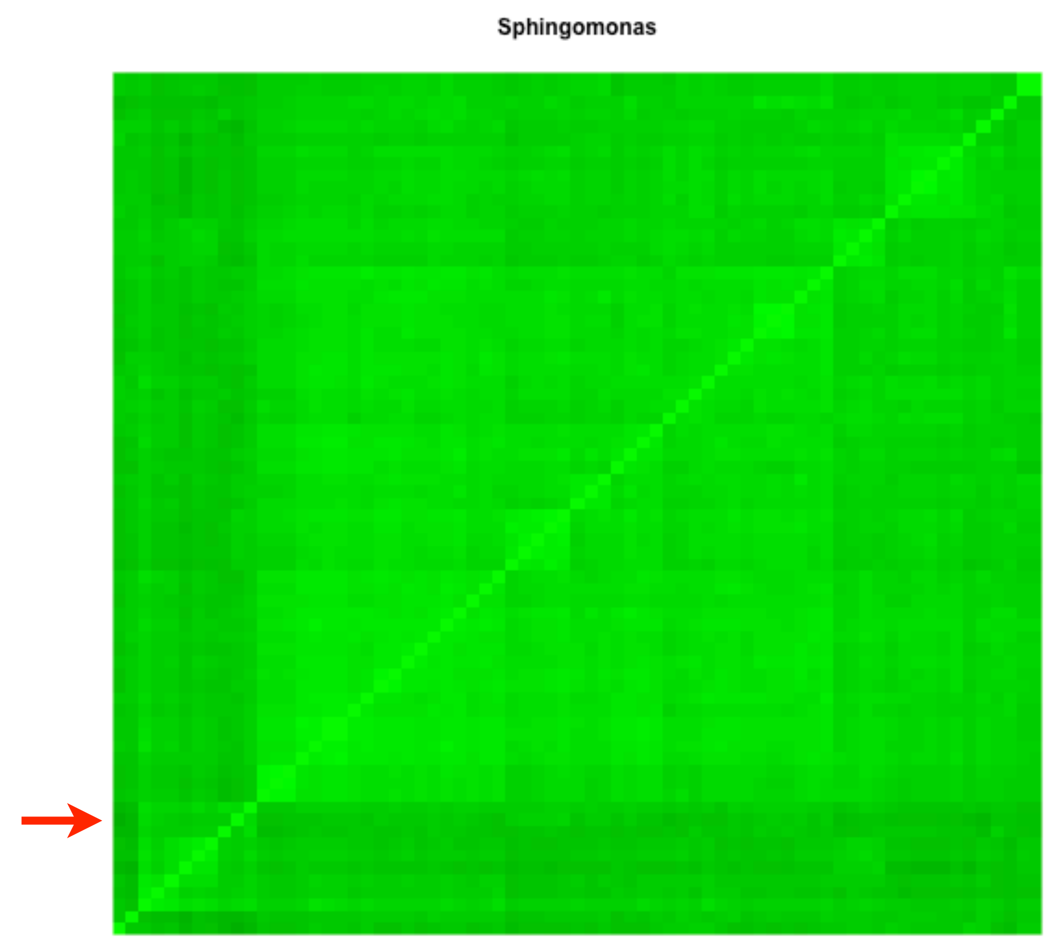


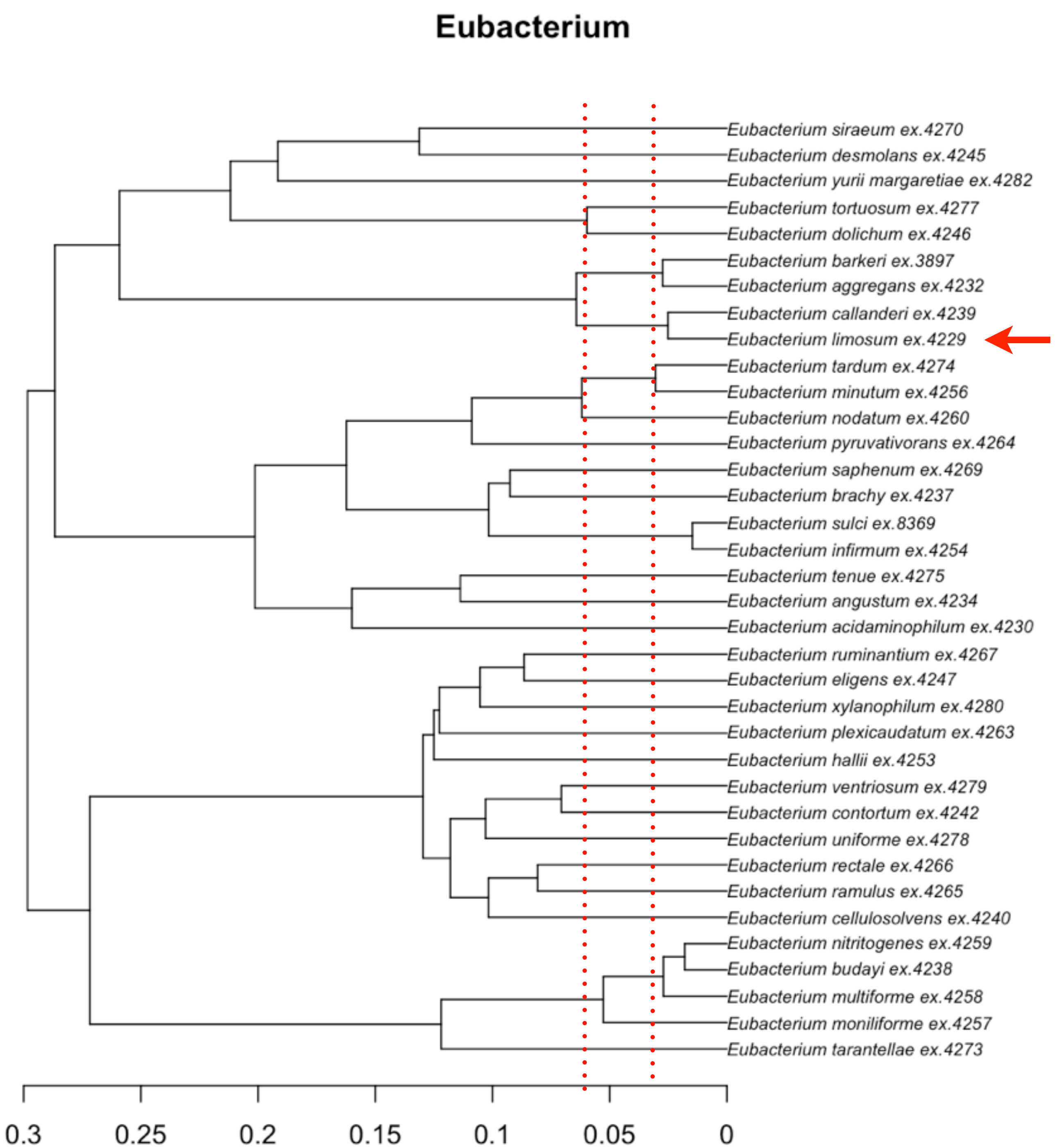
Taxonomic parameters revisited: Tarnished Gold Standards, Stackebrand, E. and Ebbers, J. Micro , may 06, 152-157

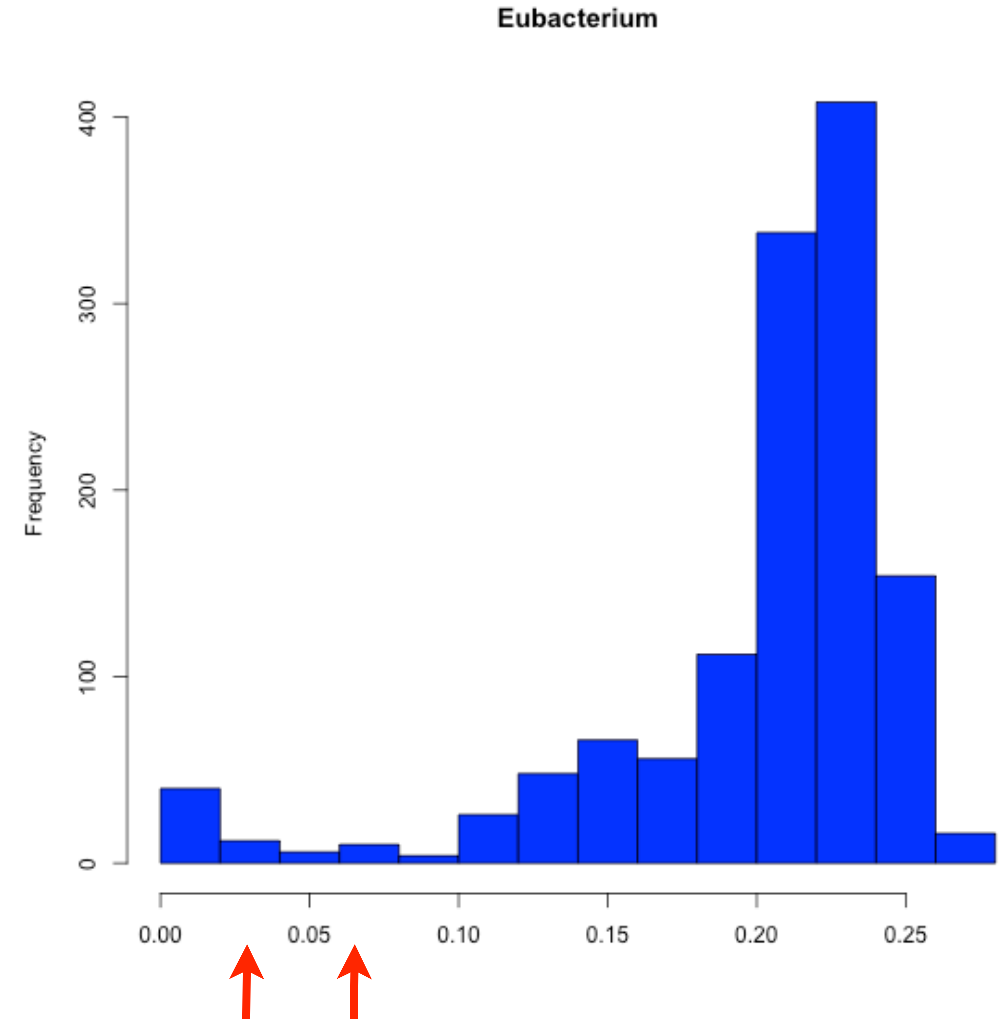

Minimum taxon boundaries
First quartile taxon boundaries
Mean taxon boundaries
Median taxon boundaries
Third quartile taxon boundaries
Maximum taxon boundaries
0.0
0.1
0.2
0.3
0.4
0.5
genus
family
order
class
phylum

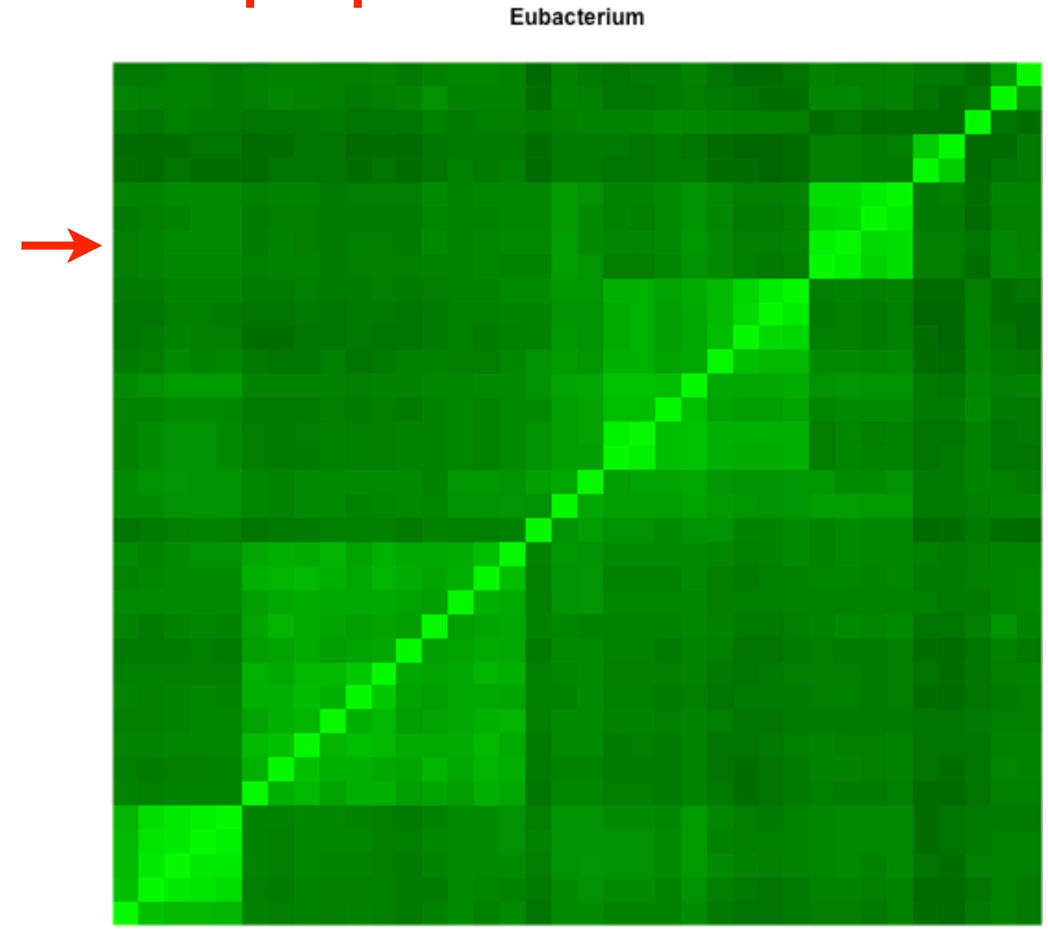

Sphingomonas
Sphingomonas gimensis ex.25983
Sphingomonas jejuensis ex.22760
Sphingomonas guangdongensis ex.25438
Sphingomonas polyaromaticivorans ex.23100
Sphingomonas natatoria ex.1477
Sphingomonas fennica ex.11285
Sphingomonas laterariae ex.23598
Sphingomonas haloaromaticamans ex.11286
Sphingomonas histidinilytica ex.18627
Sphingomonas wittichii ex.1201
Sphingomonas starnbergensis ex.23904
Sphingomonas formosensis ex.23175
Sphingomonas oryziterrae ex.22552
Sphingomonas kaistensis ex.11235
Sphingomonas ginsengisoli ex.23801
Sphingomonas astaxanthinifaciens ex.13615
Sphingomonas kyeonggiensis ex.25363
Sphingomonas japonica ex.14297
Sphingomonas molluscorum ex.11036
Sphingomonas trueperi ex.1199
Sphingomonas pituitosa ex.1188
Sphingomonas azotifigens ex.9999
Sphingomonas gei ex.26550
Sphingomonas naasensis ex.25222
Sphingomonas aestuarii ex.14366
Sphingomonas canadensis ex.24490
Sphingomonas sanxanigenens ex.14190
Sphingomonas changbaiensis ex.17882
Sphingomonas indica ex.23609
Sphingomonas rubra ex.21400
Sphingomonas insulae ex.13616
Sphingomonas endophytica ex.23174
Sphingomonas jinjuensis ex.22553
Sphingomonas desiccabilis ex.11035
Sphingomonas yunnanensis ex.9729
Sphingomonas cynarae ex.23718
Sphingomonas pseudosanguinis ex.11038
Sphingomonas yabuuchiae ex.8852
Sphingomonas sanguinis ex.1192
Sphingomonas parapaucimobilis ex.1187
Sphingomonas roseiflava ex.1191
Sphingomonas glacialis ex.21252
Sphingomonas soli ex.9986
Sphingomonas xinjiangensis ex.22403
Sphingomonas ginsenosidimutans ex.22648
Sphingomonas koreensis ex.1182
Sphingomonas yantingensis ex.25231
Sphingomonas mucosissima ex.11037
Sphingomonas dokdonensis ex.10288
Sphingomonas hankookensis ex.14852
Sphingomonas panni ex.9744
Sphingomonas leidyi ex.1258
Sphingomonas abaci ex.9743
Sphingomonas melonis ex.1185
Sphingomonas aquatilis ex.1171
Sphingomonas mali ex.1184
Sphingomonas kyungheensis ex.24613
Sphingomonas faeni ex.1180
Sphingomonas aurantiaca ex.1174
Sphingomonas aerolata ex.1169
Sphingomonas aerophila ex.25221
Sphingomonas paucimobilis ex.2659
Sphingomonas oligoaromativorans ex.25437
Sphingomonas daechungensis ex.25335
Sphingomonas sediminicola ex.23802
Sphingomonas jaspsi ex.11214
Sphingomonas alpina ex.23171
Sphingomonas echinoides ex.2598
Sphingomonas oligophenolica ex.8850
Sphingomonas phyllosphaerae ex.8851
Sphingomonas adhaesiva ex.1168
0.2
0.15
0.1
0.05
0

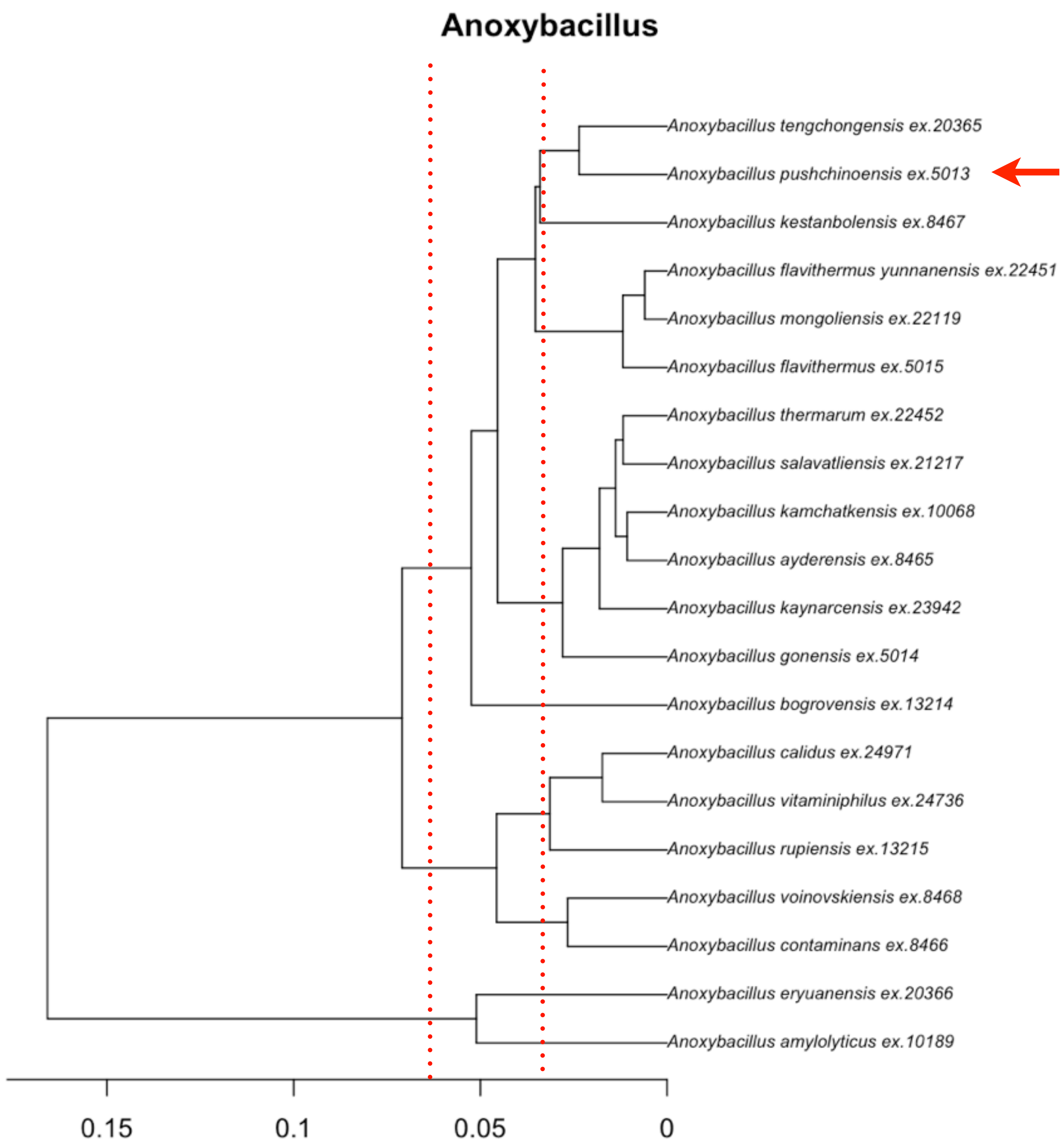

Sphingomonas
Frequency
0
200
400
600
800
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14

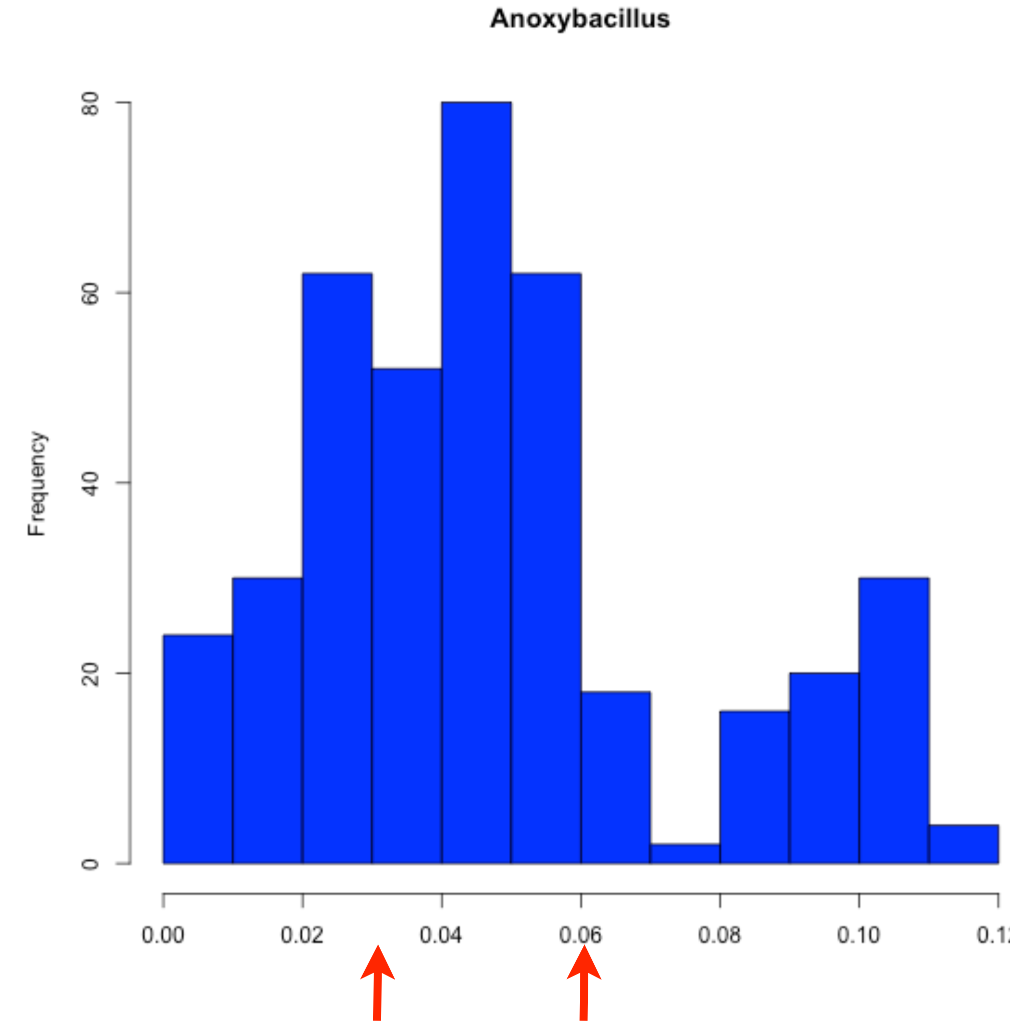

Sphingomonas

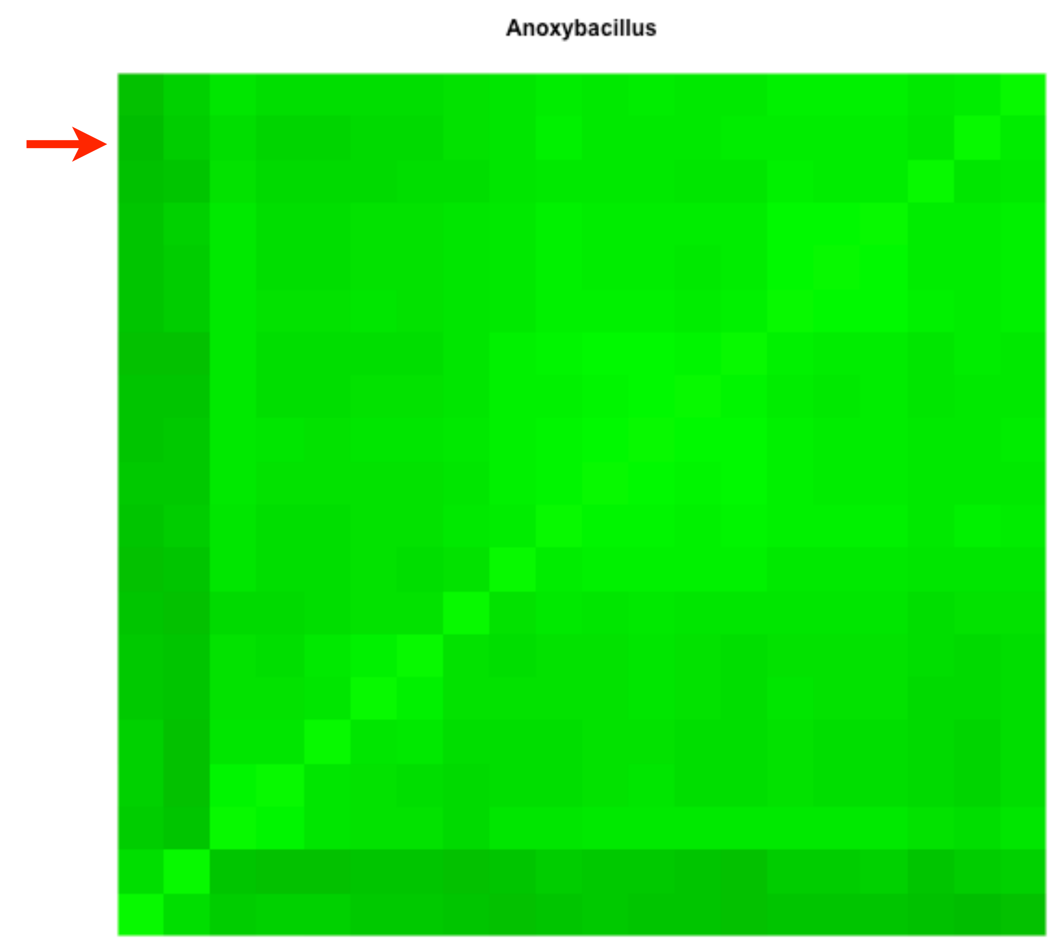

Eubacterium
Eubacterium siraeum ex.4270
Eubacterium desmolans ex.4245
Eubacterium yurii margaretiae ex.4282
Eubacterium tortuosum ex.4277
Eubacterium dolichum ex.4246
Eubacterium barkeri ex.3897
Eubacterium aggregans ex.4232
Eubacterium callanderi ex.4239
Eubacterium limosum ex.4229
Eubacterium tardum ex.4274
Eubacterium minutum ex.4256
Eubacterium nodatum ex.4260
Eubacterium pyruvativorans ex.4264
Eubacterium saphenum ex.4269
Eubacterium brachy ex.4237
Eubacterium sulci ex.8369
Eubacterium infirmum ex.4254
Eubacterium tenue ex.4275
Eubacterium angustum ex.4234
Eubacterium acidaminophilum ex.4230
Eubacterium ruminantium ex.4267
Eubacterium eligens ex.4247
Eubacterium xylanophilum ex.4280
Eubacterium plexicaudatum ex.4263
Eubacterium hallii ex.4253
Eubacterium ventriosum ex.4279
Eubacterium contortum ex.4242
Eubacterium uniforme ex.4278
Eubacterium rectale ex.4266
Eubacterium ramulus ex.4265
Eubacterium cellulosolvens ex.4240
Eubacterium nitritogenes ex.4259
Eubacterium budayi ex.4238
Eubacterium multiforme ex.4258
Eubacterium moniliforme ex.4257
Eubacterium tarantellae ex.4273
0.3
0.25
0.2
0.15
0.1
0.05
0

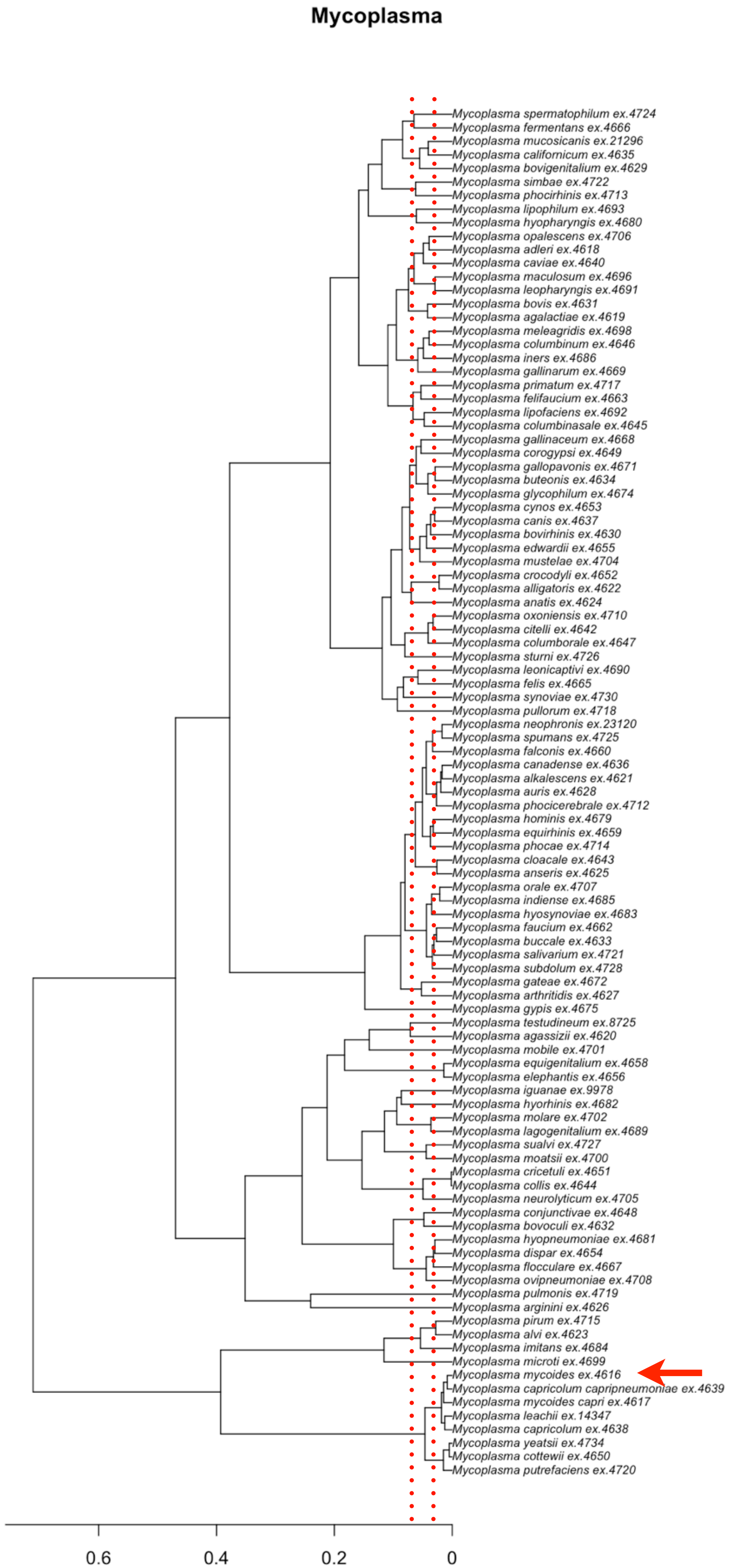

Eubacterium
Frequency
0
100
200
300
400
0.00
0.05
0.10
0.15
0.20
0.25

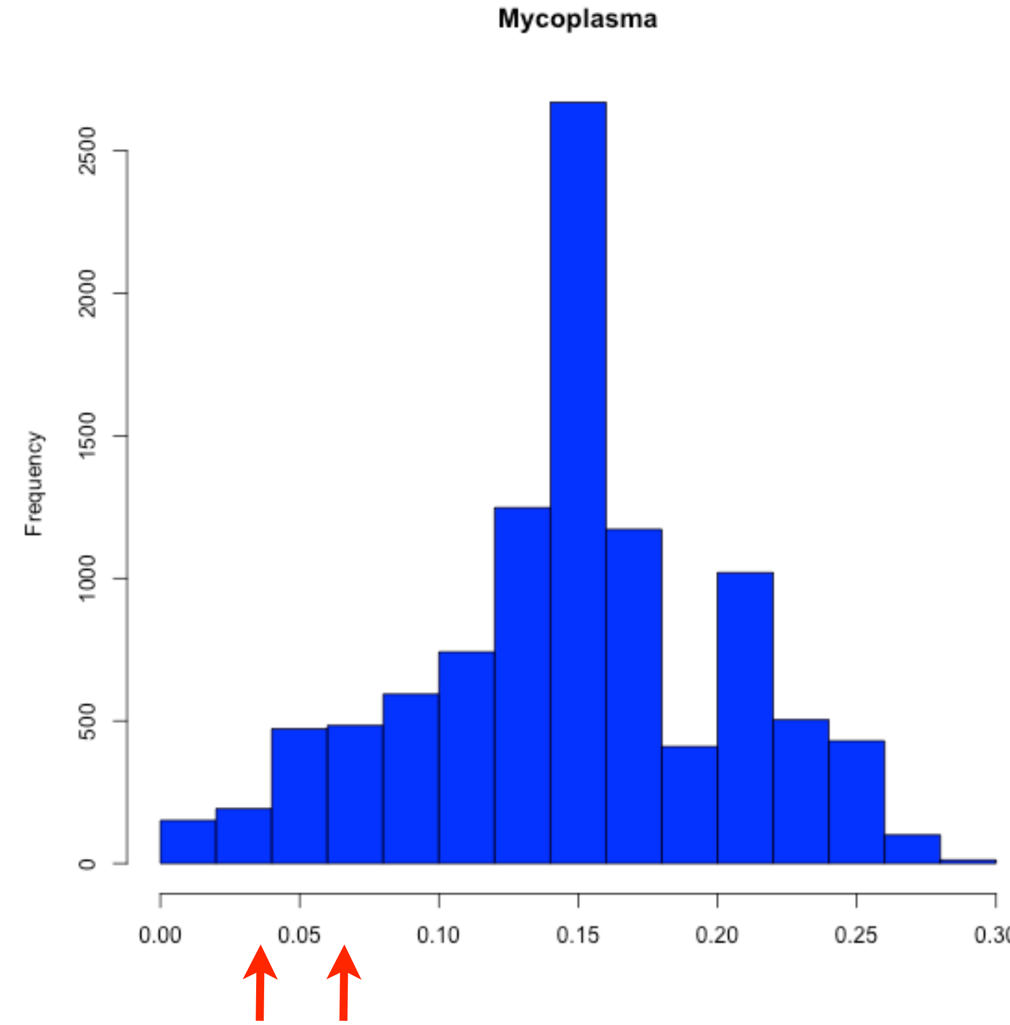

Eubacterium

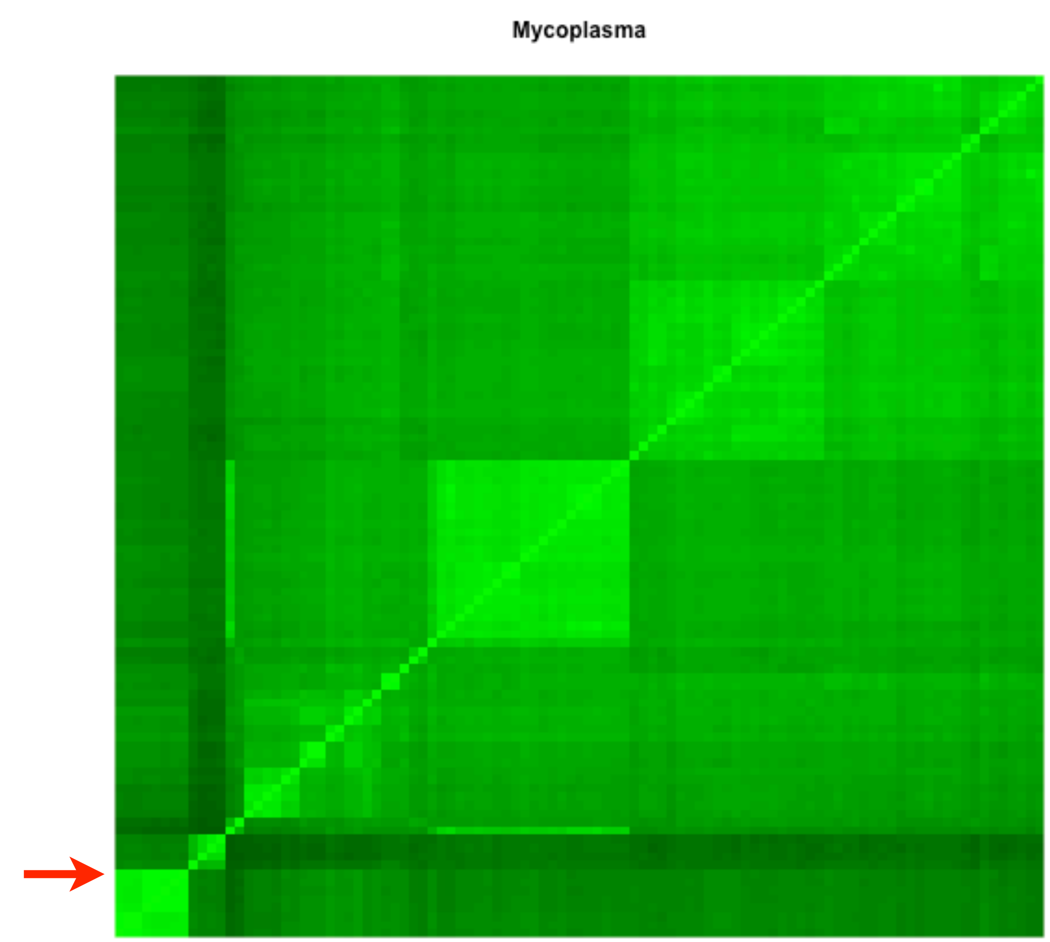

Anoxybacillus
Anoxybacillus tengchongensis ex.20365
Anoxybacillus pushchinoensis ex.5013
Anoxybacillus kestanbolensis ex.8467
Anoxybacillus flavithermus yunnanensis ex.22451
Anoxybacillus mongoliensis ex.22119
Anoxybacillus flavithermus ex.5015
Anoxybacillus thermarum ex.22452
Anoxybacillus salavatliensis ex.21217
Anoxybacillus kamchatkensis ex.10068
Anoxybacillus ayderensis ex.8465
Anoxybacillus kaynarcensis ex.23942
Anoxybacillus gonensis ex.5014
Anoxybacillus bogrovensis ex.13214
Anoxybacillus calidus ex.24971
Anoxybacillus vitaminiphilus ex.24736
Anoxybacillus rupiensis ex.13215
Anoxybacillus voinovskiensis ex.8468
Anoxybacillus contaminans ex.8466
Anoxybacillus eryuanensis ex.20366
Anoxybacillus amylolyticus ex.10189
0.15
0.1
0.05
0

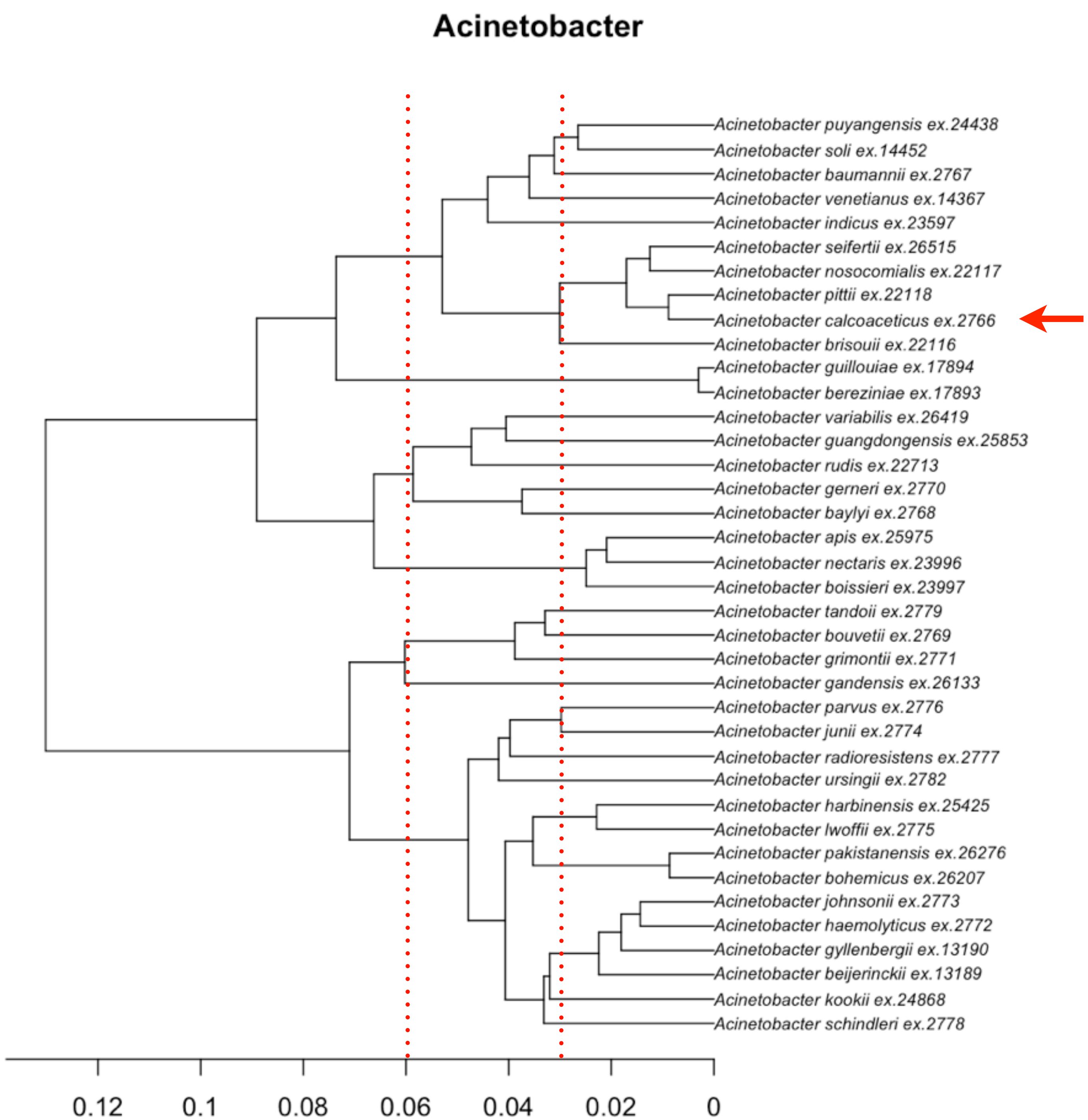

Anoxybacillus
Frequency
0
20
40
60
80
0.00
0.02
0.04
0.06
0.08
0.10
0.12

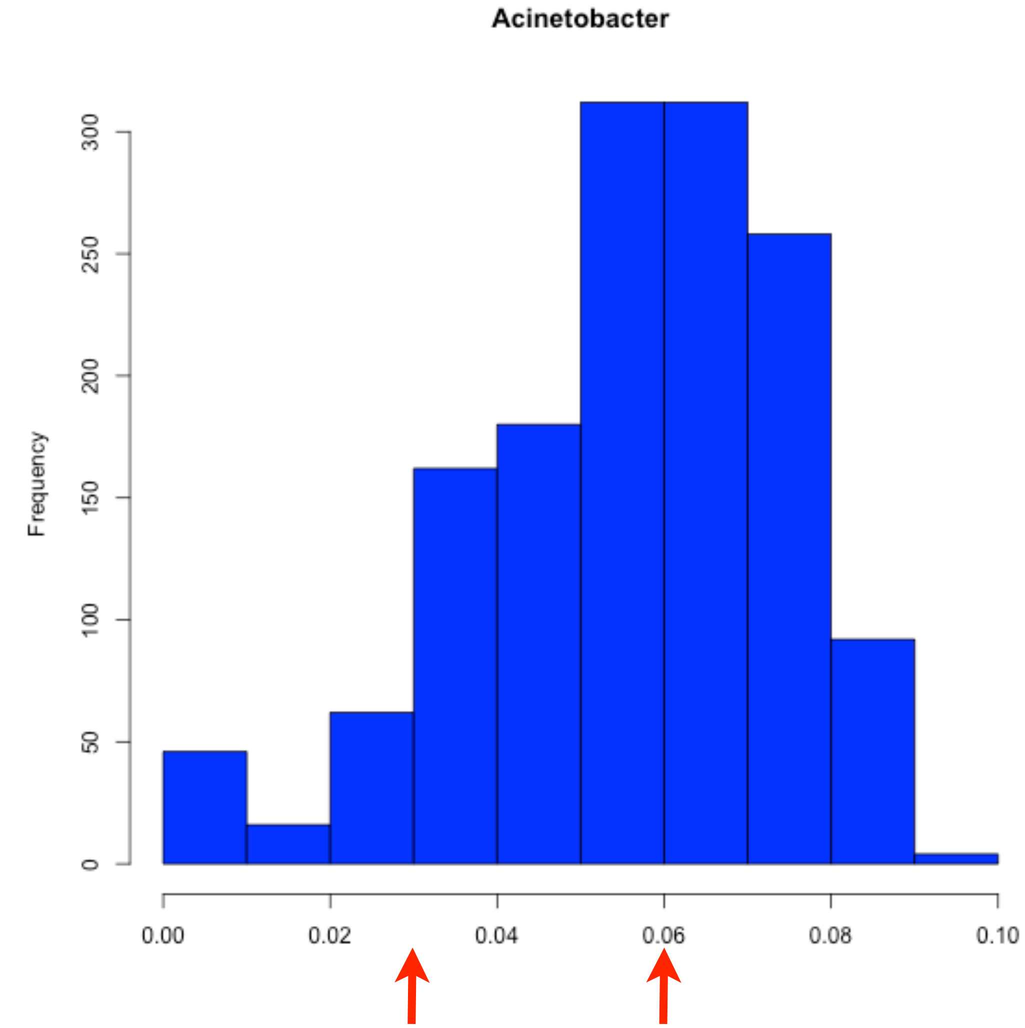

Anoxybacillus

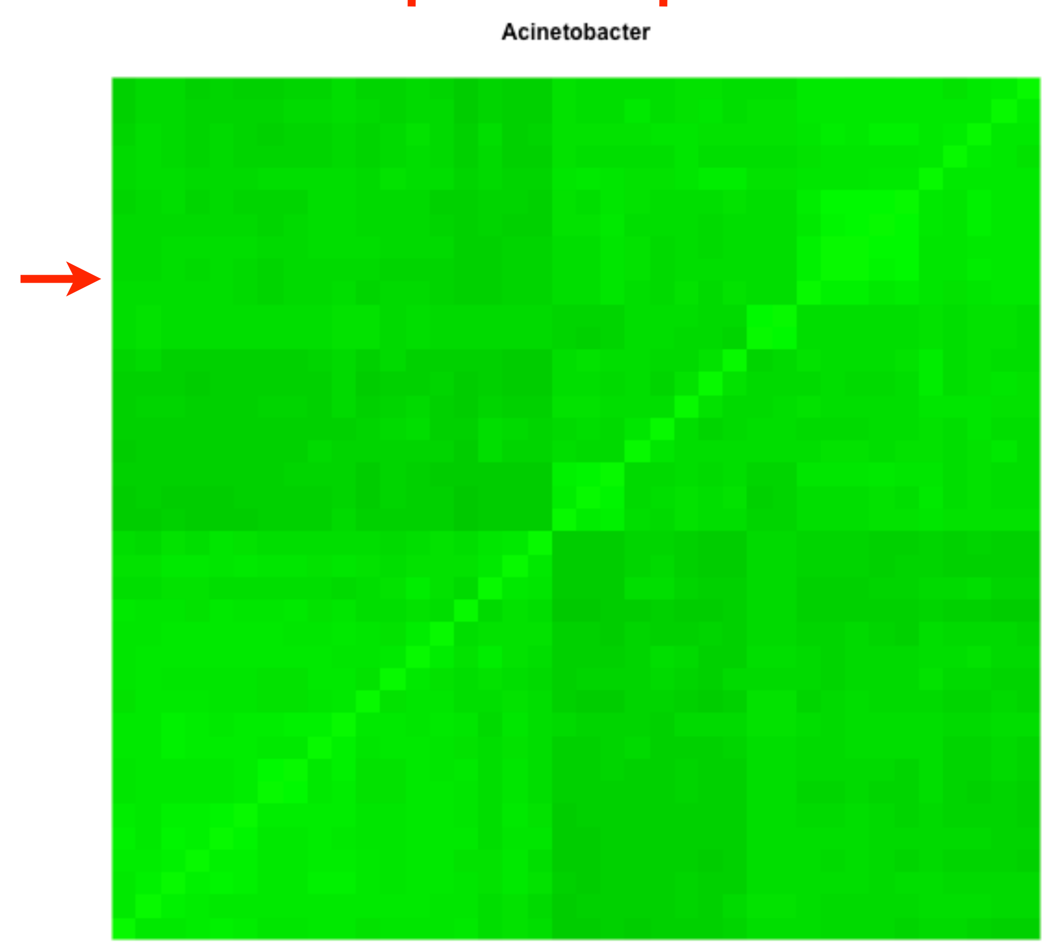

Mycoplasma
Mycoplasma spermatophilum ex.4724
Mycoplasma fermentans ex.4666
Mycoplasma mucosicanis ex.21296
Mycoplasma californicum ex.4635
Mycoplasma bovigenitalium ex.4629
Mycoplasma simbae ex.4722
Mycoplasma phocirhinis ex.4713
Mycoplasma lipophilum ex.4693
Mycoplasma hyopharyngis ex.4680
Mycoplasma opalescens ex.4706
Mycoplasma adleri ex.4618
Mycoplasma caviae ex.4640
Mycoplasma maculosum ex.4696
Mycoplasma leopharyngis ex.4691
Mycoplasma bovis ex.4631
Mycoplasma agalactiae ex.4619
Mycoplasma meleagridis ex.4698
Mycoplasma columbinum ex.4646
Mycoplasma iners ex.4686
Mycoplasma gallinarum ex.4669
Mycoplasma primatum ex.4717
Mycoplasma felifaucium ex.4663
Mycoplasma lipofaciens ex.4692
Mycoplasma columbinasale ex.4645
Mycoplasma gallinaceum ex.4668
Mycoplasma corogypsi ex.4649
Mycoplasma gallopavonis ex.4671
Mycoplasma buteonis ex.4634
Mycoplasma glycophilum ex.4674
Mycoplasma cynos ex.4653
Mycoplasma canis ex.4637
Mycoplasma bovirhinis ex.4630
Mycoplasma edwardii ex.4655
Mycoplasma mustelae ex.4704
Mycoplasma crocodyli ex.4652
Mycoplasma alligatoris ex.4622
Mycoplasma anatis ex.4624
Mycoplasma oxoniensis ex.4710
Mycoplasma citelli ex.4642
Mycoplasma columborale ex.4647
Mycoplasma sturni ex.4726
Mycoplasma leonicaptivi ex.4690
Mycoplasma felis ex.4665
Mycoplasma synoviae ex.4730
Mycoplasma pullorum ex.4718
Mycoplasma neophronis ex.23120
Mycoplasma spumans ex.4725
Mycoplasma falconis ex.4660
Mycoplasma canadense ex.4636
Mycoplasma alkalescens ex.4621
Mycoplasma auris ex.4628
Mycoplasma phocicerebrale ex.4712
Mycoplasma hominis ex.4679
Mycoplasma equirhinis ex.4659
Mycoplasma phocae ex.4714
Mycoplasma cloacale ex.4643
Mycoplasma anseris ex.4625
Mycoplasma orale ex.4707
Mycoplasma indiense ex.4685
Mycoplasma hyosynoviae ex.4683
Mycoplasma faucium ex.4662
Mycoplasma buccale ex.4633
Mycoplasma salivarium ex.4721
Mycoplasma subdolum ex.4728
Mycoplasma gateae ex.4672
Mycoplasma arthritidis ex.4627
Mycoplasma gypis ex.4675
Mycoplasma testudineum ex.8725
Mycoplasma agassizii ex.4620
Mycoplasma mobile ex.4701
Mycoplasma equigenitalium ex.4658
Mycoplasma elephantis ex.4656
Mycoplasma iguanae ex.9978
Mycoplasma hyorhinis ex.4682
Mycoplasma molare ex.4702
Mycoplasma lagogenitalium ex.4689
Mycoplasma sualvi ex.4727
Mycoplasma moatsii ex.4700
Mycoplasma cricetuli ex.4651
Mycoplasma collis ex.4644
Mycoplasma neurolyticum ex.4705
Mycoplasma conjunctivae ex.4648
Mycoplasma bovoculi ex.4632
Mycoplasma hyopneumoniae ex.4681
Mycoplasma dispar ex.4654
Mycoplasma flocculare ex.4667
Mycoplasma ovipneumoniae ex.4708
Mycoplasma pulmonis ex.4719
Mycoplasma arginini ex.4626
Mycoplasma pirum ex.4715
Mycoplasma alvi ex.4623
Mycoplasma imitans ex.4684
Mycoplasma microti ex.4699
Mycoplasma mycoides ex.4616
Mycoplasma capricolum capripneumoniae ex.4639
Mycoplasma mycoides capri ex.4617
Mycoplasma leachii ex.14347
Mycoplasma capricolum ex.4638
Mycoplasma yeatsii ex.4734
Mycoplasma cottewii ex.4650
Mycoplasma putrefaciens ex.4720
0.6
0.4
0.2
0

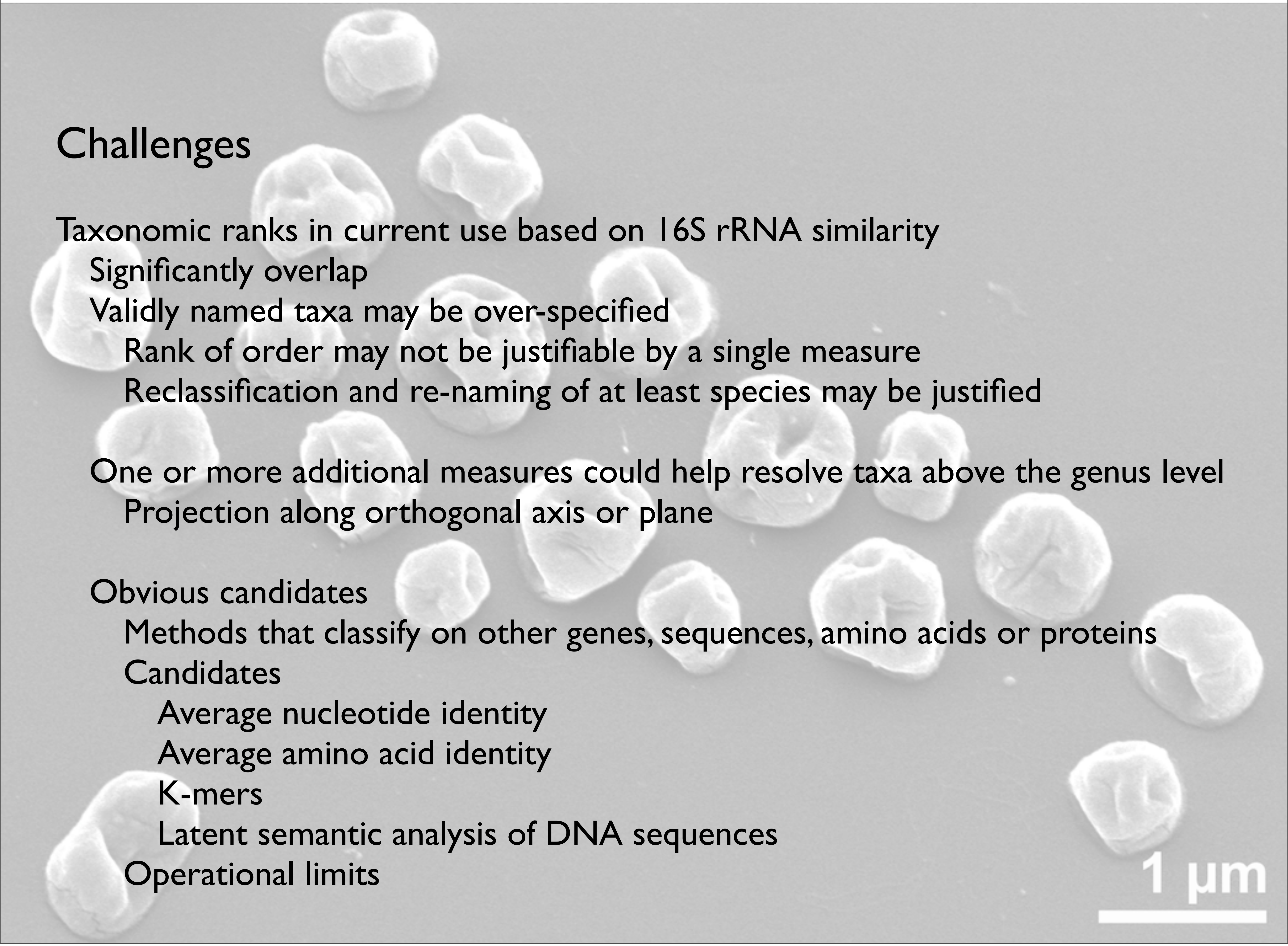

Mycoplasma
Frequency
0
500
1000
1500
2000
2500
0.00
0.05
0.10
0.15
0.20
0.25
0.30


Mycoplasma


Acinetobacter
Acinetobacter puyangensis ex.24438
Acinetobacter soli ex.14452
Acinetobacter baumannii ex.2767
Acinetobacter venetianus ex.14367
Acinetobacter indicus ex.23597
Acinetobacter seifertii ex.26515
Acinetobacter nosocomialis ex.22117
Acinetobacter pittii ex.22118
Acinetobacter calcoaceticus ex.2766
Acinetobacter brisouii ex.22116
Acinetobacter guillouiae ex.17894
Acinetobacter bereziniae ex.17893
Acinetobacter variabilis ex.26419
Acinetobacter guangdongensis ex.25853
Acinetobacter rudis ex.22713
Acinetobacter gerneri ex.2770
Acinetobacter baylyi ex.2768
Acinetobacter apis ex.25975
Acinetobacter nectaris ex.23996
Acinetobacter boissieri ex.23997
Acinetobacter tandoii ex.2779
Acinetobacter bouvetii ex.2769
Acinetobacter grimontii ex.2771
Acinetobacter gandensis ex.26133
Acinetobacter parvus ex.2776
Acinetobacter junii ex.2774
Acinetobacter radioresistens ex.2777
Acinetobacter ursingii ex.2782
Acinetobacter harbinensis ex.25425
Acinetobacter lwoffii ex.2775
Acinetobacter pakistanensis ex.26276
Acinetobacter bohemicus ex.26207
Acinetobacter johnsonii ex.2773
Acinetobacter haemolyticus ex.2772
Acinetobacter gyllenbergii ex.13190
Acinetobacter beijerinckii ex.13189
Acinetobacter kookii ex.24868
Acinetobacter schindleri ex.2778
0.12
0.1
0.08
0.06
0.04
0.02
0


Acinetobacter
Frequency
0
50
100
150
200
250
300
0.00
0.02
0.04
0.06
0.08
0.10


Acinetobacter



Challenges

- Taxonomic ranks in current use based on 16S rRNA similarity
  - Significantly overlap
  - Validly named taxa may be over-specified
    - Rank of order may not be justifiable by a single measure
    - Reclassification and re-naming of at least species may be justified

  - One or more additional measures could help resolve taxa above the genus level
    - Projection along orthogonal axis or plane

  - Obvious candidates
    - Methods that classify on other genes, sequences, amino acids or proteins
    - Candidates
      - Average nucleotide identity
      - Average amino acid identity
      - K-mers
      - Latent semantic analysis of DNA sequences
    - Operational limits

Experimental plan

- Assemble a reasonable sized collections of type strains with sequenced genomes
  - Strains representing the type (genus) of 170/202 validly named orders
    - Varying quality (finished, high quality draft, draft)
    - Sequenced/assembled by different groups
    - Collaborators perform analysis using preferred method for all strains
    - Visualized against 16S sequence similarly data for same strains
- NamesforLife - Charles Parker
- DOE JGI - Nikos Kyrpides, Neha Varghese
- ORNL - Dave Ussery, Miriam Land, Se-Ran Jun, Intewat Nookaew, Visanu Wanchai
- Univ. Nebraska, Lincoln - Khalid Sayood, Ufuk Ubantagalu

Trends between gene content and genome size in prokaryotic species with larger genomes

Konstantinos T. Konstantinidis*† and James M. Tiedje*†‡§

*Center for Microbial Ecology and Departments of †Crop and Soil Sciences and ‡Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824-1325

Contributed by James M. Tiedje, December 24, 2003

Trends between gene content and genome size

Genomic insights that advance the species definition for prokaryotes

Konstantinos T. Konstantinidis*† and James M. Tiedje*†‡§

*Center for Microbial Ecology, and Departments of †Crop and Soil Sciences and ‡Microbiology and Molecular Genetics, Michigan State University, East Lansing, MI 48824

Contributed by James M. Tiedje, December 24, 2004

To help advance the species definition for prokaryotes, we have compared the gene content of 70 closely related and fully se-

informative, with respect to the species definition, because it concerns genes that largely determine the organism's phenotype.

niversity,

Trends between gene content and genome size

Genomic insights that advance the species definition

JOURNAL OF BACTERIOLOGY, Sept. 2005, p. 6258–6264 Vol. 187, No. 18
0021-9193/05/$08.00+0 doi:10.1128/JB.187.18.6258–6264.2005

Towards a Genome-Based Taxonomy for Prokaryotes

Konstantinos T. Konstantinidis[1,2] and James M. Tiedje[1,2,3]*

Center for Microbial Ecology[1] and Departments of Crop and Soil Sciences[2] and Microbiology and Molecular Genetics,[3] Michigan State University, East Lansing, Michigan

Sunday, January 31, 16

Trends between gene content and genome size

Genomic insights that advance the species definition

Journal of Bacteriology, Sept. 2005, p. 6258–6264 Vol. 187, No. 18
0021-9193/05/$08.00+0 doi:10.1128/JB.187.18.6258–6264.2005

Towards a Genome-Based Taxonomy for Prokaryotes

Applied and Environmental Microbiology, Nov. 2006, p. 7286–7293 Vol. 72, No. 1
0099-2240/06/$08.00+0 doi:10.1128/AEM.01398-06

Toward a More Robust Assessment of Intraspecies Diversity, Using Fewer Genetic Markers▽†

Konstantinos T. Konstantinidis,*¶ Alban Ramette,§¶ and James M. Tiedje
Center for Microbial Ecology, Michigan State University, East Lansing, Michigan

Received 16 June 2006/Accepted 5 September 2006

Trends between gene content and genome size

Genomic insights that advance the species definition

JOURNAL OF BACTERIOLOGY, Sept. 2005, p. 6258–6264 Vol. 187, No. 18
0021-9193/05/$08.00+0 doi:10.1128/JB.187.18.6258–6264.2005

Towards a Genome-Based Taxonomy for Prokaryotes

APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Nov. 2006, p. 7286–7293 Vol. 72, No. 1
0099-2240/06/$08.00+0 doi:10.1128/AEM.01398-06

Toward a More Robust Assessment of Intraspecies Diversity,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2006) **361**, 1929–1940
doi:10.1098/rstb.2006.1920
Published online 11 October 2006

The bacterial species definition in the genomic era

Konstantinos T. Konstantinidis*, Alban Ramette† and James M. Tiedje

Center for Microbial Ecology, Michigan State University, East Lansing, MI 48824, USA

Trends between gene content and genome size

Genomic insights that advance the species definition

Journal of Bacteriology, Sept. 2005, p. 6258–6264 Vol. 187, No. 18
0021-9193/05/$08.00+0 doi:10.1128/JB.187.18.6258–6264.2005

Towards a Genome-Based Taxonomy for Prokaryotes

Applied and Environmental Microbiology, Nov. 2006, p. 7286–7293 Vol. 72, No. 1
0099-2240/06/$08.00+0 doi:10.1128/AEM.01398-06

Toward a More Robust Assessment of Intraspecies Diversity,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2006) **361**, 1929–1940
doi:10.1098/rstb.2006.1920
Published online 11 October 2006

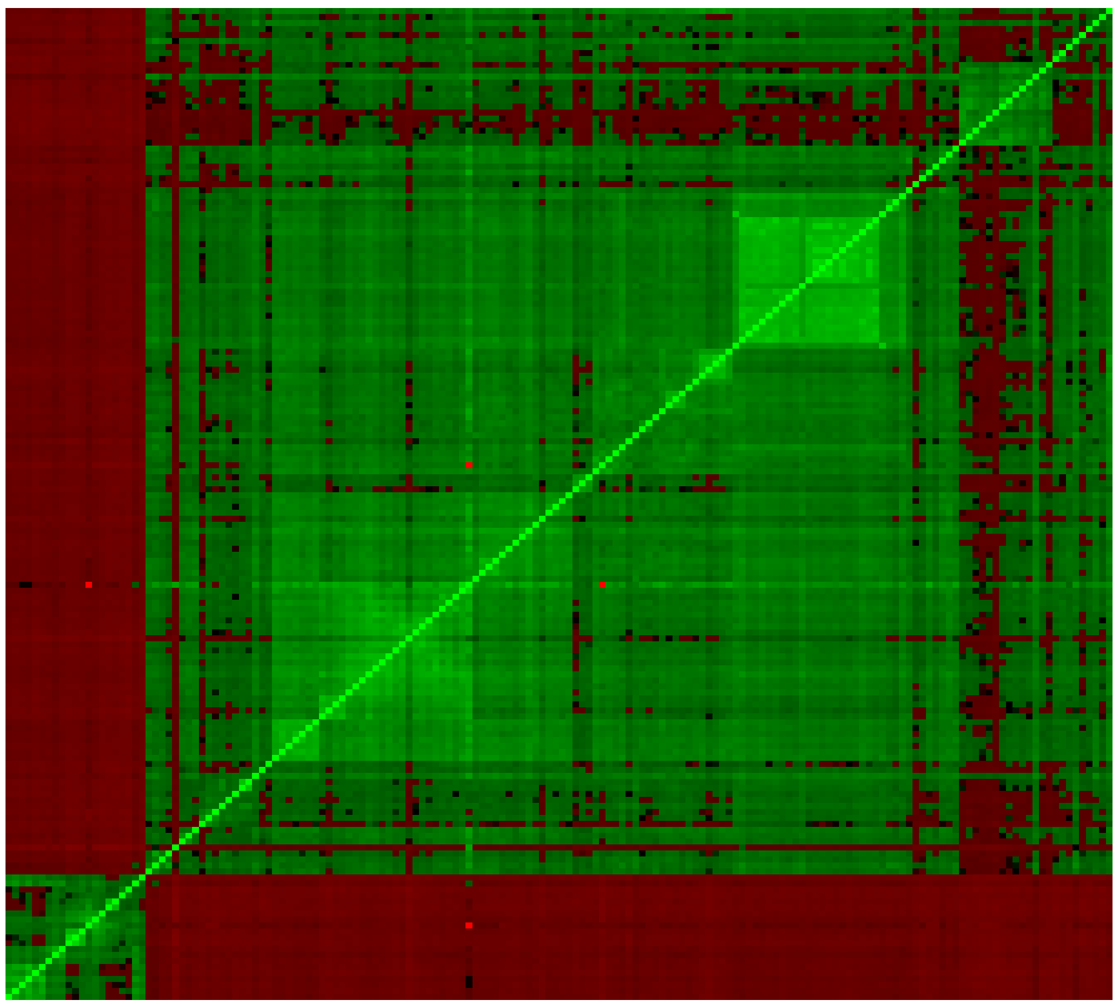
The bacterial species definition in the genomic era

Standards in Genomic Sciences (2009) **2**:117-134 DOI:10.4056/sigs.531120

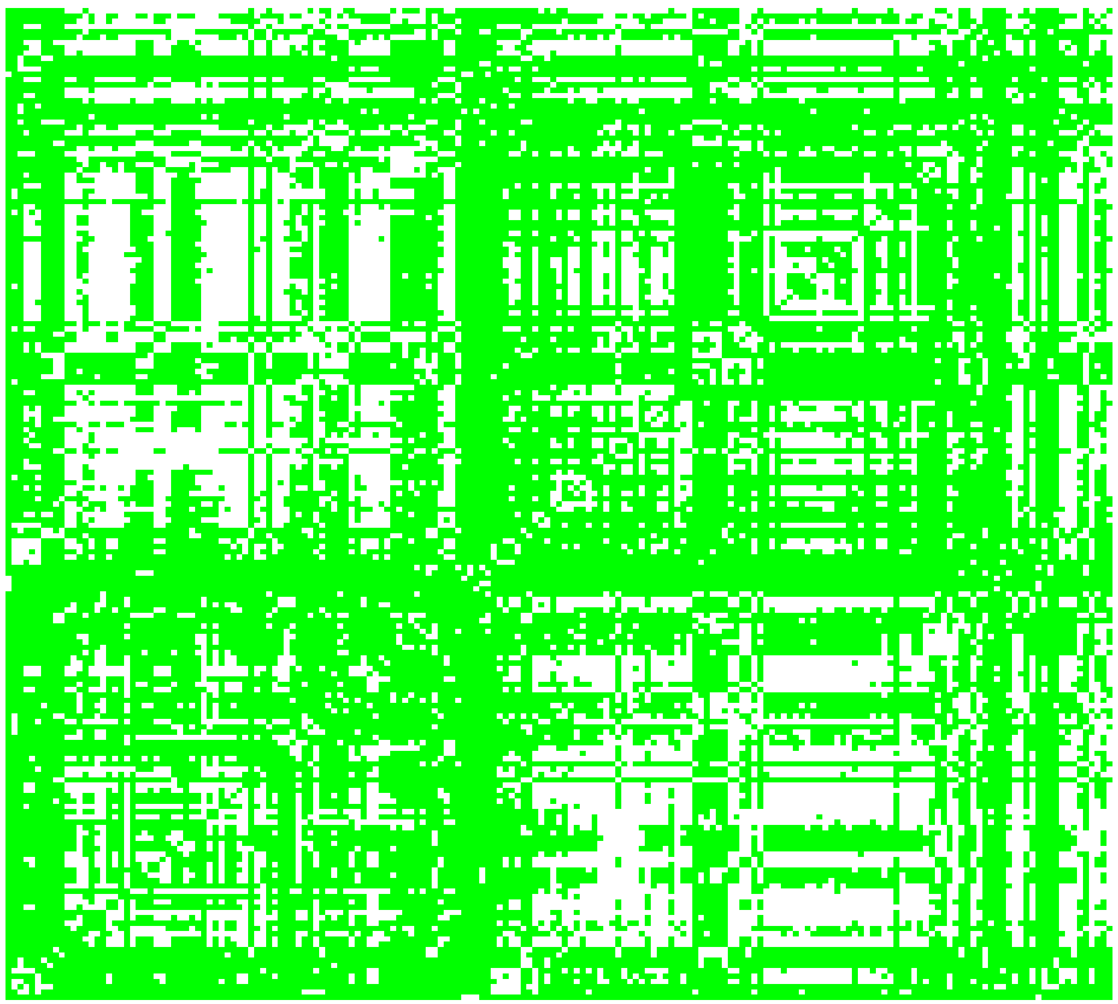
Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison

Alexander F. Auch[1], Mathias von Jan[2], Hans-Peter Klenk[2*], Markus Göker[2]

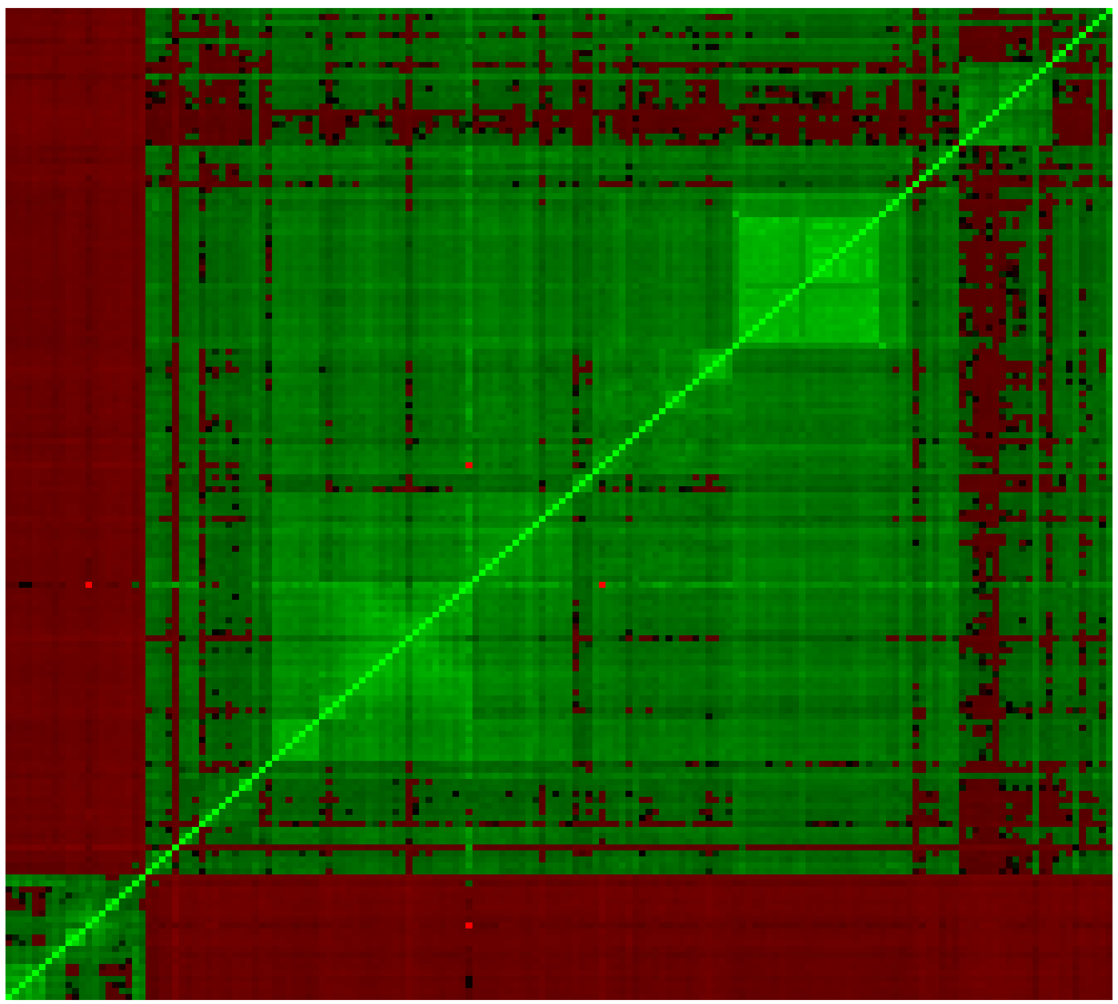
16S rRNA sequence subset



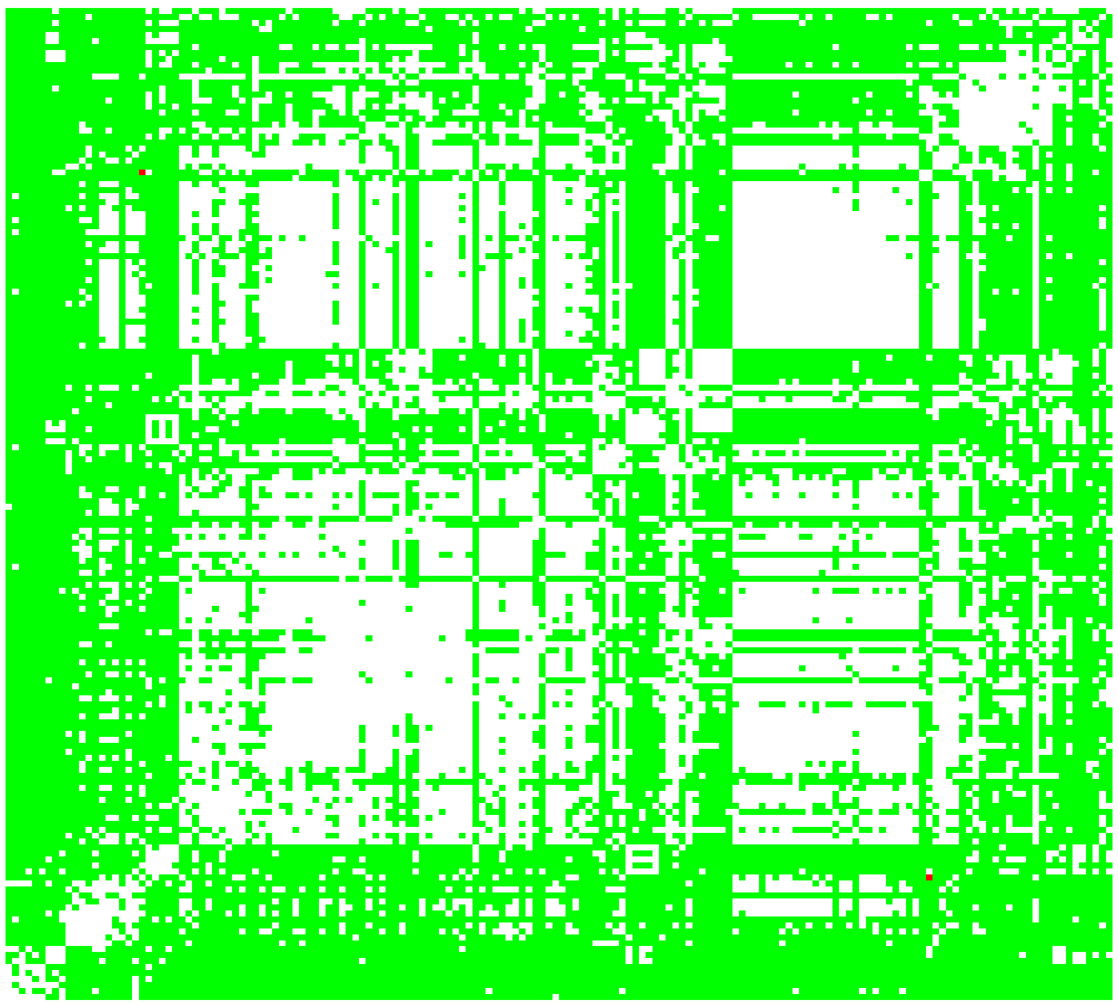
ANIa, JGI, n=187



16S rRNA sequence subset



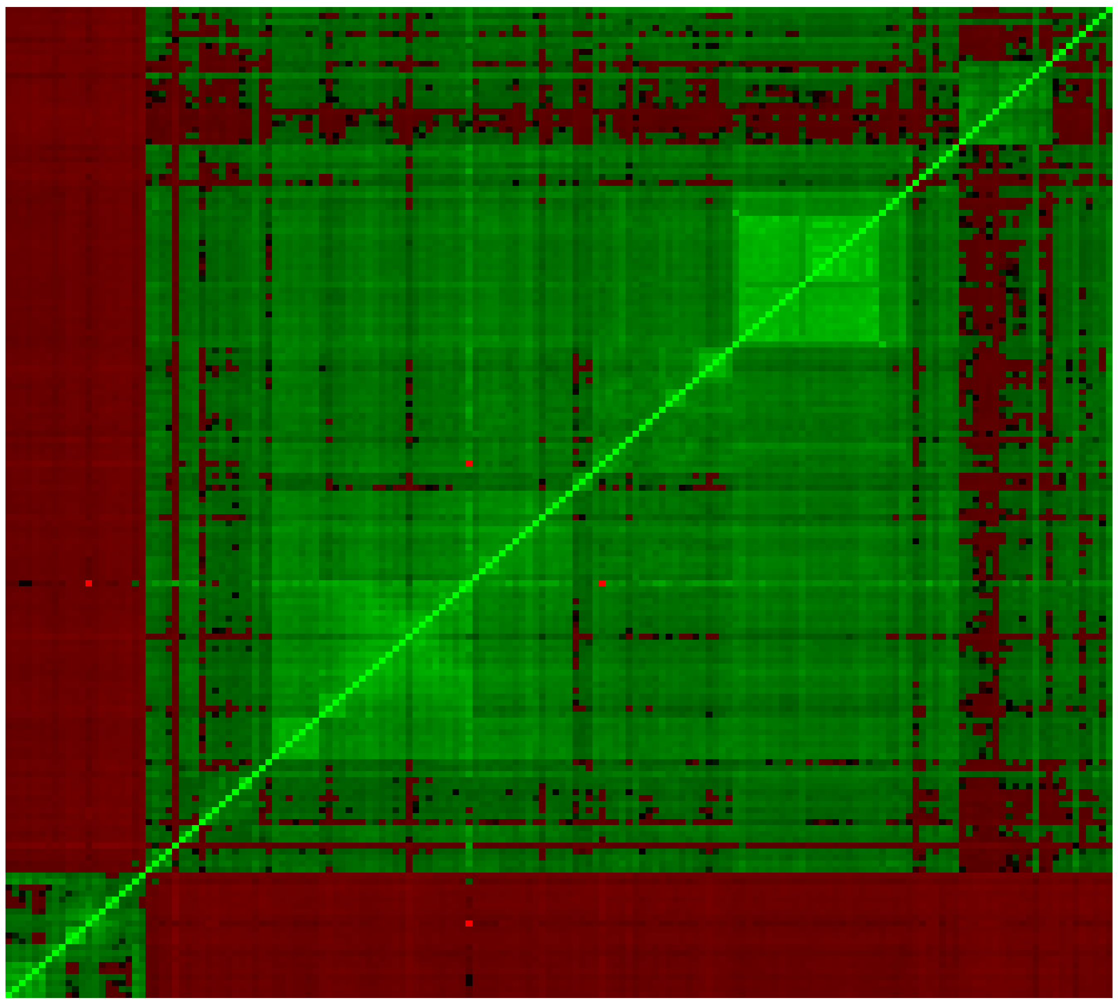
ANI, Oak Ridge National Laboratories



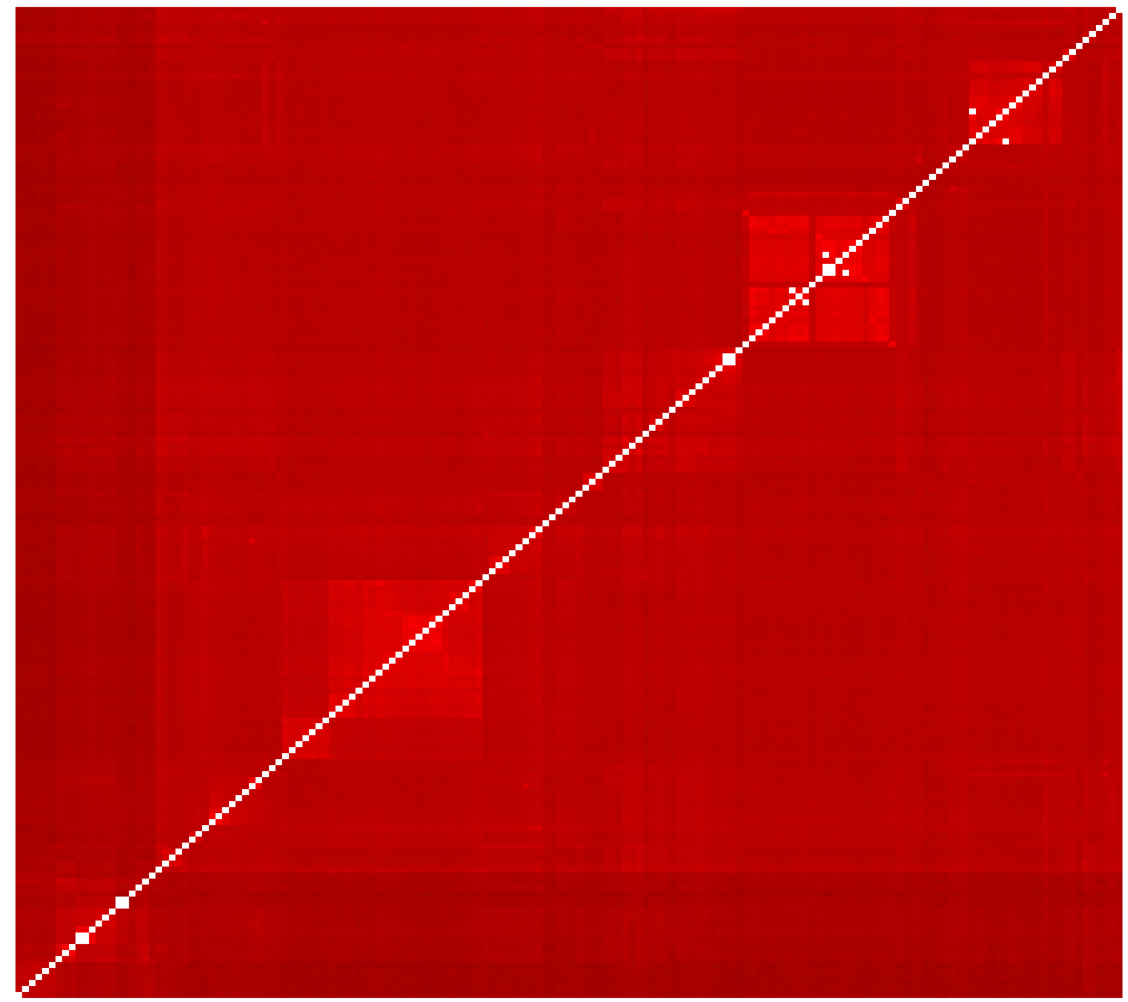
166 type strains, with conditioning of AAI signal

cor = -0.46807

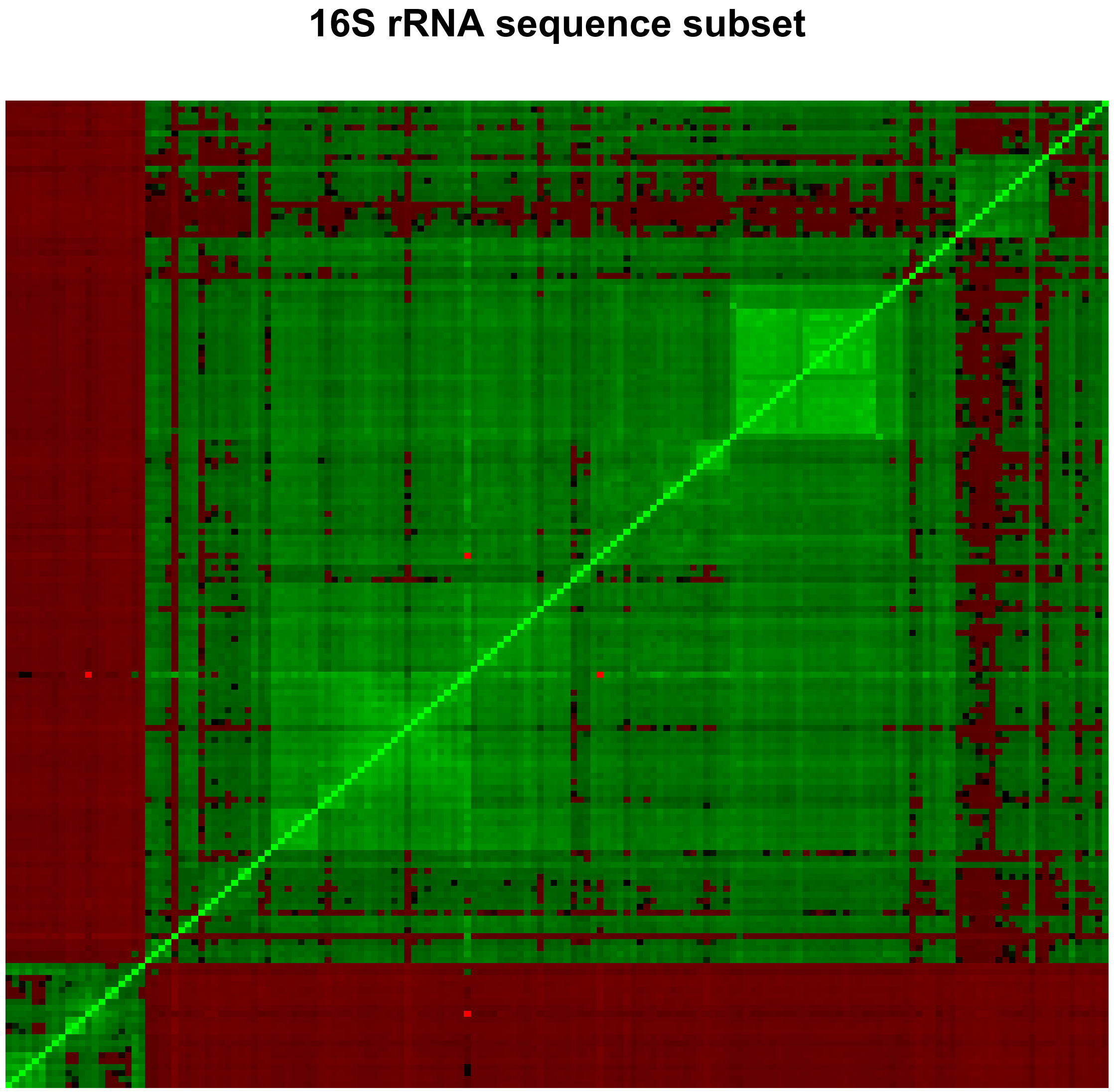
16S rRNA sequence subset

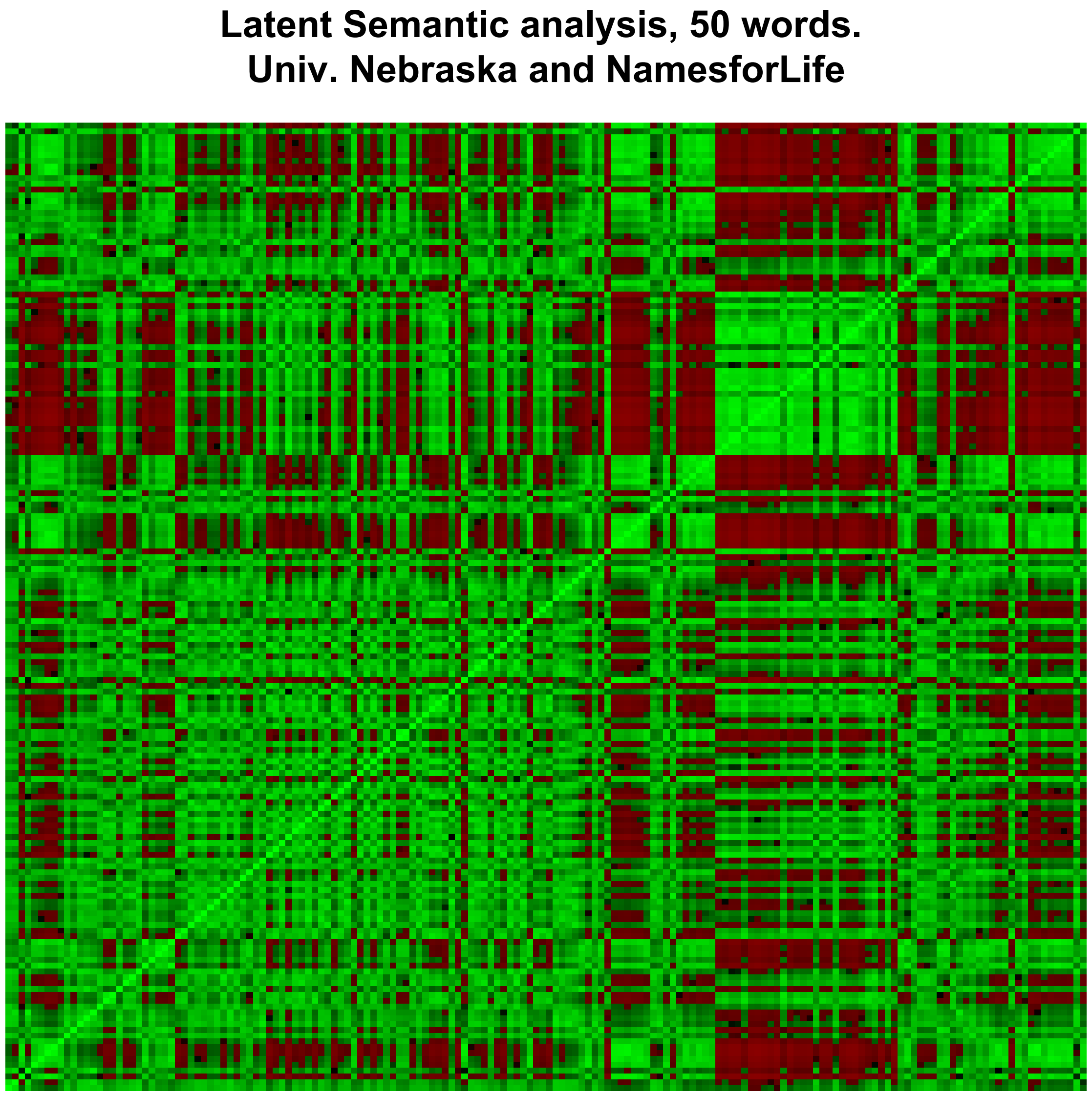


AAI, Oak Ridge National Laboratories



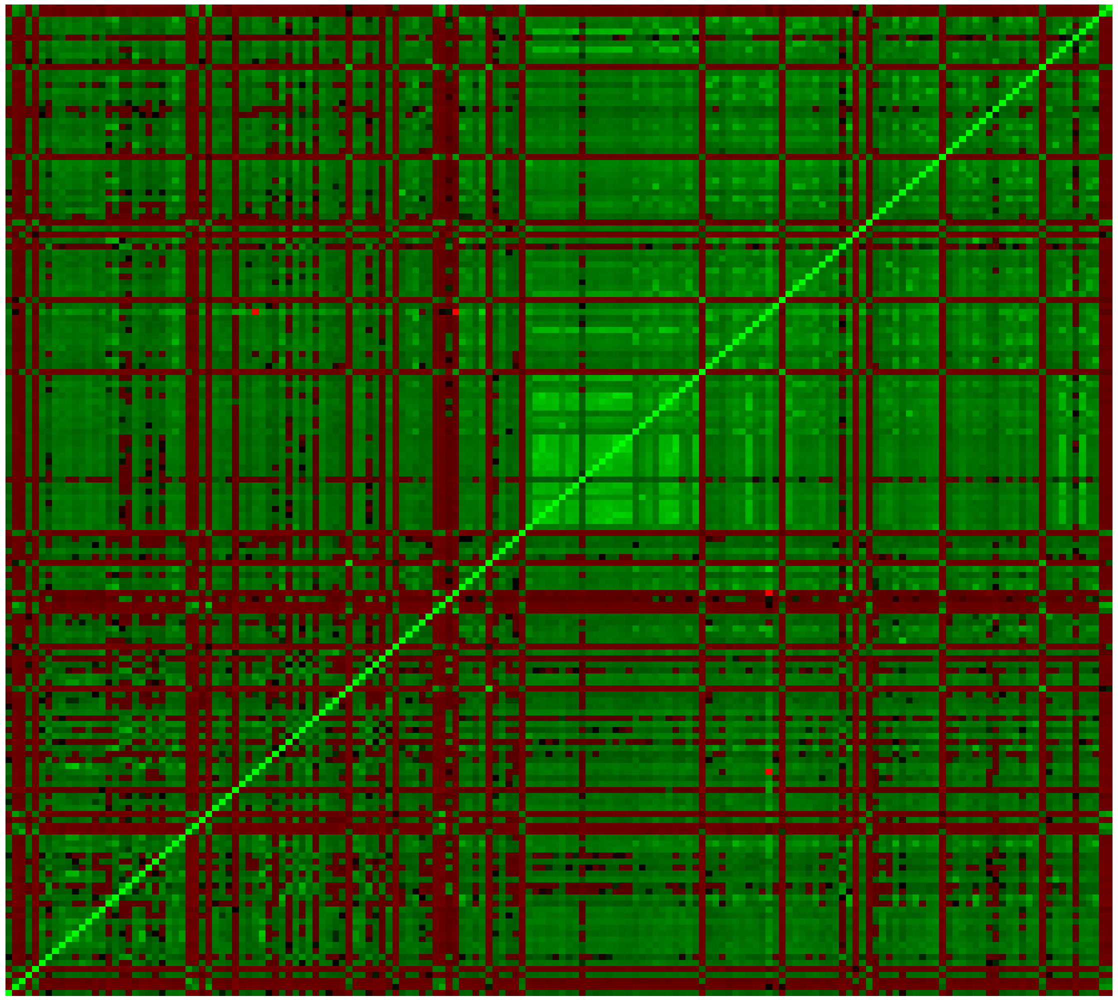
166 type strains, with conditioning of AAI signal
cor = -0.6631114



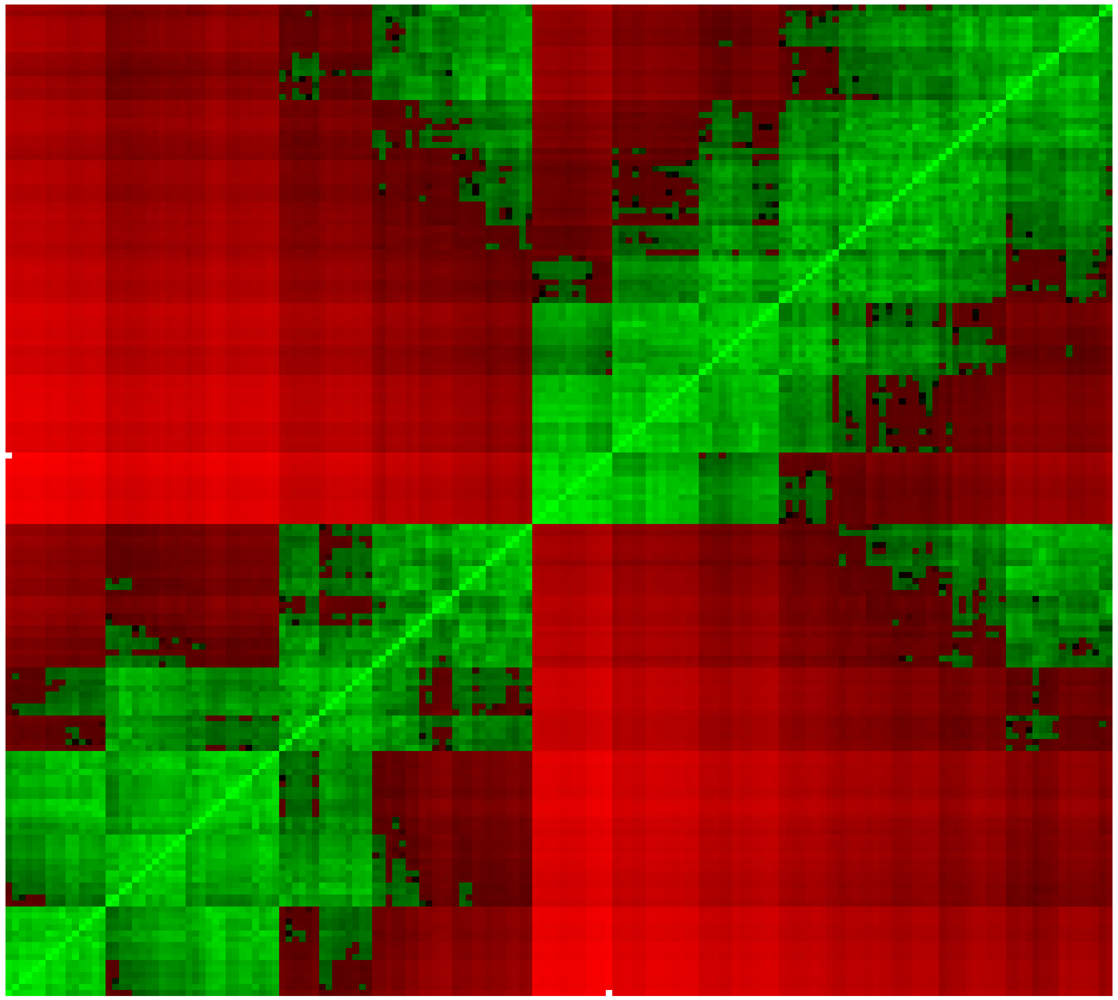


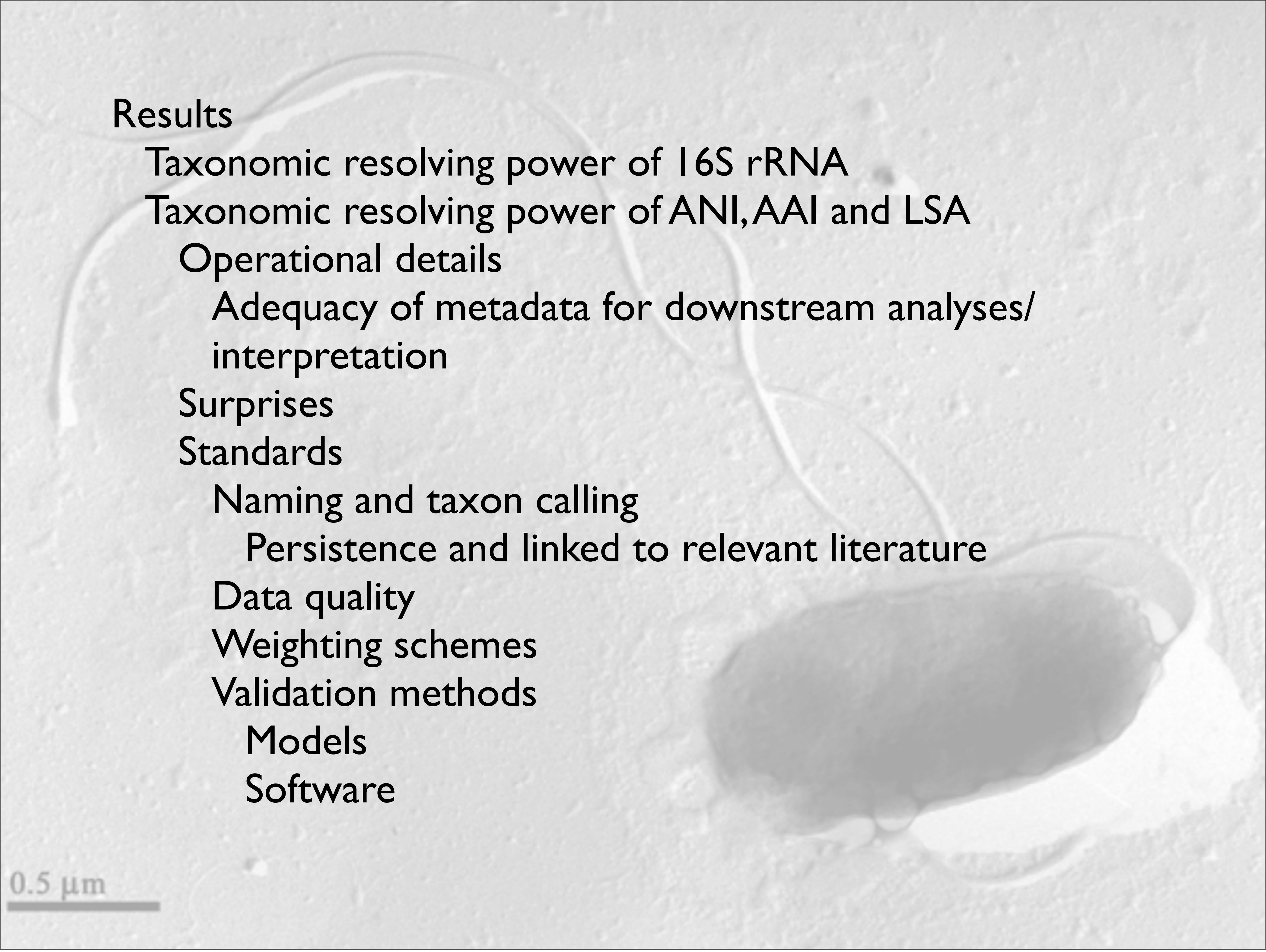
166 type strains, with conditioning of LSA signal
cor = -0.02849482

Type strains reordered according to LSA classification

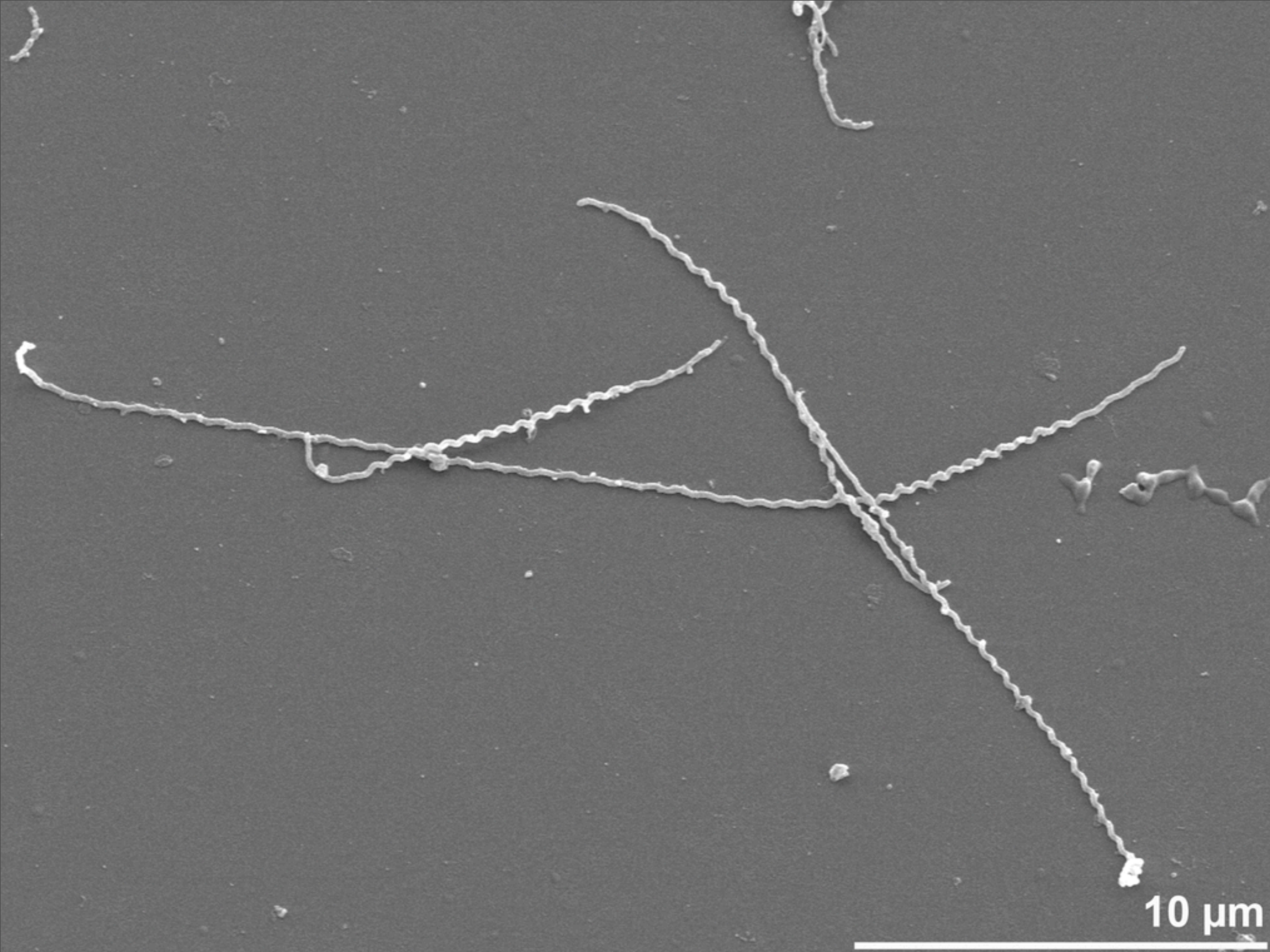


Type strain genomes reordered according to LSA





- Results
  - Taxonomic resolving power of 16S rRNA
  - Taxonomic resolving power of ANI, AAI and LSA
    - Operational details
      - Adequacy of metadata for downstream analyses/ interpretation
    - Surprises
    - Standards
      - Naming and taxon calling
        - Persistence and linked to relevant literature
      - Data quality
      - Weighting schemes
      - Validation methods
        - Models
        - Software


10 µm

NamesforLife™
Bringing meaning to life ...

Charles T. Parker
Vo Phan Chuong
Dorothea Taylor
Sarah Wigley
Nicole Osier
Grace Rodriguez

INTERNATIONAL COMMITTEE ON SYSTEMATICS OF PROKARYOTES

Brian Tindall
Aharon Oren
David Labeda

JGI
DOE JOINT GENOME INSTITUTE

Nikos Kyrpides
Neha Varghese

SOCIETY FOR GENERAL MICROBIOLOGY

Ron Fraser
Robin Dunford
Karen Rowlett
Sue Andrews
Rachel Walker
Leighton Chipperfield

ATCC

Timothy G. Lilburn

USDA ARS

David Labeda

SIGS Standards in Genomic Sciences

Oranmiyan Nelson
Editorial Board

MICHIGAN STATE UNIVERSITY

Kevin Petersen
Nenad Krdzavac

UNIVERSITY OF Nebraska Lincoln

Khalid Sayood
Ufuk Nalbantoglu
Sam Way

DTU

Dave Ussery

OAK RIDGE National Laboratory

Dave Ussery
Miriam Land
Se-Ran Jun
intewat Nookaew
Visanu Wanchai

10 µm

Ed Pence

Norman Paskin

Gerry Roston
Kurt Riegger

This work was funded through the Small Business technology transfer program of the United States Department of Energy under grants number DE-FG02-07ER86321 and DE-SC0006191. Funding for business development was provided through grants and loans from the Michigan Economic Development Corporation and the Michigan Universities Commercialization Initiative. NamesforLife semantic resolution technology is covered under US Patent 7,925,444 B2. The SOSCC systems and methods is covered under US Patent 8,036,997. Semiotic fingerprinting is covered under US 8,903,824. Semantic tagging, indexing, equivalency mapping, and latent semantic analysis of molecular sequence data are the subjects of pending US and EPO patent applications.

1 µm