Metagenomics-Enabled Understanding of Metal-Reducing Communities at ORIFC

James M. Tiedje. Terry Marsh, Jizhong Zhou, Shelly Kelly, Ken Kemner

OAK RIDGE NATIONAL LABORATORY

Components

Community Structure Shifts: in response to biostimulation, and oxidation

- Stable isotope probing (SIP): the active members
- 454FLX 16S rRNA sequencing: depth & replication
- CCA with biogeochemical features

Desulfitobacter, a Gram positive metal reducer

Functional gene targets: eventual markers of activity

- Strategies for recovery of gene and operons

The metagenome of an Area 3 microbe

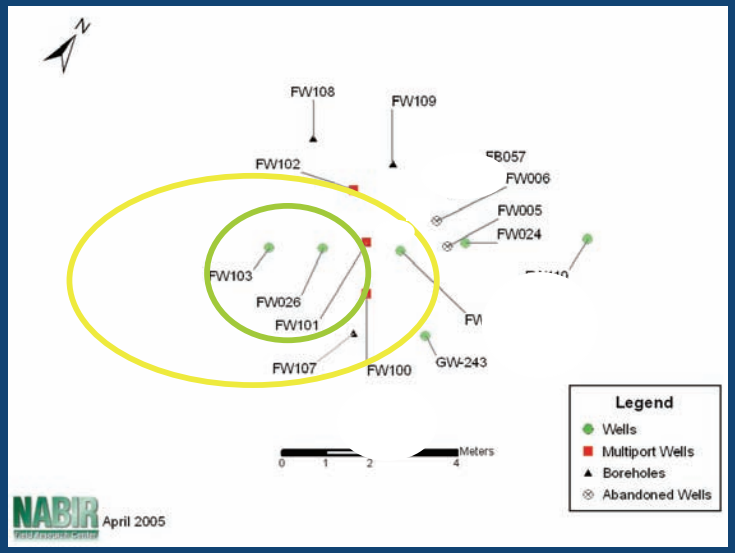
Dominant bacteria

Family	Dominant Genus	Relative abundance of bacteria (% of clones detected) in wells						
		Treatment zone						Before treatment
		104	101-2	101-3	102-2	102-3	026	**106**
Desulfovibrionaceae	*Desulfovibrio* [a.b.d]	7	15	5	4	12	5	**0**
Geobacteraceae	*Geobacter* [a.b.]	2	1	1	11	1	1	**0**
Rhodocyclaceae	*Ferribacterium* [b.c.]	12	6	38	10	17	18	**0**
Hydrogenophilaceae	*Thiobacillus* [c.e.]	5	27	0	1	4	5	**0**
Acidobacteraceae	*Geothrix* [b.c.]	12	7	10	4	10	16	**0**
Oxalobacteraceae	*Duganella*	9	10	2	2	11	2	**7**
Xhantomonadaceae	*Rhodanobacter*	6	2	0	5	5	0	**84**
Commanonadaceae	*Acidovorax* [a.c.]	2	1	2	1	2	6	**0**
Sphingomonadaceae	*Sphingomonas* [c.]	6	0	1	2	2	1	**0**
other families		39	31	41	60	36	46	**8**

a. uranium reducers b. iron reducers c. nitrate reducers d. sulfate reducers e. iron oxidizers

How do environmental variables shape the microbial community?

- Geochemical parameters (pH, U, sulfate, nitrate, etc.)
- Microbial community analysis
  - Massively parallel sequencing (454 Life Sciences)
  - >57,000 sequences from treatment zone
  - Still not fully sampled!
  - Formed 1,868 OTUs (95% similarity cutoff)
- Canonical correspondence analysis
  - Shows ecological preference of each species



How does biostimulation work?

Injection well FW104

Sampling well FW101

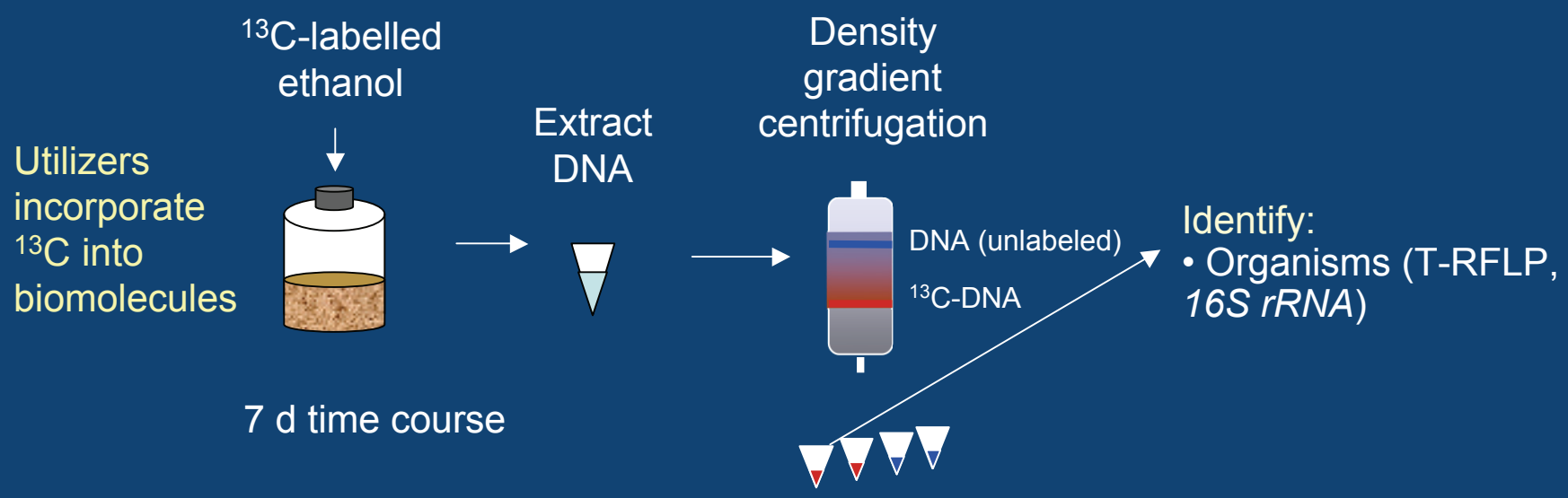
Extraction well FW026

Investigate dynamics of:

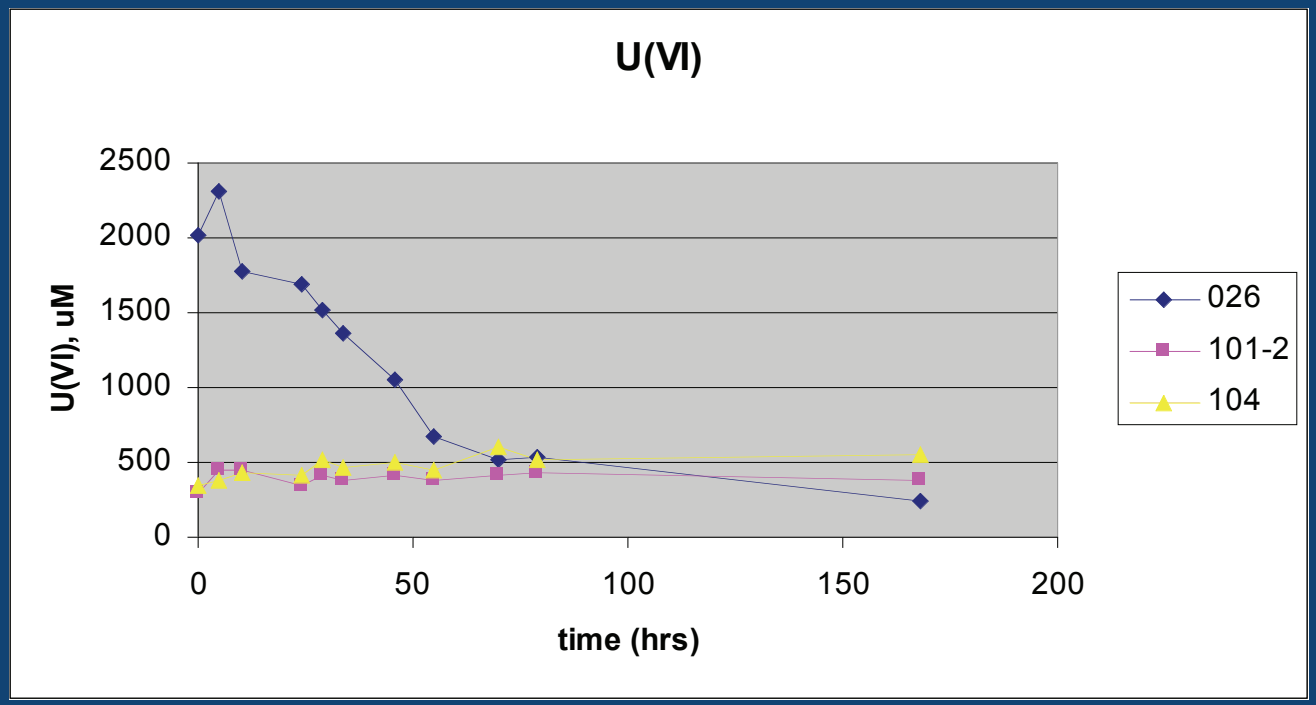
- Ethanol utilization
- Reduction of
  - U(VI)
  - nitrate
  - sulfate

Identify bacteria that derive C from ethanol

Stable isotope probing (SIP)



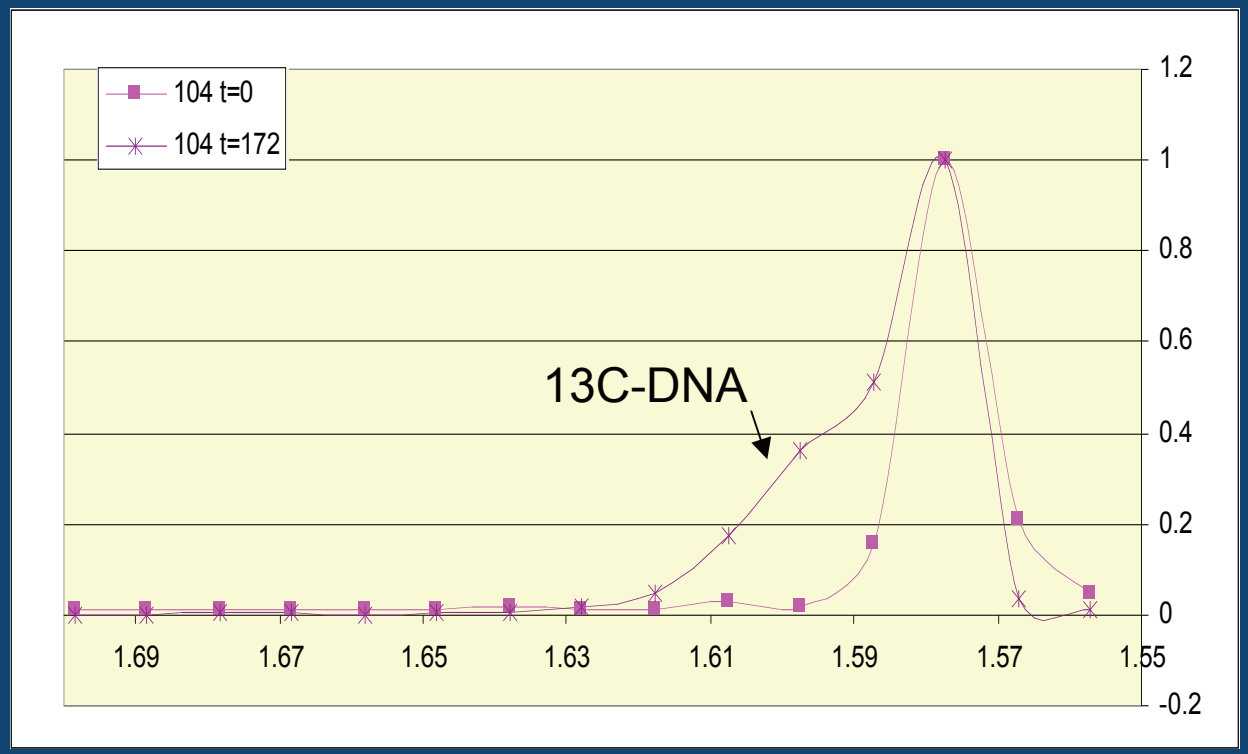
U(VI) reduction in microcosms



101-2, 104:

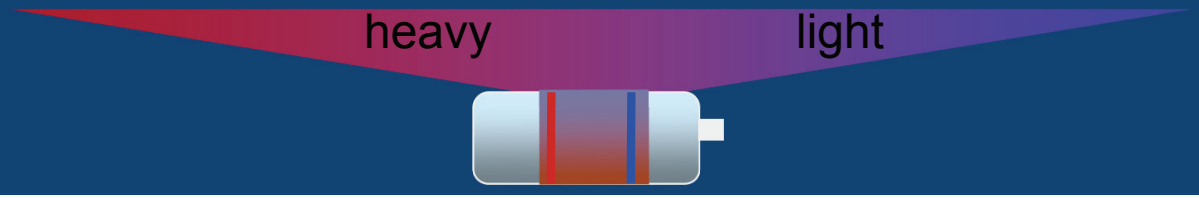
- Kinetics?
- U(VI) Dropped after 30 d
- pH and U(VI) desorption?

Use of ethanol for cell growth (Injection well FW104)



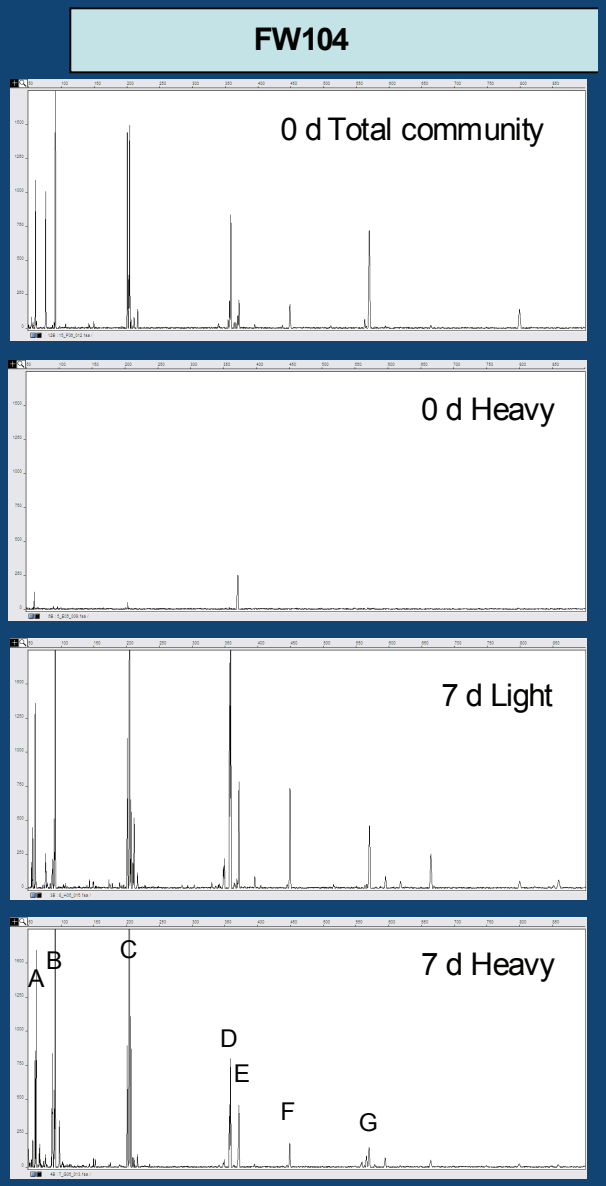
16S rRNA gene copies (ratio of max detected)

Buoyant density (g/ml)



- Growth on ^{13}C occurred between 1-7 days incubation

Community fingerprinting (T-RFLP) of SIP fractions



Terminal restriction fragment length polymorphism

A. *Duganella**, *Ferribacterium*, *Sterolibacterium*, *Thiobacillus*;

B. *Desulfovibrio**, *Sphingomonas**, *Geobacter, Geothrix, Anaeromyxobacter;*

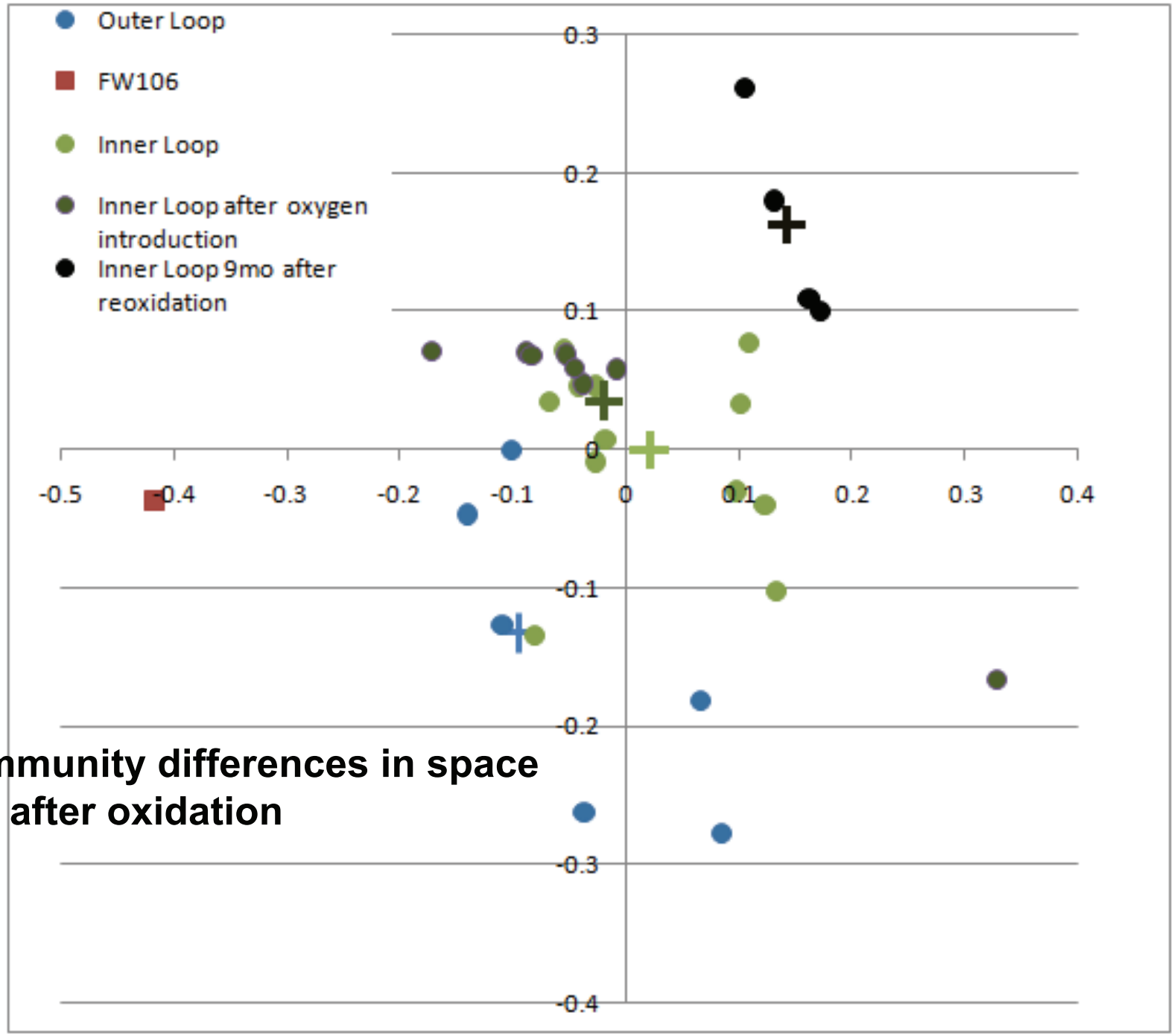
C. *Ferribacterium***, *Rhodanobacter***, *Rhodoferax**, *Anaerolinea**, *Acidovorax** ;

D. *Gemmatimonas*,

E. *Geothrix***,

F. *Thiobacillus**,

G. *Thiobacillus*.



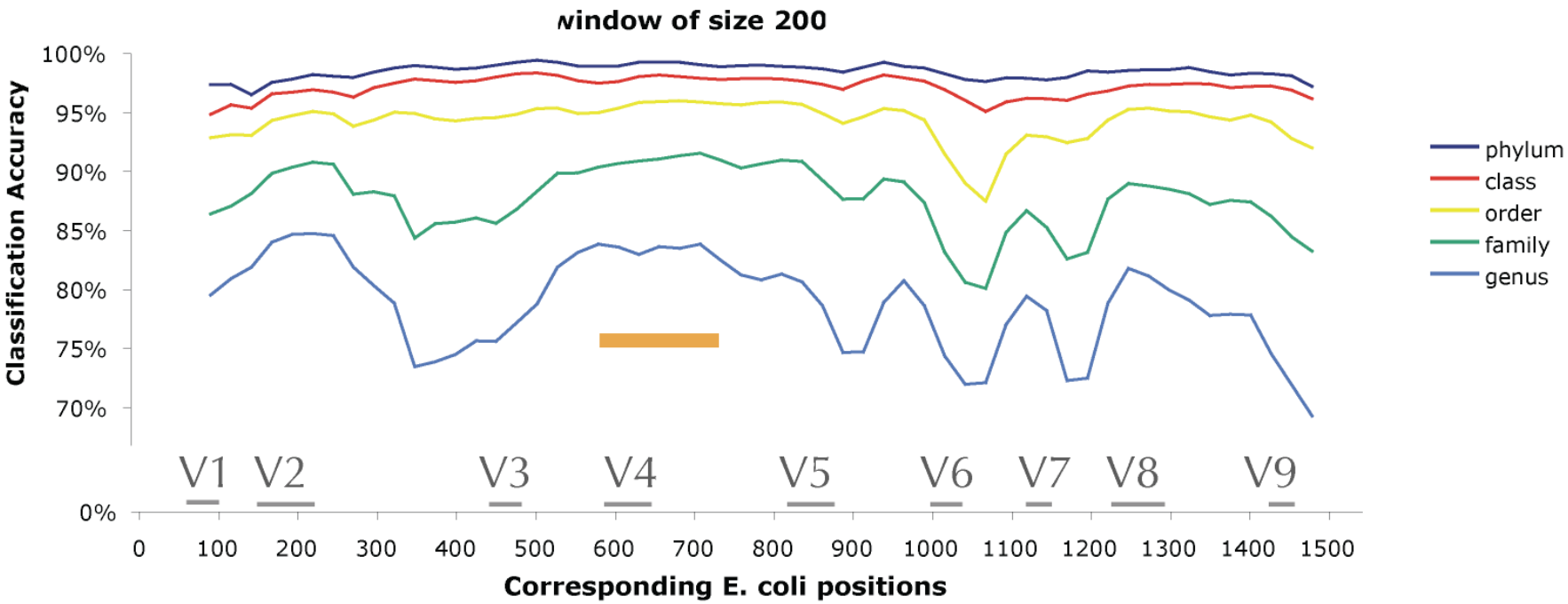
Community differences in space and after oxidation

Strategies for reducing complexity for metagenomics: targeted metagenomics

- Tagged (key) for 16S rRNA genes (classify, cluster)
- Tagged (key) for functional genes (resolve at finer scale, AND eventually relate to activity)
- Use above to develop probes to recover active, targeted genes to study their function
- Selected subsets of community DNA for sequencing
- Functional screening of community DNA clones
- A variety of other means to select subsets of community DNA

Classification Accuracy of 200 bp Regions

Region covered by FLX primers shown colored bar



Using 454FLX sequencing
By key/tag of amplified rRNA genes

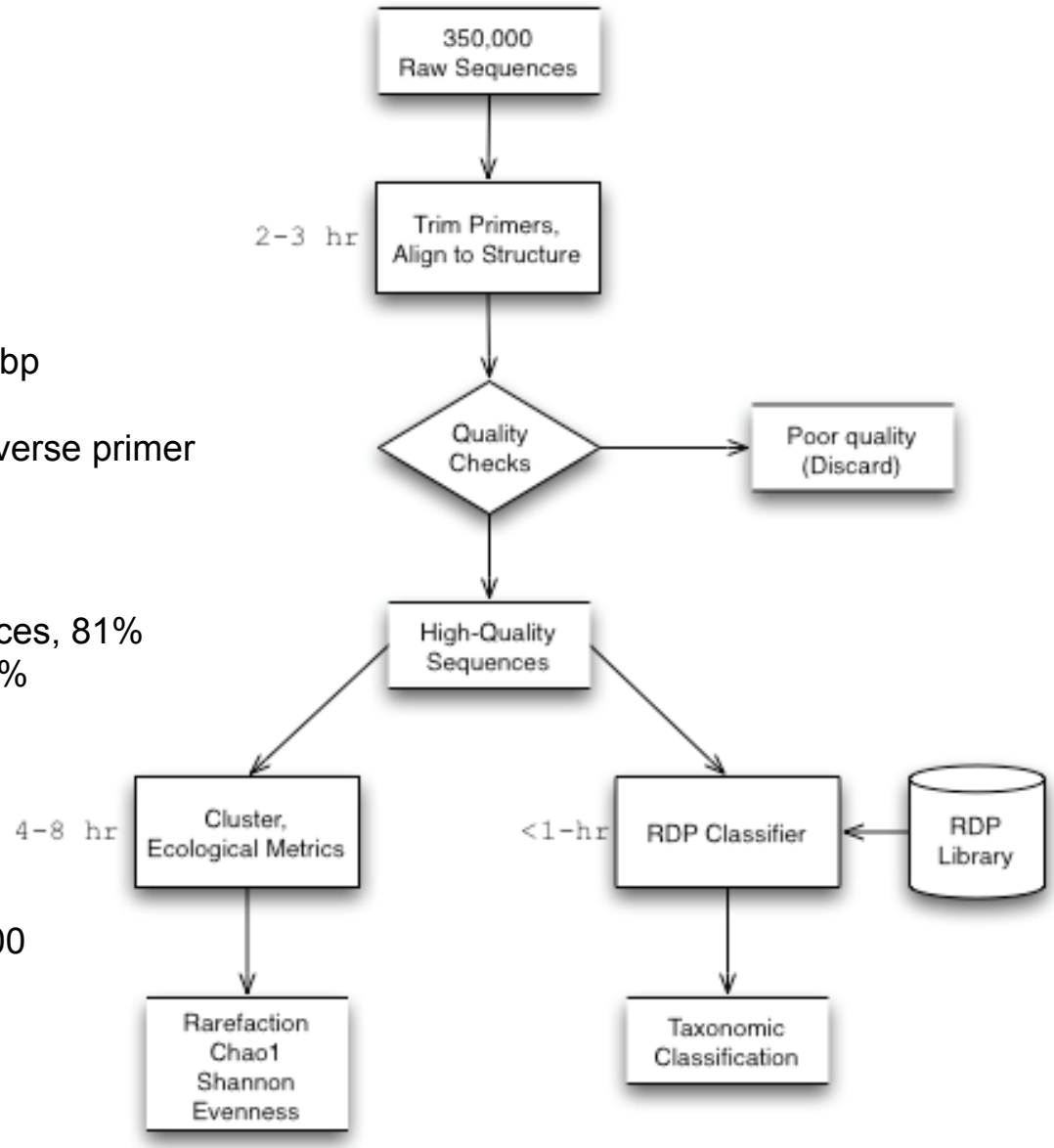
From Wang et. al., AEM, 2007

Our High Throughput Processing of 454 Data

Quality criteria
- no reads , 150 bp
- no Ns
- no errors in reverse primer

Quality sequences, 81%
Error rate, 0.20%

MLC of 200,000 sequences



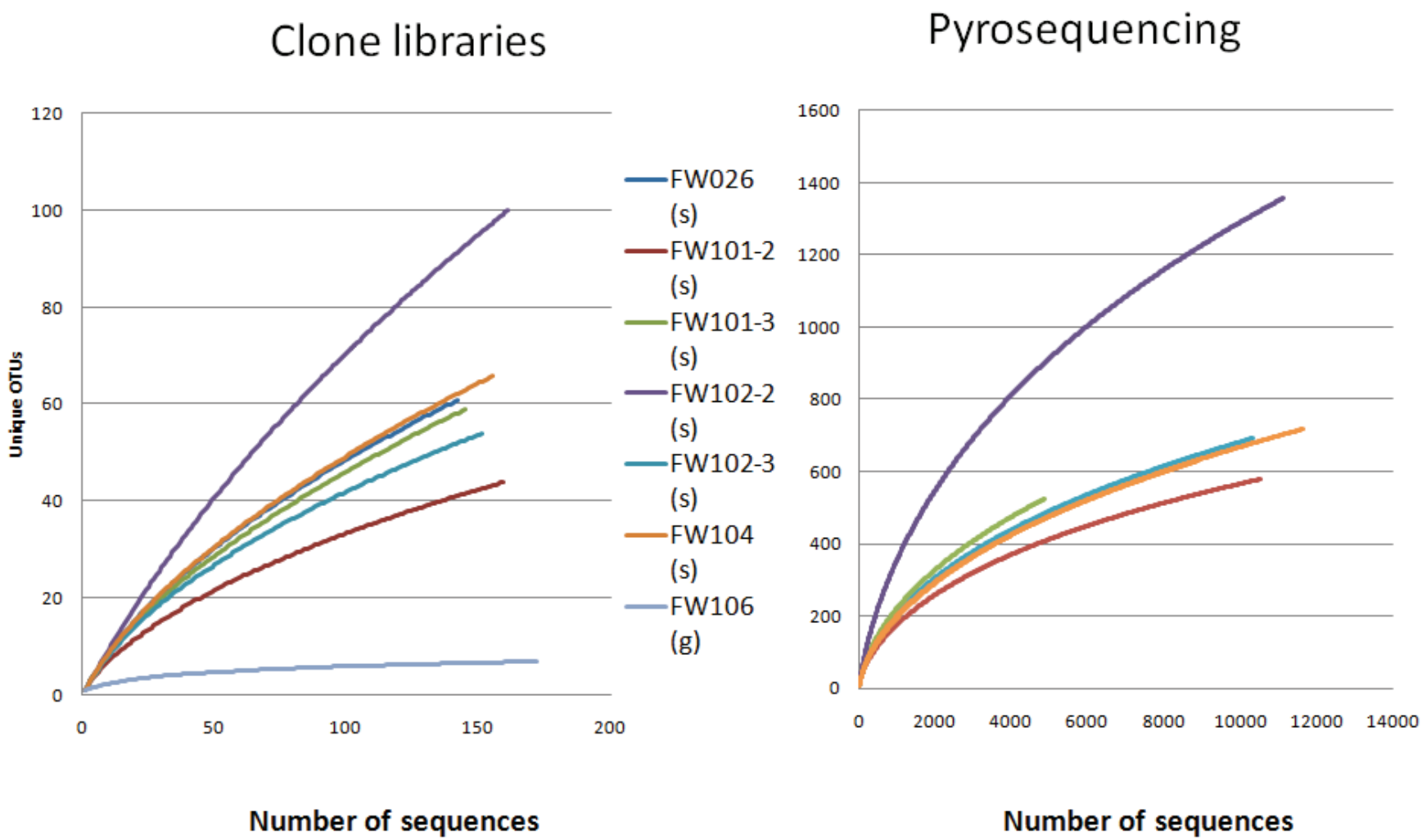
16S rRNA gene Pyrosequencing

Well	Sequences	Phyla	Classes	Orders	Families	Genera
FW026	10328	15	71	34	63	100
FW101-2	10523	15	68	36	56	95
FW101-3	4857	15	70	33	61	70
FW102-2	11098	18	87	36	79	133
FW102-3	8905	14	73	31	70	88
FW104	11649	15	69	33	62	86
Total	57359	19	20	42	95	197

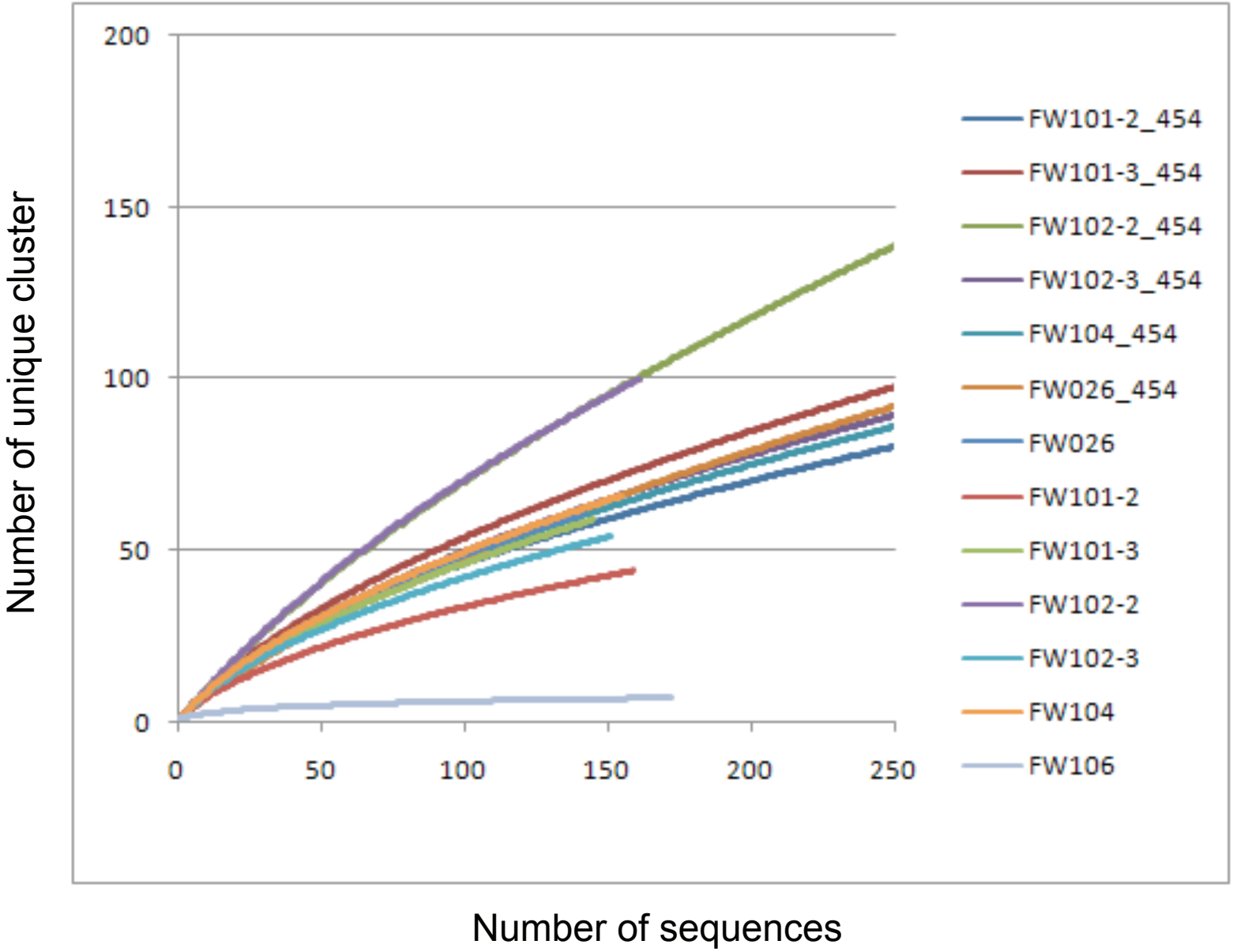
As classified by RDP Classifier using a 80% confidence threshold

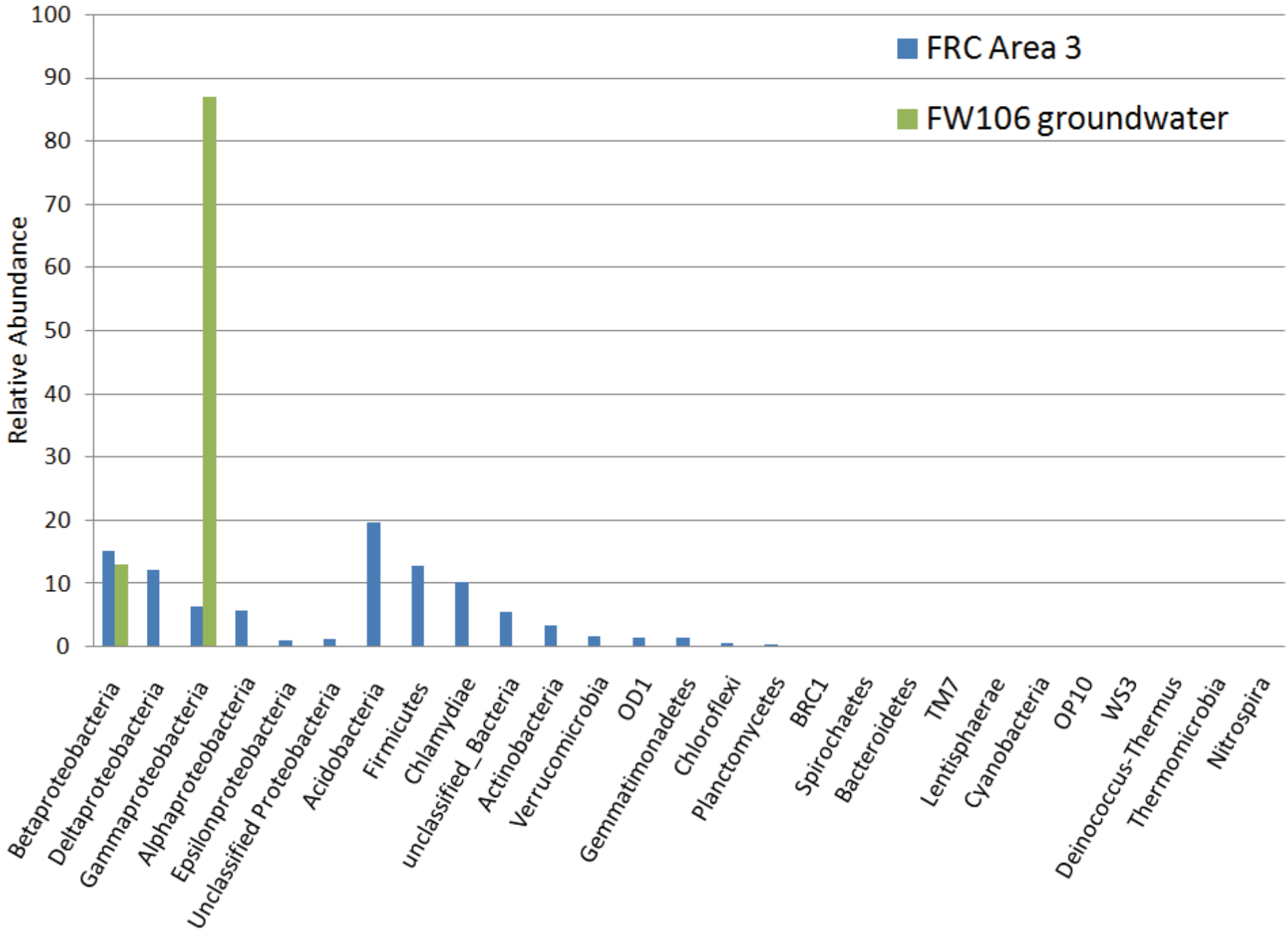
	Sequencing	
	Traditional Sanger	454 Pyrosequencing
Av. length	600-800	200(400)
Chimeras	Yes	Unlikely
Cloning bias	Yes	No
PCR bias	Yes	Yes
Intensive sampling	No	Yes

Comparison of rarefaction curves



Rarefaction curves for Pyrosequencing and Clone Libraries
Clusters are defined at 3%




FRC Area 3
FW106 groundwater
Relative Abundance
100
90
80
70
60
50
40
30
20
10
0
Betaproteobacteria
Deltaproteobacteria
Gammaproteobacteria
Alphaproteobacteria
Epsilonproteobacteria
Unclassified Proteobacteria
Acidobacteria
Firmicutes
Chlamydiae
unclassified_Bacteria
Actinobacteria
Verrucomicrobia
OD1
Gemmatimonadetes
Chloroflexi
Planctomycetes
BRC1
Spirochaetes
Bacteroidetes
TM7
Lentisphaerae
Cyanobacteria
OP10
WS3
Deinococcus-Thermus
Thermomicrobia
Nitrospira

Top 10 clusters (5%) in sediments from the FRC Area3

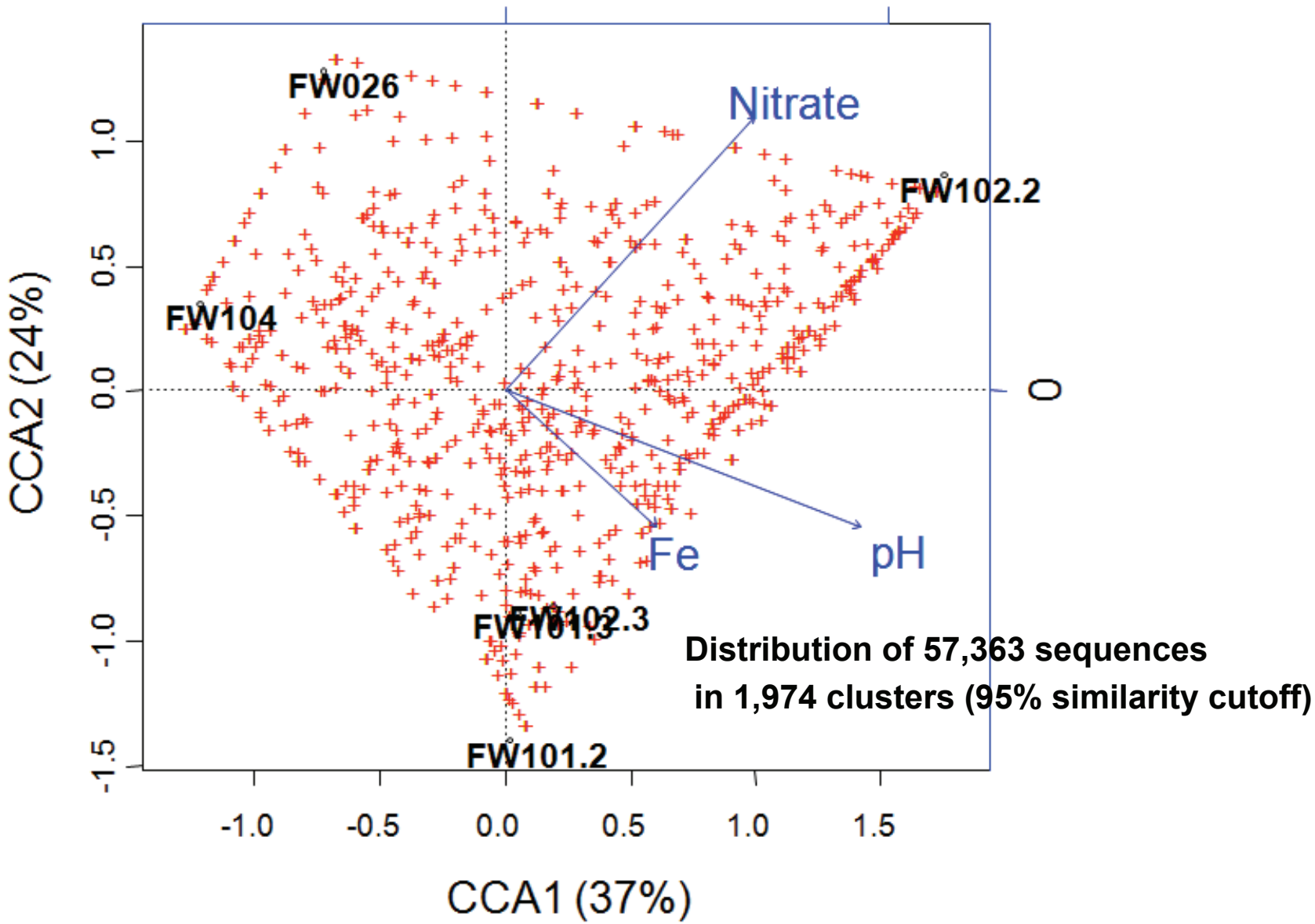
Rank	Occurrence	Frequency in sample (%)	Classification *		Sab score	Closest Match **	Final electron acceptor used by relatives	
			Phylum	Genus			U(VI)	Fe(II)
1	8103	14.1	Acidobacteria	Gp8 (Geothrix)	0.942	Uncultured *Holophaga* sp.; GuBH2-AD-9; AJ519665		Yes
2	2603	4.54	Firmicutes	Desulfosporosins	0.915	*Desulfosporosinus* sp. 5apy; AF159120	Yes	Yes
3	2298	4.01	Proteobacteria-δ	Desulfovibrio	0.946	*Desulfovibrio putealis* (T); DSM 16056; AY574979	Yes	Yes
4	2179	3.80	Proteobacteria-γ	Unclassified Xhantomonadace	0.959	Uncultured bacterium; AKAU3534; DQ125555		
5	1976	3.44	Chlamydiae	Parachlamydia	0.862	Undosymbiont of *Acanthamoeba* sp. UWE25; AF083615		
6	1686	2.94	Chlamydiae	Rhabdochlamydia	0.706	Uncultured bacterium; LT110PlE2; DQ444442		
7	1686	2.94	Proteobacteria-β	Ferribacterium	0.947	*Ferribacterium limneticum* (T); cda-1; Y17060		Yes
8	1125	1.96	Acidobacteria	Gp8 (Geothrix)	0.946	Uncultured *Geothrix* sp.; S15B-MN21; AJ583203		Yes
9	831	1.45	Proteobacteria-δ	Desulfovibrio	0.921	Uncultured bacterium; LS4-150; AB234259	Yes	Yes
10	811	1.41	Firmicutes	Pelotomaculum	0.949	Uncultured bacterium; AKAU4159; DQ125898		Yes

*Using RDP Classifier at 80% confidence treshold
** Using RDP SeqMatch against a Good quality,near full length(>1300bp) dataset with isolates and clones

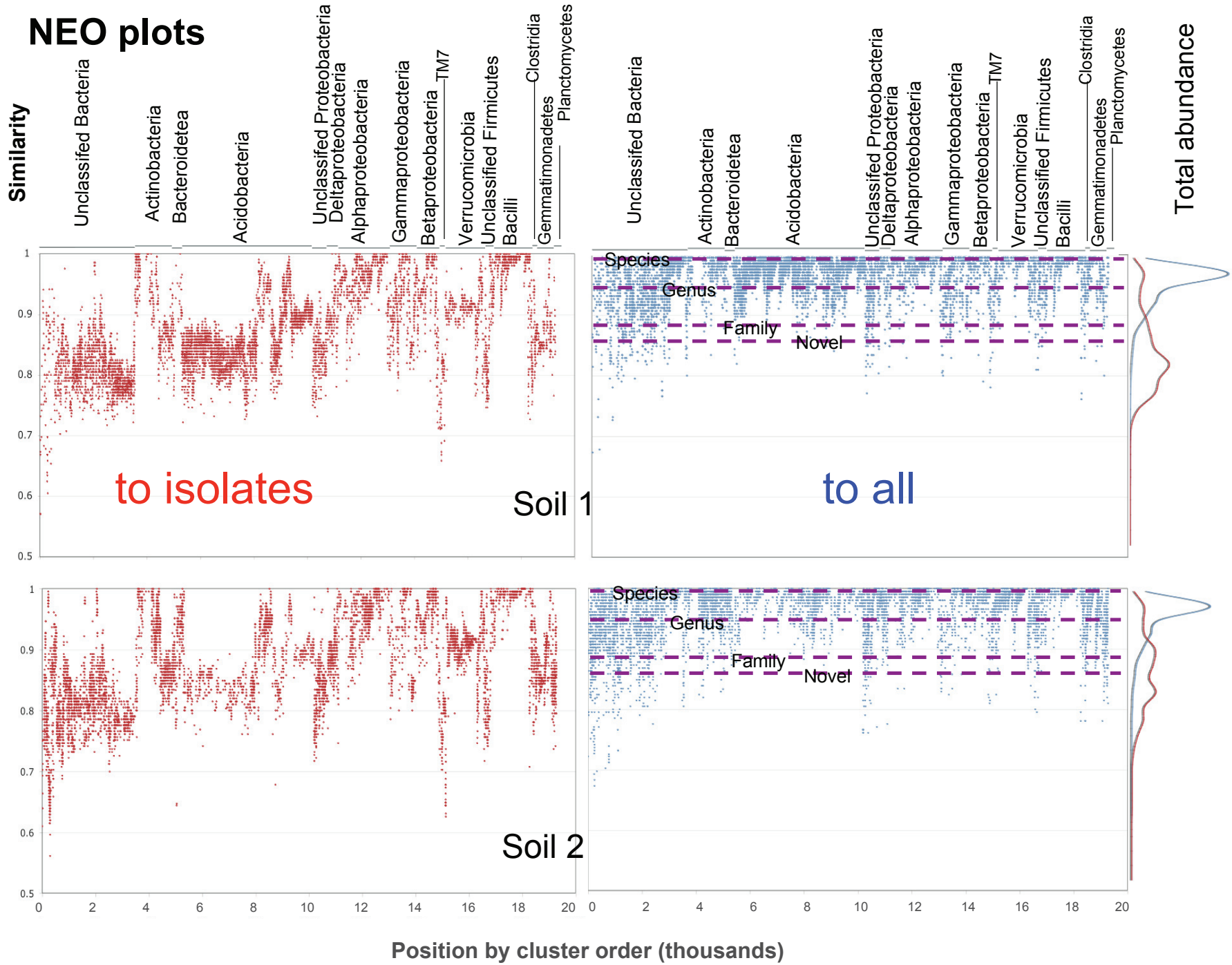
Relative abundance of known U(VI) reducers in libraries

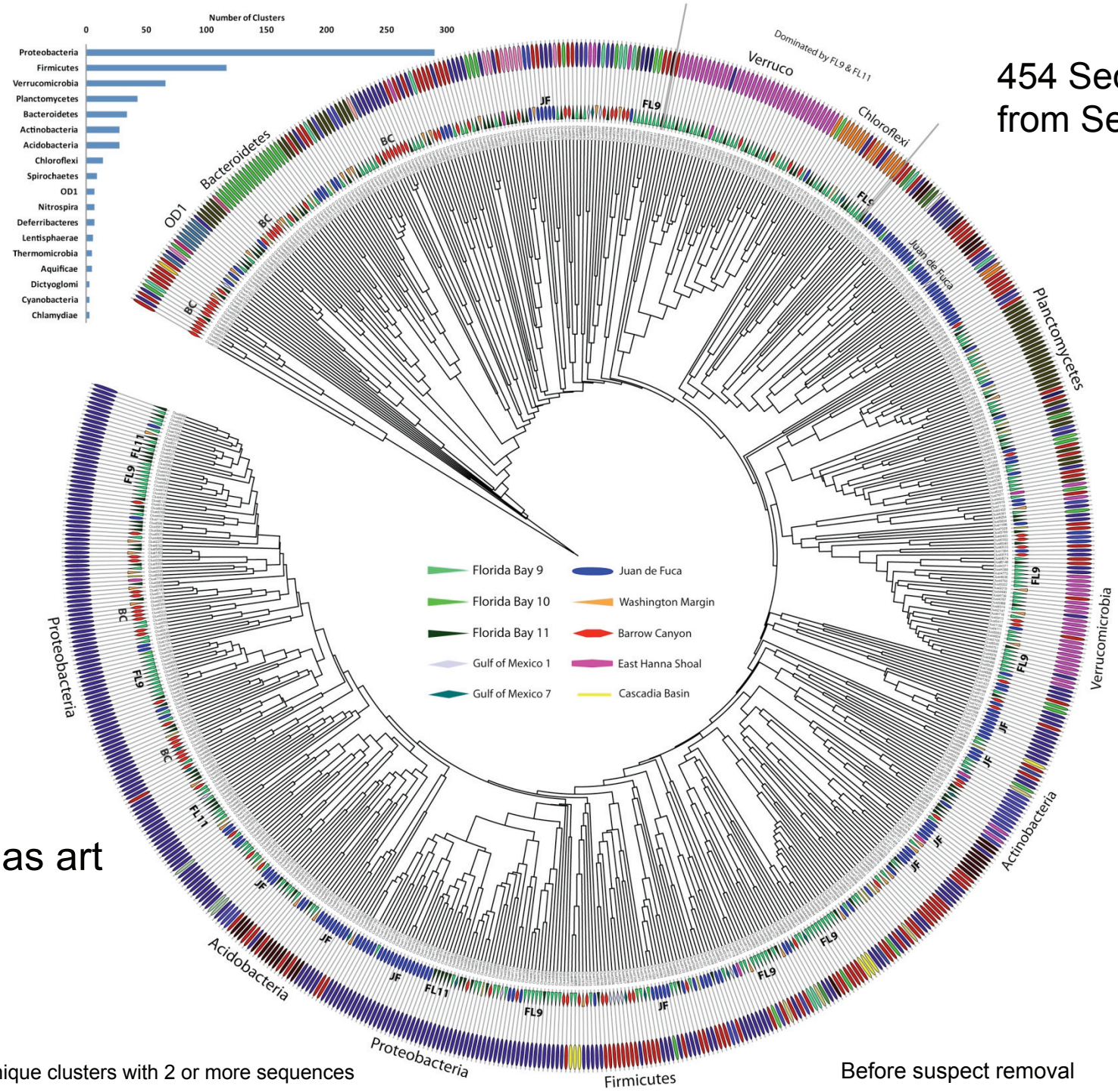
U(VI) Reducer	454 Libraries*	Clone libraries**
Desulfovibrio	8.35	7.08
Desulfosporosinus	6.09	1.10
Geobacter	1.47	2.58
Anaeromyxobacter	1.02	1.84
Acidovorax	0.02	0.37
Desulfitobacterium	0.01	0.000
Deinococcus	0.005	0.000
Shewanella	0.003	0.000

*RDP Classifier at 80% confidence (4/3/08)
**RDP Classifier at 80% confidence (4/6/08)



The first CCA is strongly influences by pH, while the second axis is strongly influenced by nitrate. Anova test supports the significance (p=0.01, when alpha=0.05, beta=0.01, step=100, Permutations .max=10000)


NEO plots
Similarity
Unclassifed Bacteria
Actinobacteria
Bacteroidetea
Acidobacteria
Unclassifed Proteobacteria
Deltaproteobacteria
Alphaproteobacteria
Gammaproteobacteria
Betaproteobacteria
TM7
Verrucomicrobia
Unclassified Firmicutes
Bacilli
Clostridia
Gemmatimonadetes
Planctomycetes
Total abundance
Species
Genus
Family
Novel
to isolates
to all
Soil 1
Soil 2
Position by cluster order (thousands)

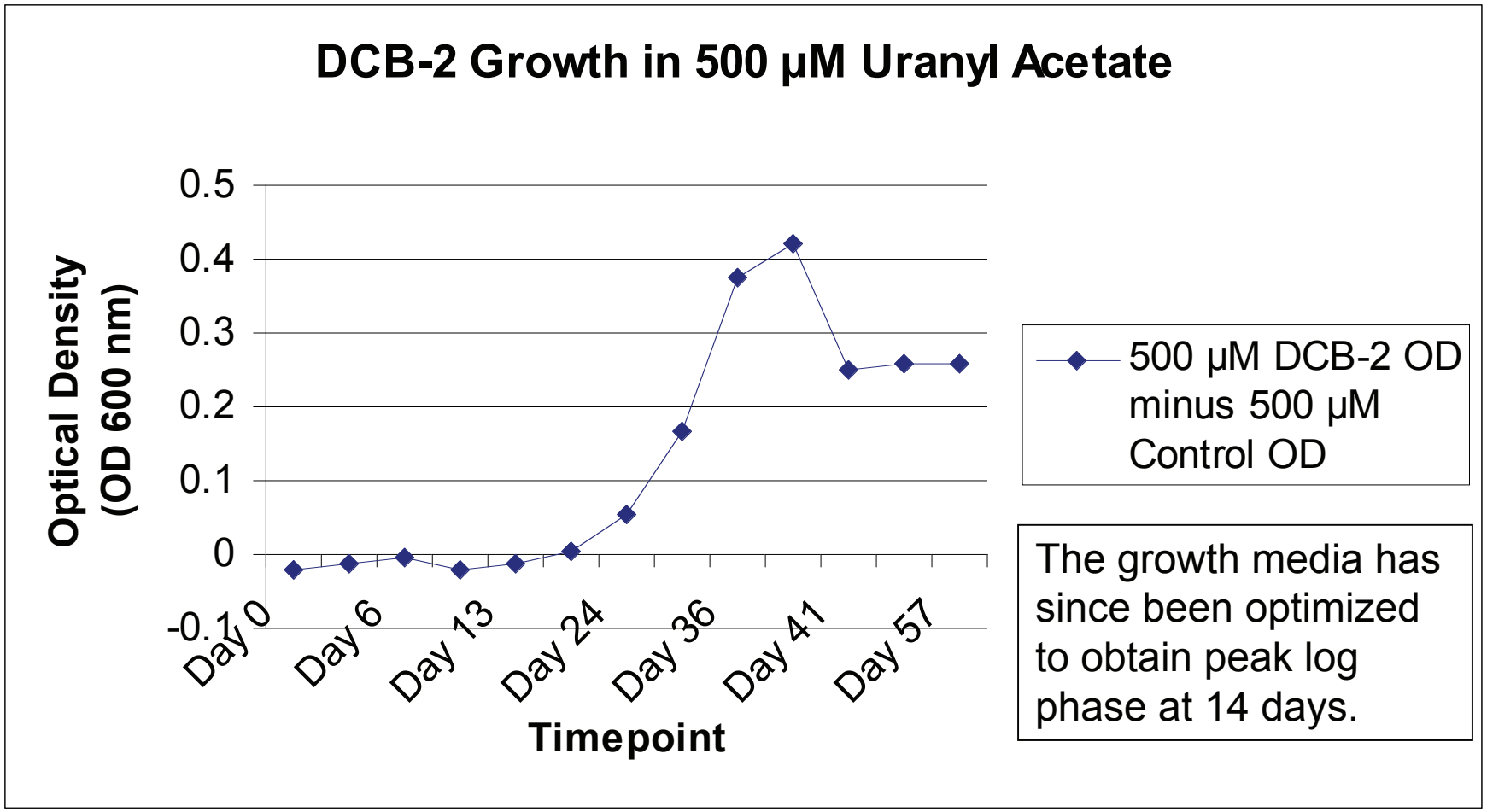


454 Sequences from Sediments

Trees as art forms!

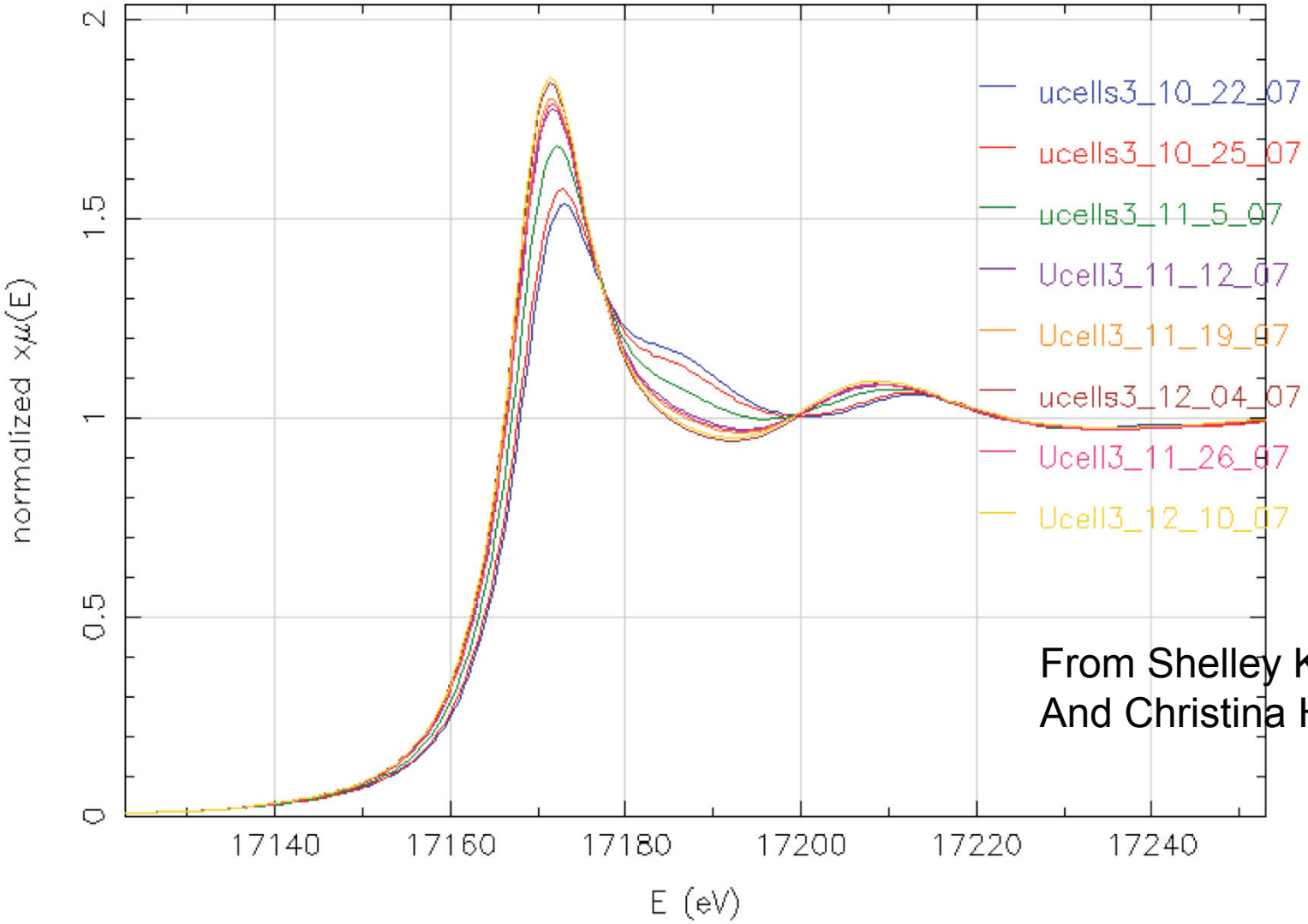
696 Unique clusters with 2 or more sequences

Before suspect removal


DCB-2 Growth in 500 µM Uranyl Acetate
Optical Density (OD 600 nm)
0.5
0.4
0.3
0.2
0.1
0
-0.1
Day 0
Day 6
Day 13
Day 24
Day 36
Day 41
Day 57
Timepoint
500 µM DCB-2 OD minus 500 µM Control OD
The growth media has since been optimized to obtain peak log phase at 14 days.

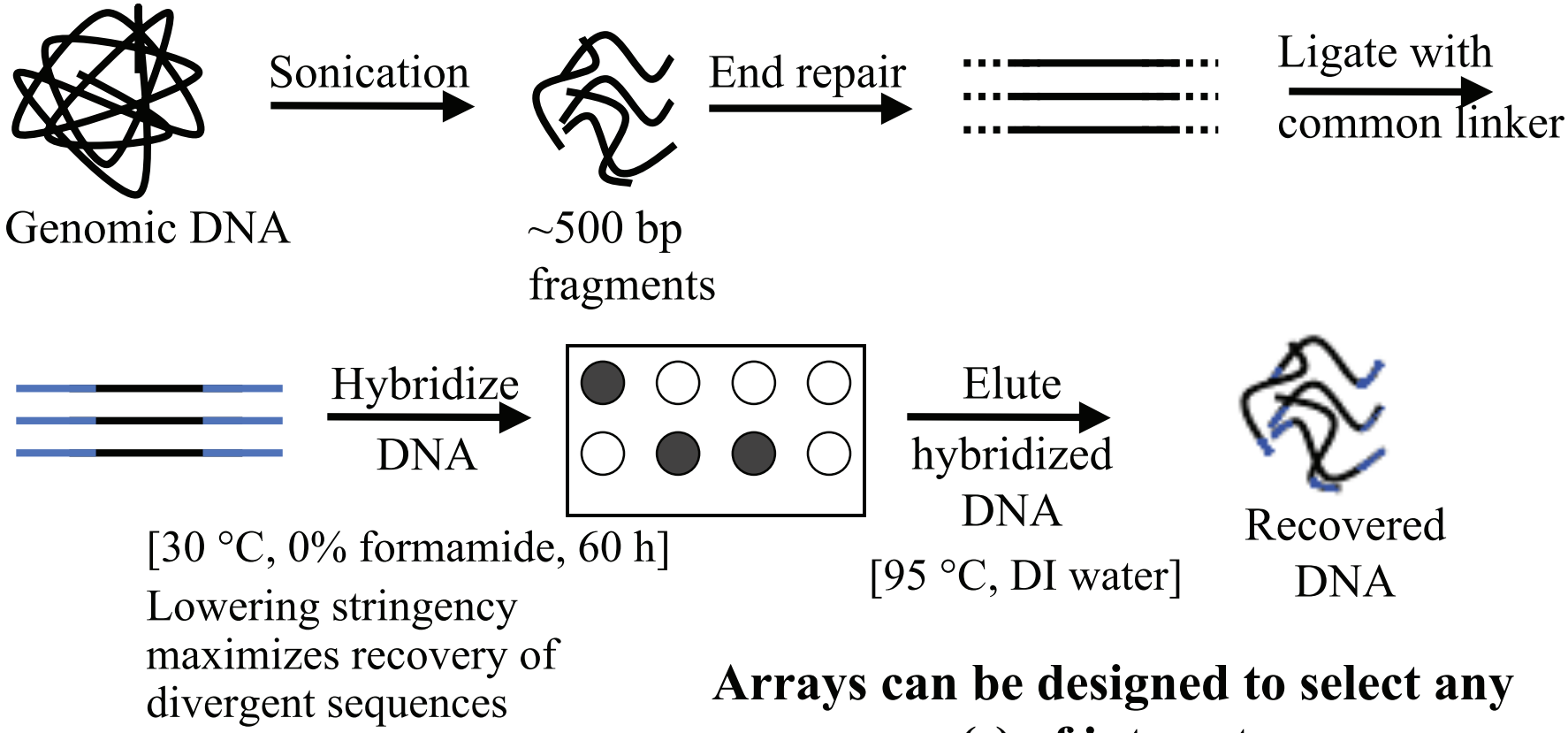
But have observed growth by DCB-2 with Uranium (VI) reduction at 0.91 mM Uranyl acetate

Time course of XANES illustrating reduction of 500 µM Uranium over time by *Desulfitobacterium hafniense* DCB-2.



From Shelley Kelly
And Christina Harzman

Sequence selection strategy



Arrays can be designed to select any gene group(s) of interest

Alberts et al., 2007; Okou et al., 2007, Nature Methods

Microarray Approach for selecting desired BAC/fosmid Clones for Sequencing

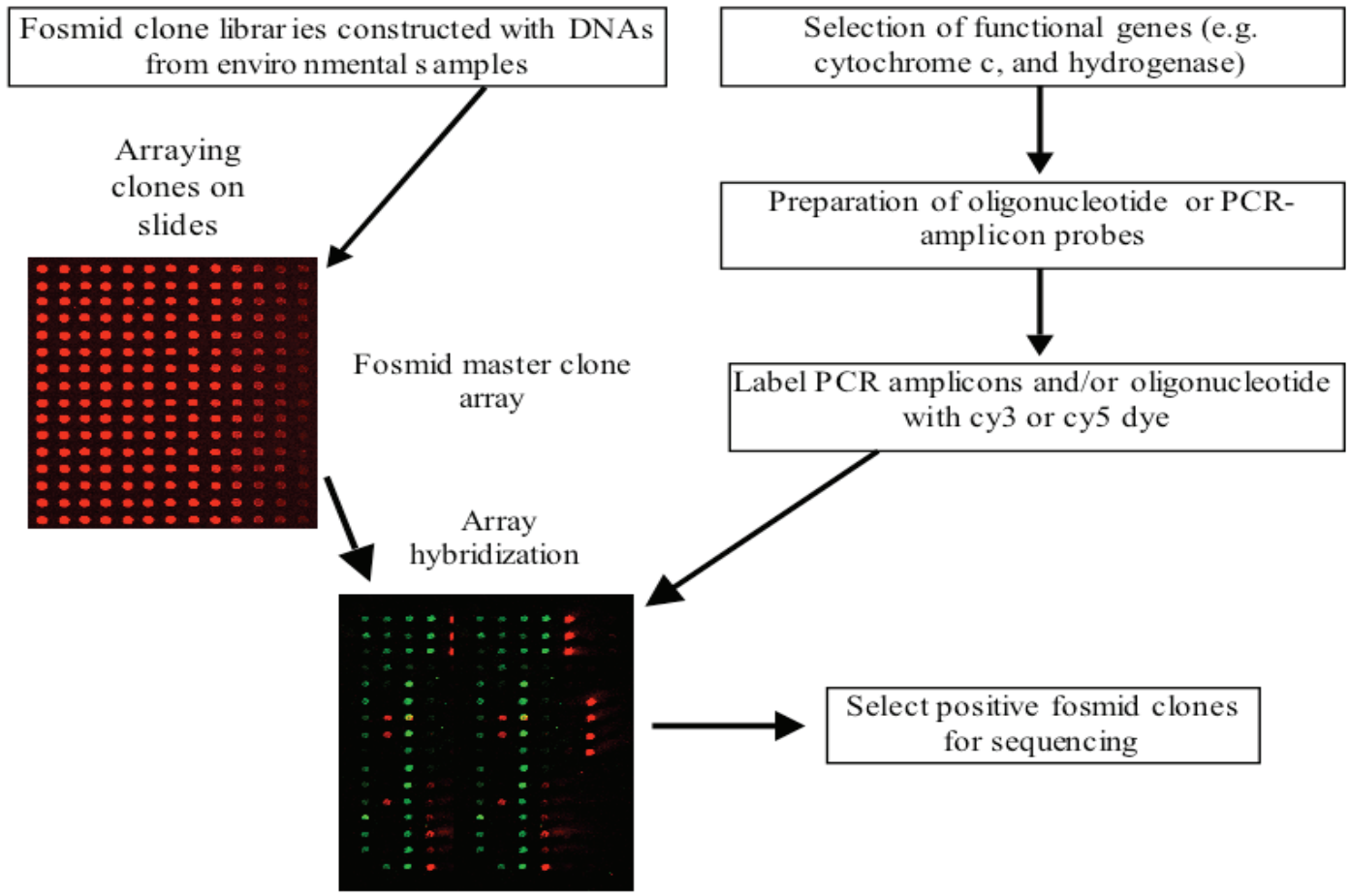
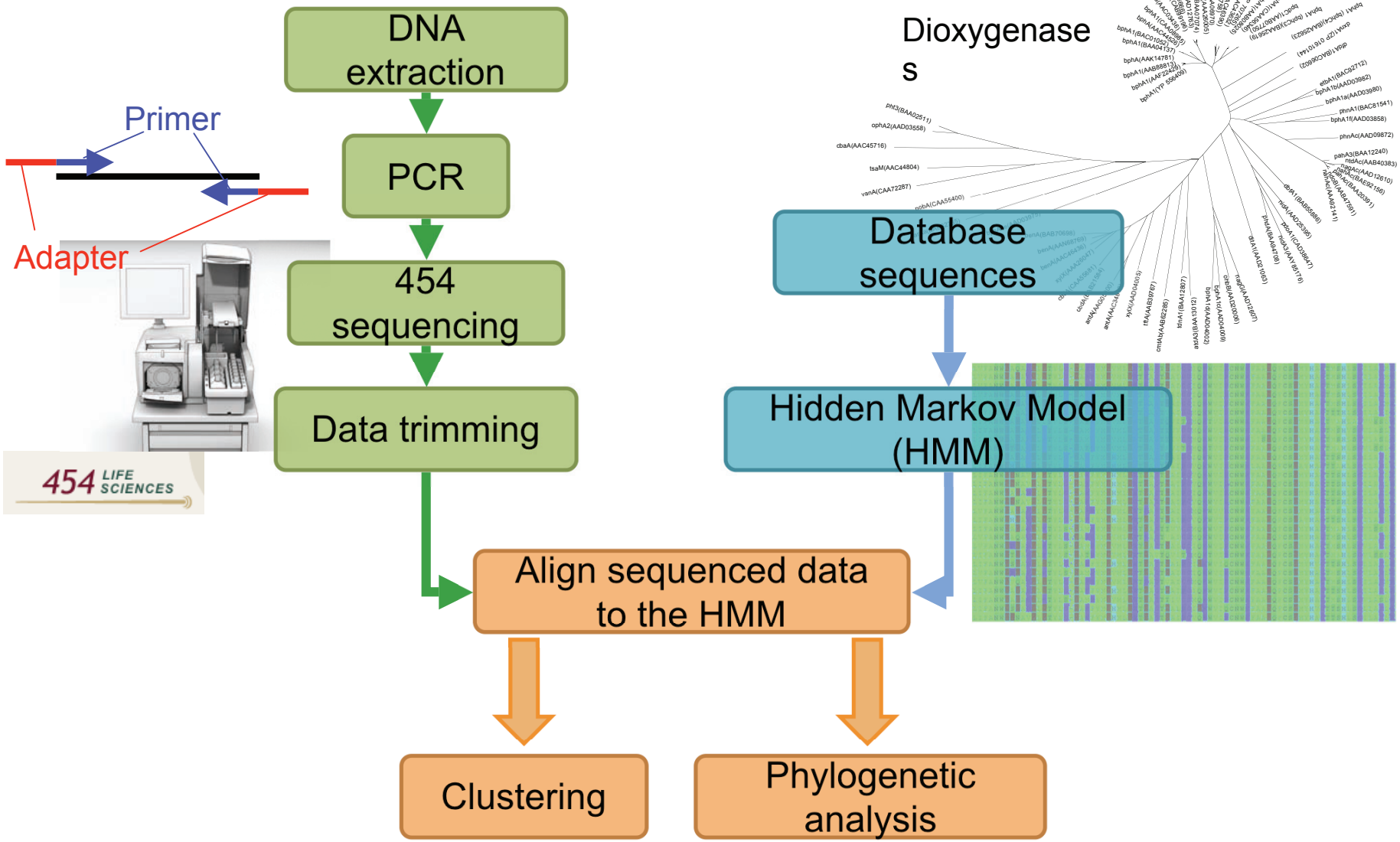


Fig. 1 Construction of fosmid libraries and selection of fosmid clones for sequencing using microarray-based meta-genomics approaches.

Scheme of 454 sequencing for biphenyl dioxygenase



Functional Gene Pipeline/Repository

[help | FGPR home | RDP home]

:: Results

Your Search Topic: narG_v2.1 :: Showing 1074 of 2240 Total Results

Analyze Selection | Jump to Display/Filtering Options | Select All

seed sequences

Motif1: Molydop_binding
Motif2: Molybdopterin

organism name

= HMM training sequence
= New to this release

Select	Score	New Hit	PID	NID	Definition	Organism	Occ.	% of HMM Match	Size (aa)	Reference	Motif1	Motif2	Notes	View Edi
☐	3580.0		AAC74308	U00096	nitrate reductase 1, alpha subunit [Escherichia coli K12]	Escherichia coli K12	1	100	1247	Blattner,F.R.	151.9	631.5		...
☐	3575.3		NP_287472	NC_002655	nitrate reductase 1, alpha subunit [Escherichia coli O157:H7 EDL933]	Escherichia coli O157:H7 EDL933	1	100	1247	Perna,N.T.	151.9	631.5		
					A, The Crystal ure Of The Narghi t Nari-K86a In ex With chlorophenol	Escherichia coli	1	100	1246	Ber				...
					tory nitrate reductase a chain [Escherichia 6]	Escherichia coli 536	1	100	1247	Hochhut,B.	153.5	631.2		...
					tase hia	Escherichia coli B7A	1	100	1247	NCBI Microbial	151.9	631.5		

active links to GenBank records

:: Analysis of 7 Selections - narG_v2.1

Download Protein Alignment | Generate Protein Tree with Selection
Download Nucleotide Alignment | Generate Nucleotide Tree with Selection
Download Phylip (Protein)
Download Phylip (Nucleotide)

Note: Inclusion of poorly alignable sequences may produce an aberrant tree.

custom analysis

Probe Match Search

Please enter your probe:

5' _____ 3'

Probe targets + strand. Search Nucleotide

Request Custom Analysis

Give brief instructions of the type of analysis you wish to be performed on your selected sequences.

Your Name:

display/filter options

Order search results by: score **Make any other changes below before pressing** Submit

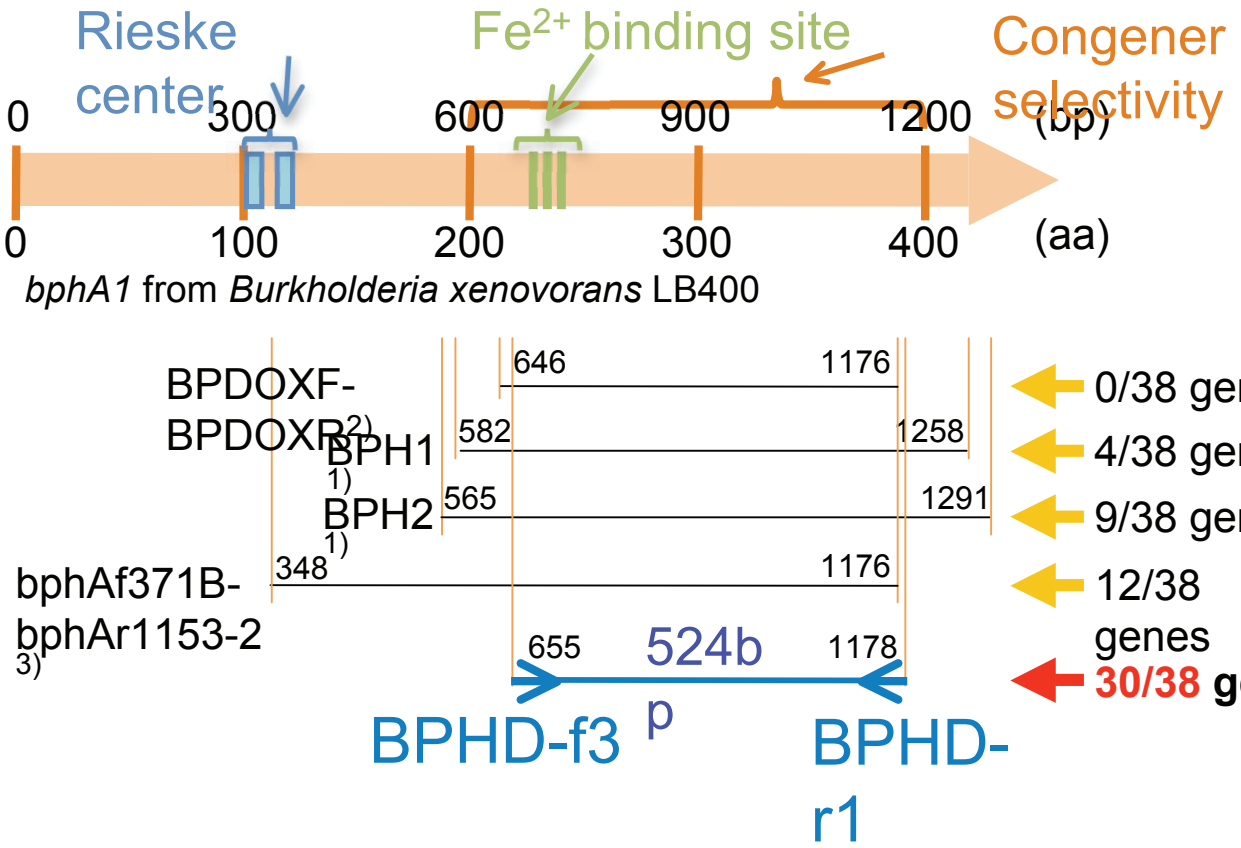
Display the following columns:

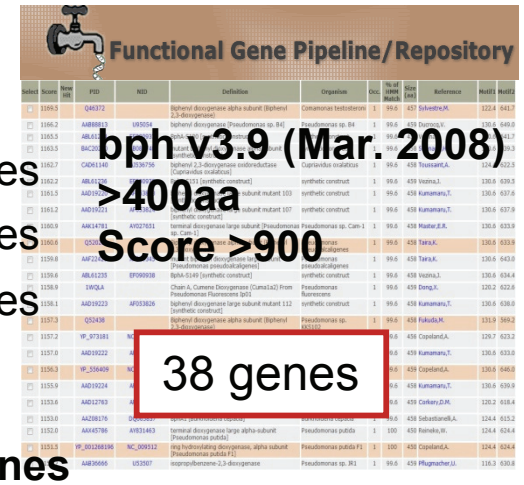
☑ Score ☑ New Hit ☑ Definition ☑ Size (aa)
☑ PID ☑ Organism ☑ Reference ☑ HMM Match
☑ NID ☑ Occurrence ☑ Motifs ☑ Notes

Filter by:

Size > 0 Size ≤ 10000000
HMMmatch% ≥ 0 Score ≥ 151.3
○ Clones ○ Isolates ◉ Both
☐ Redundant entries ☐ New Hits

Dioxygenase primer sets

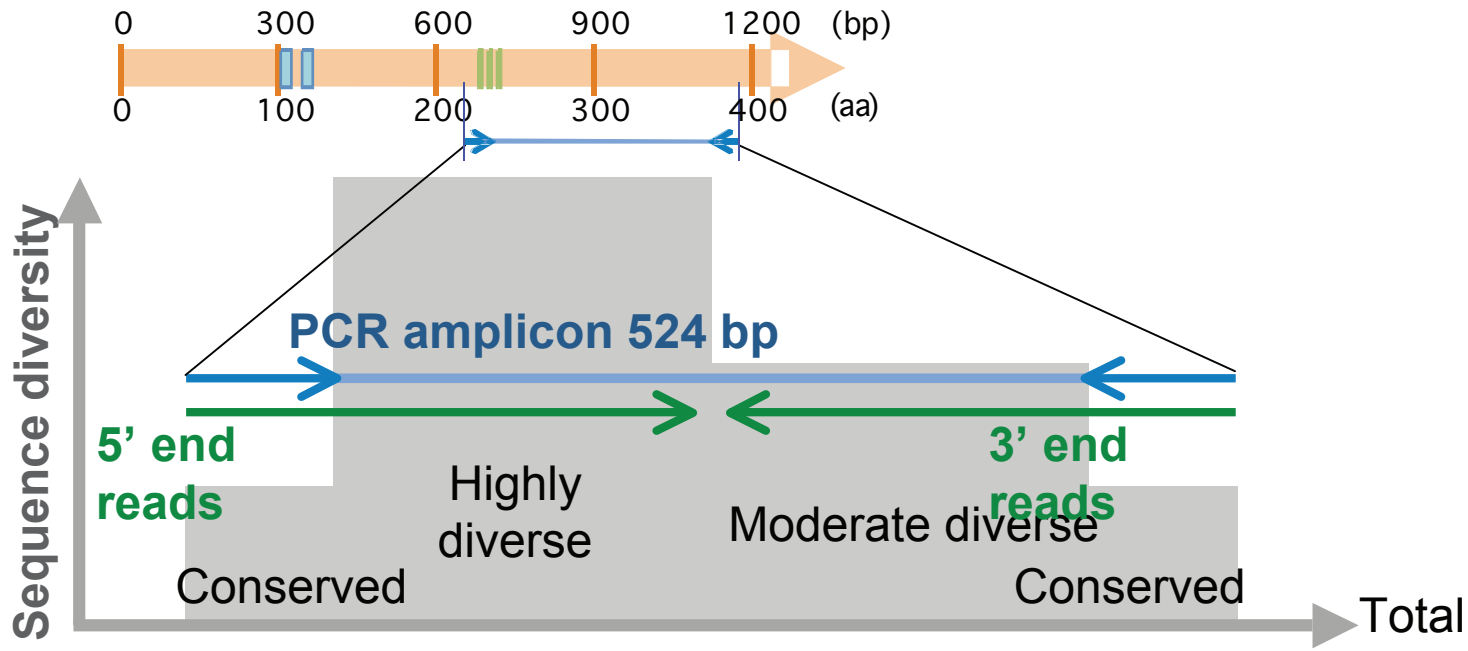




- ✓ Specific to toluene/biphenyl dioxygenase
- ✓ Amplify 524 bp of congener selectivity site
- ✓ Cover 30/38 sequences in Fungene database bph v3.9 (Mar. 2008)

1) Baldwin *et al*. (2003), 2) Yeates *et al.* (2000), 3) Witzig *et al.*, (2006)

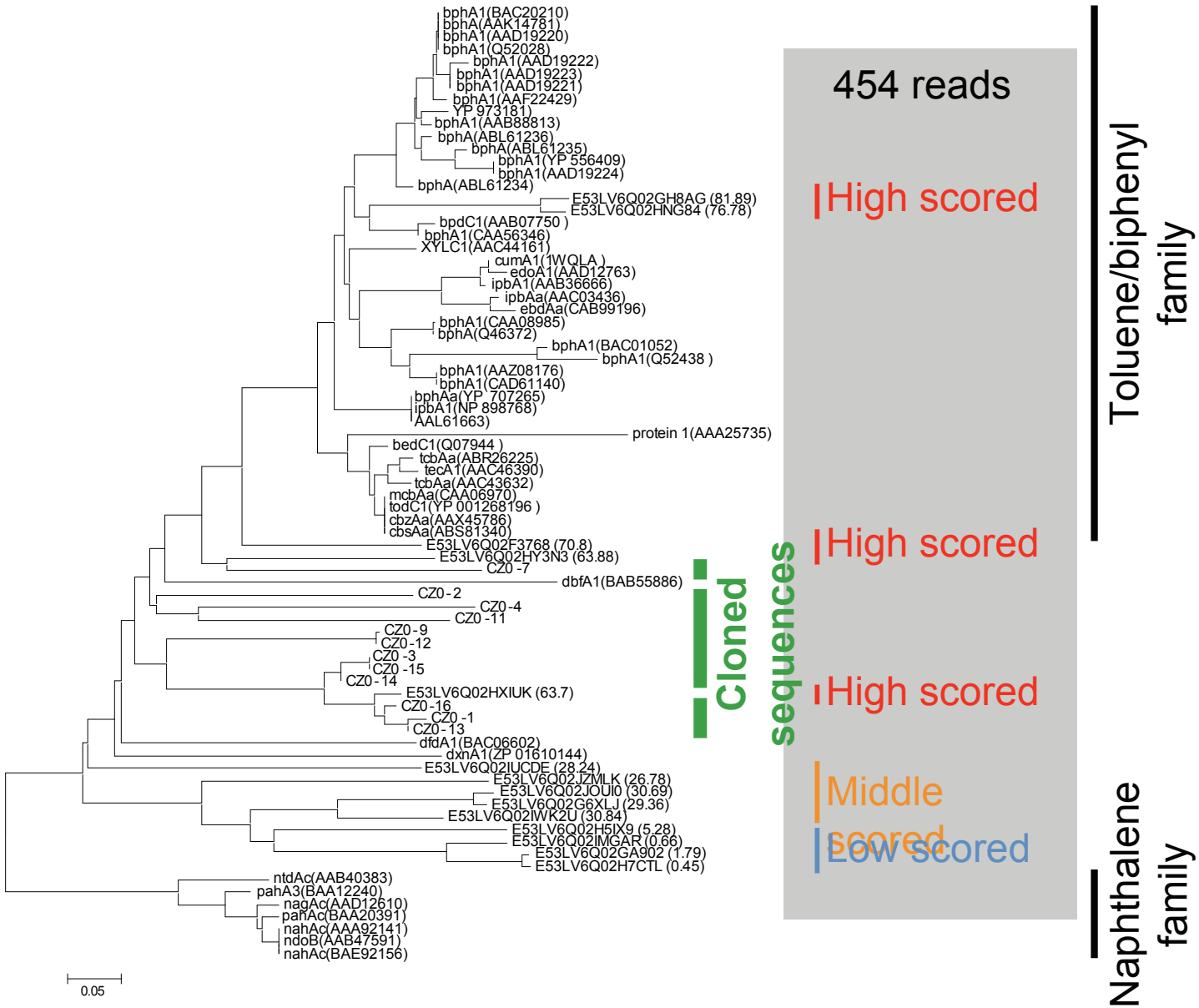
454 sequencing reads



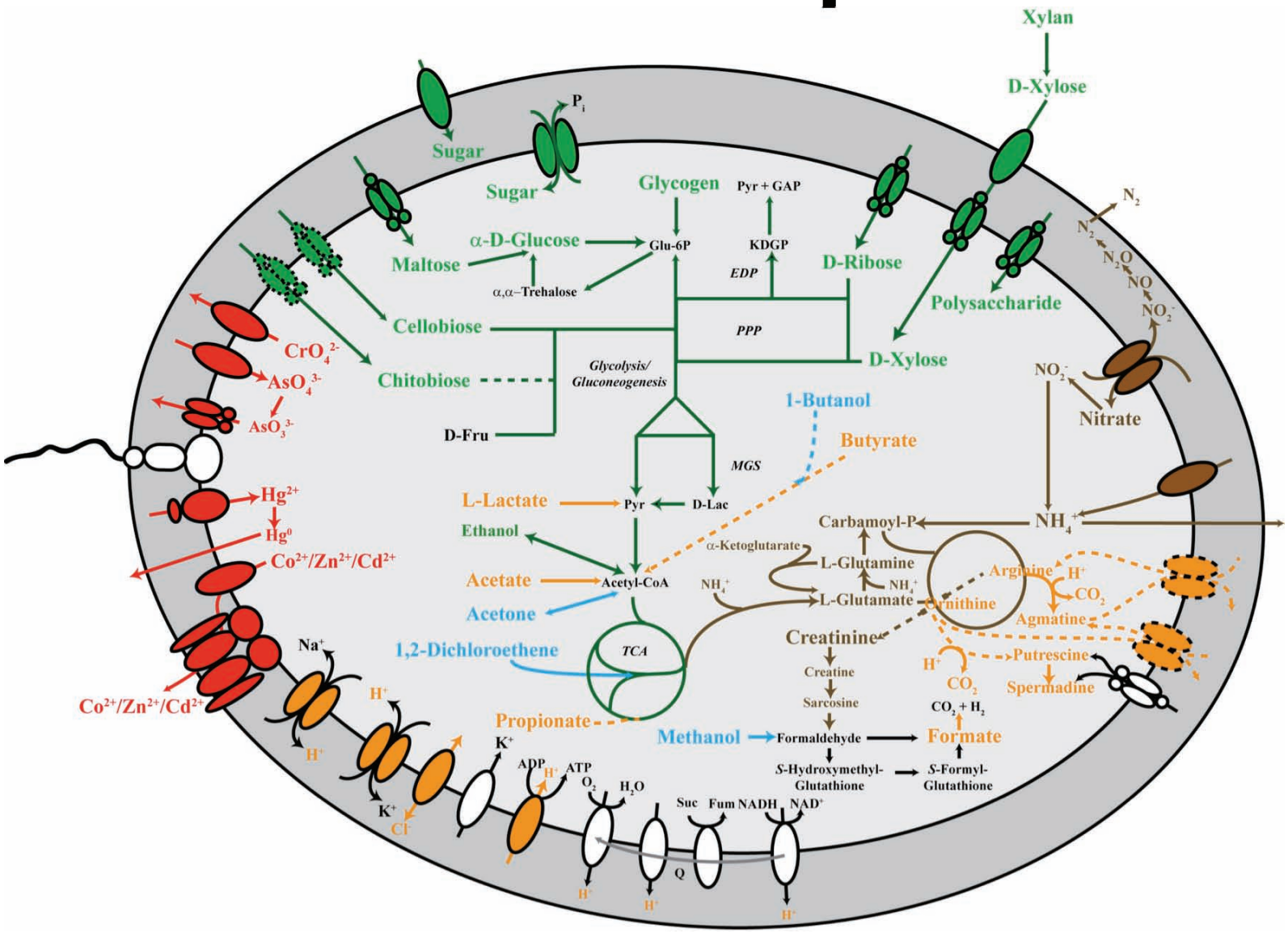
Reads	2,486	835	3,322
Aligned without primer sequences (Trimmed)	80	787	867
Aligned with primer sequences	2,302	830	3,132 (94%)

- At least 94 % of the reads are dioxygenases

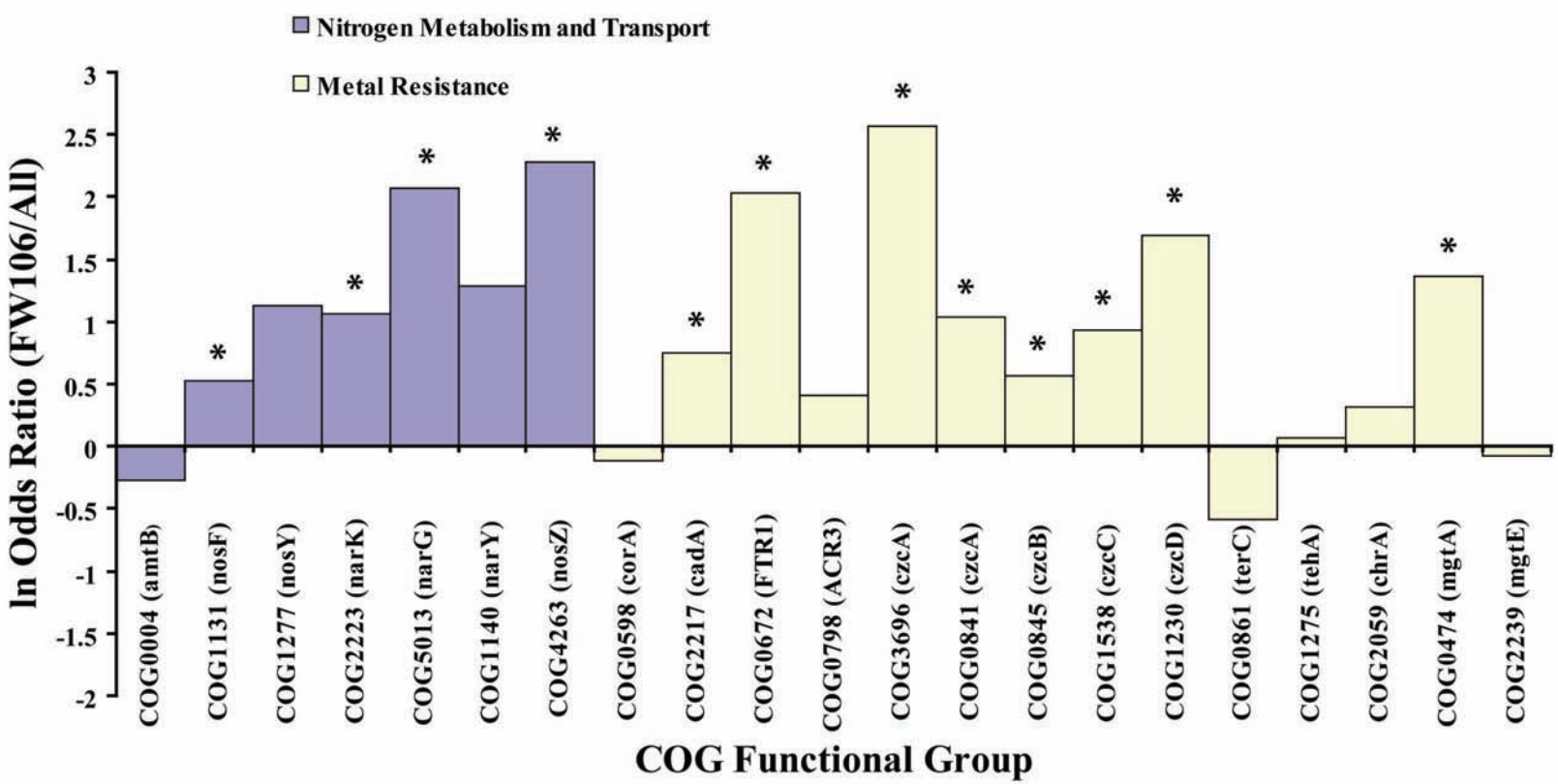
Comparison to the clone library



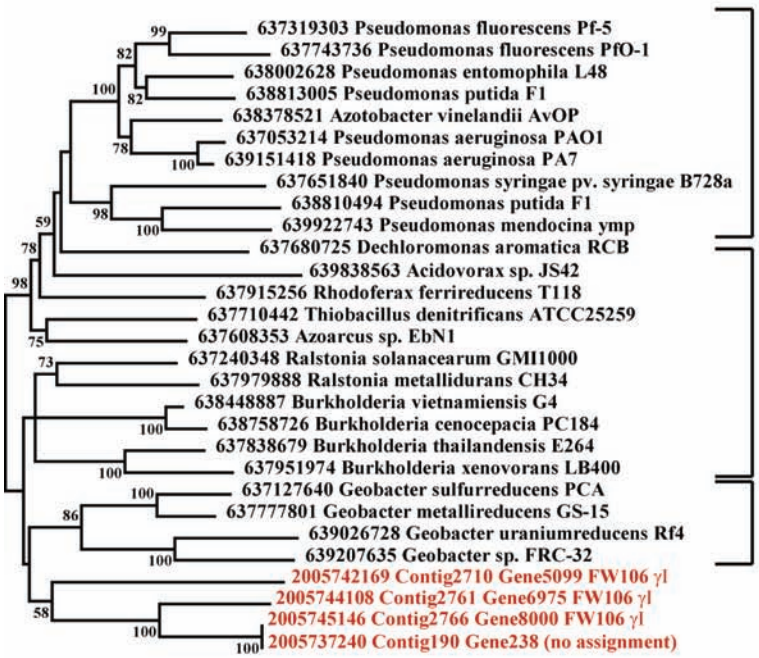
FW106 γI


Xylan
D-Xylose
Sugar
Sugar
Pi
Glycogen
Pyr + GAP
α-D-Glucose
Glu-6P
KDGP
EDP
D-Ribose
Maltose
α,α–Trehalose
Polysaccharide
Cellobiose
PPP
D-Xylose
Chitobiose
Glycolysis/
Gluconeogenesis
D-Fru
1-Butanol
Butyrate
MGS
L-Lactate
Pyr
D-Lac
Ethanol
Acetate
Acetyl-CoA
Acetone
1,2-Dichloroethene
TCA
Propionate
α-Ketoglutarate
L-Glutamine
L-Glutamate
NH4+
Carbamoyl-P
NH4+
Arginine
Ornithine
Agmatine
Creatinine
Creatine
Sarcosine
Putrescine
Spermidine
Methanol
Formaldehyde
Formate
S-Hydroxymethyl-
Glutathione
S-Formyl-
Glutathione
CO2 + H2
N2
N2O
NO
NO2-
Nitrate
CrO4 2-
AsO4 3-
AsO3 3-
Hg2+
Hg0
Co2+/Zn2+/Cd2+
Co2+/Zn2+/Cd2+
Na+
H+
K+
Cl-
ADP
ATP
O2
H2O
Suc
Fum
NADH
NAD+
Q

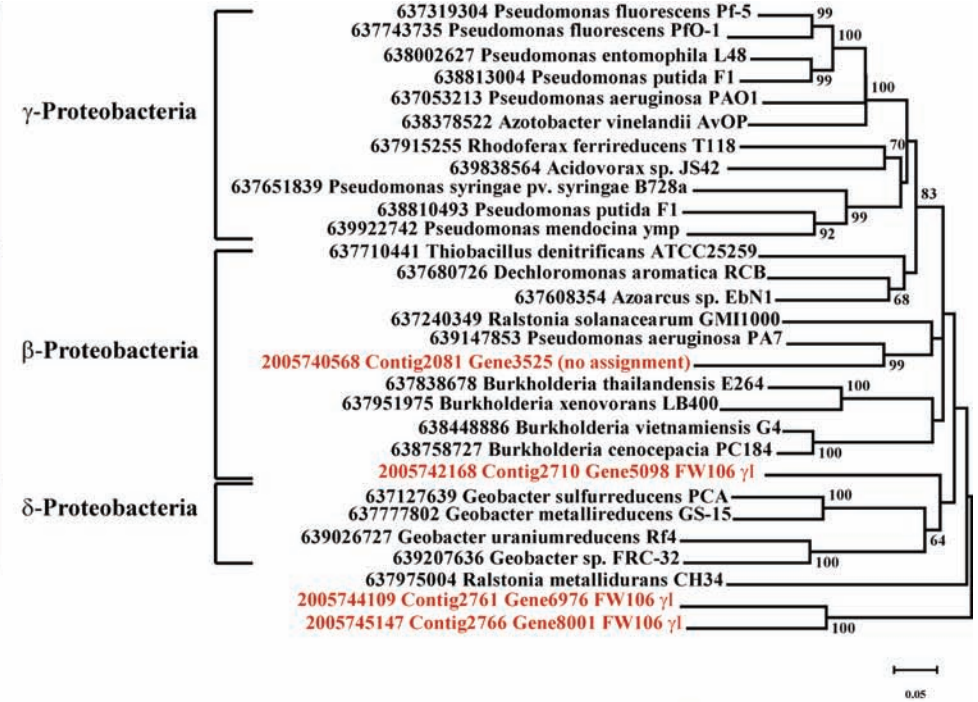
Abundance of COGs in FW106



LGT in FW106 Community

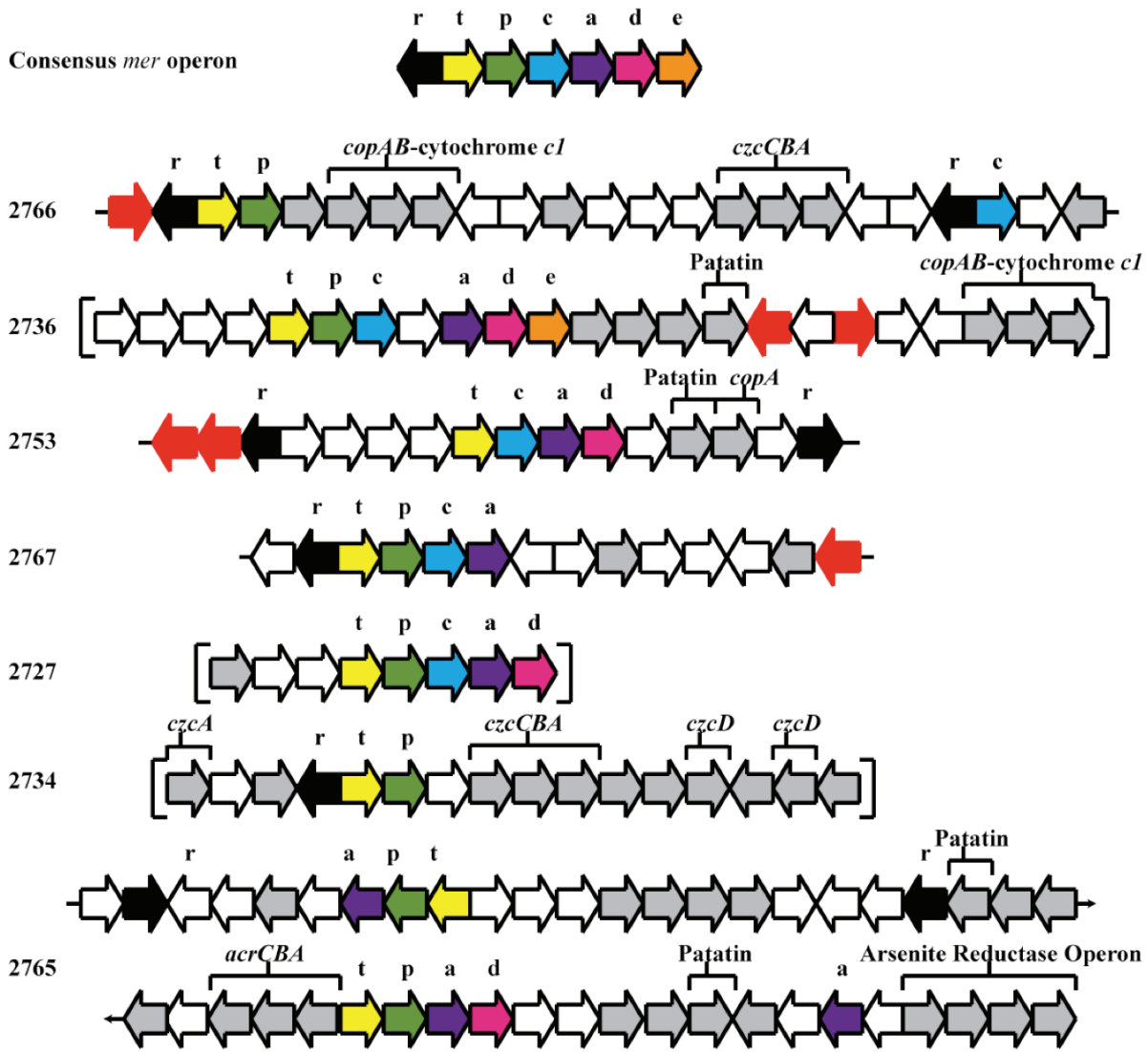
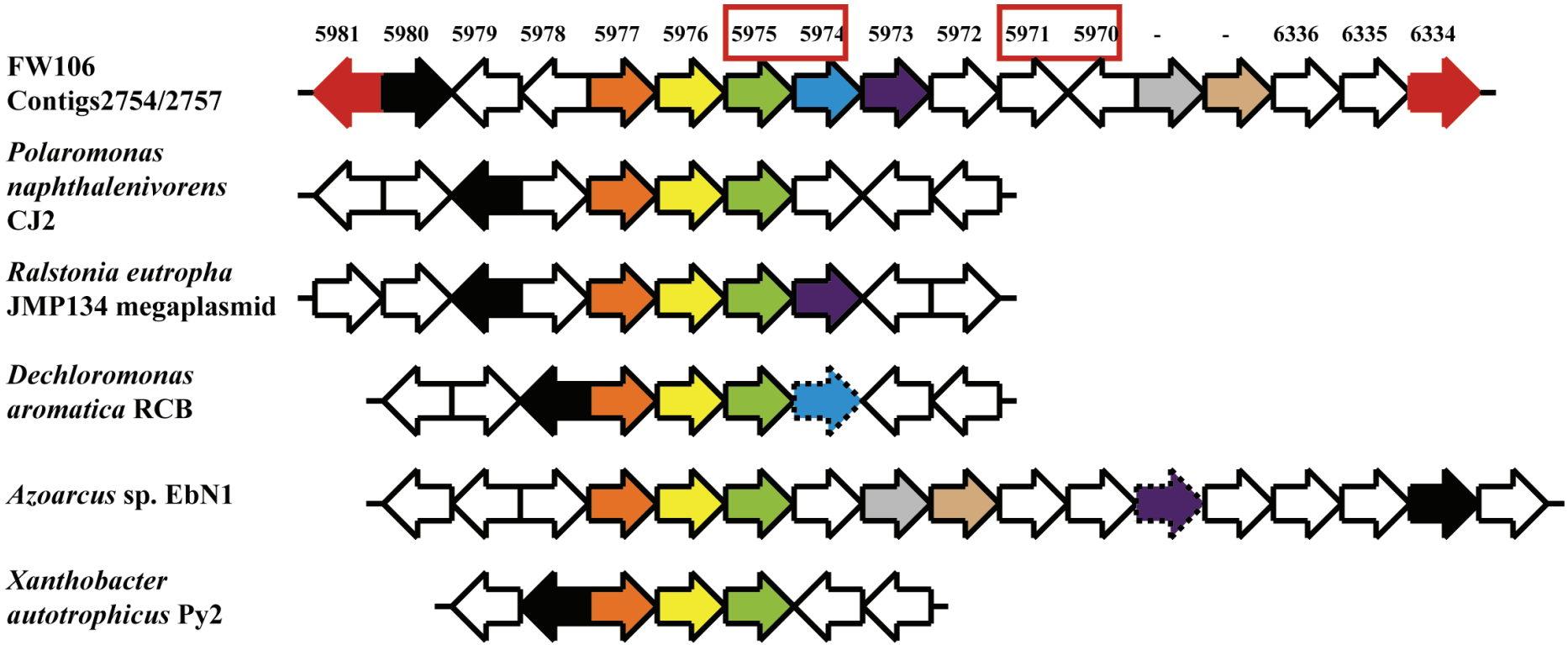


Sensor Histidine Kinase
(IrlS/CopS/CzcS)



Response Regulator
(IrlR/CopR/CzcR)

LGT in FW106 Community



LGT in FW106 Community