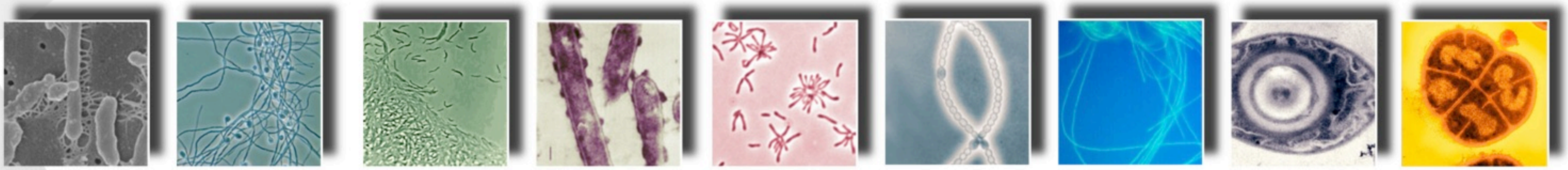
Single-cell genomics

2015 Workshop on Genomics, Cesky Krumlov
January 22, 2015

Tanja Woyke, DOE-JGI



JGI JOINT GENOME INSTITUTE
UNITED STATES DEPARTMENT OF ENERGY

- **A few words about the JGI**
  - who we are & what we do
- **Single-cell genomics**
  - the why, the how & what to expect from it

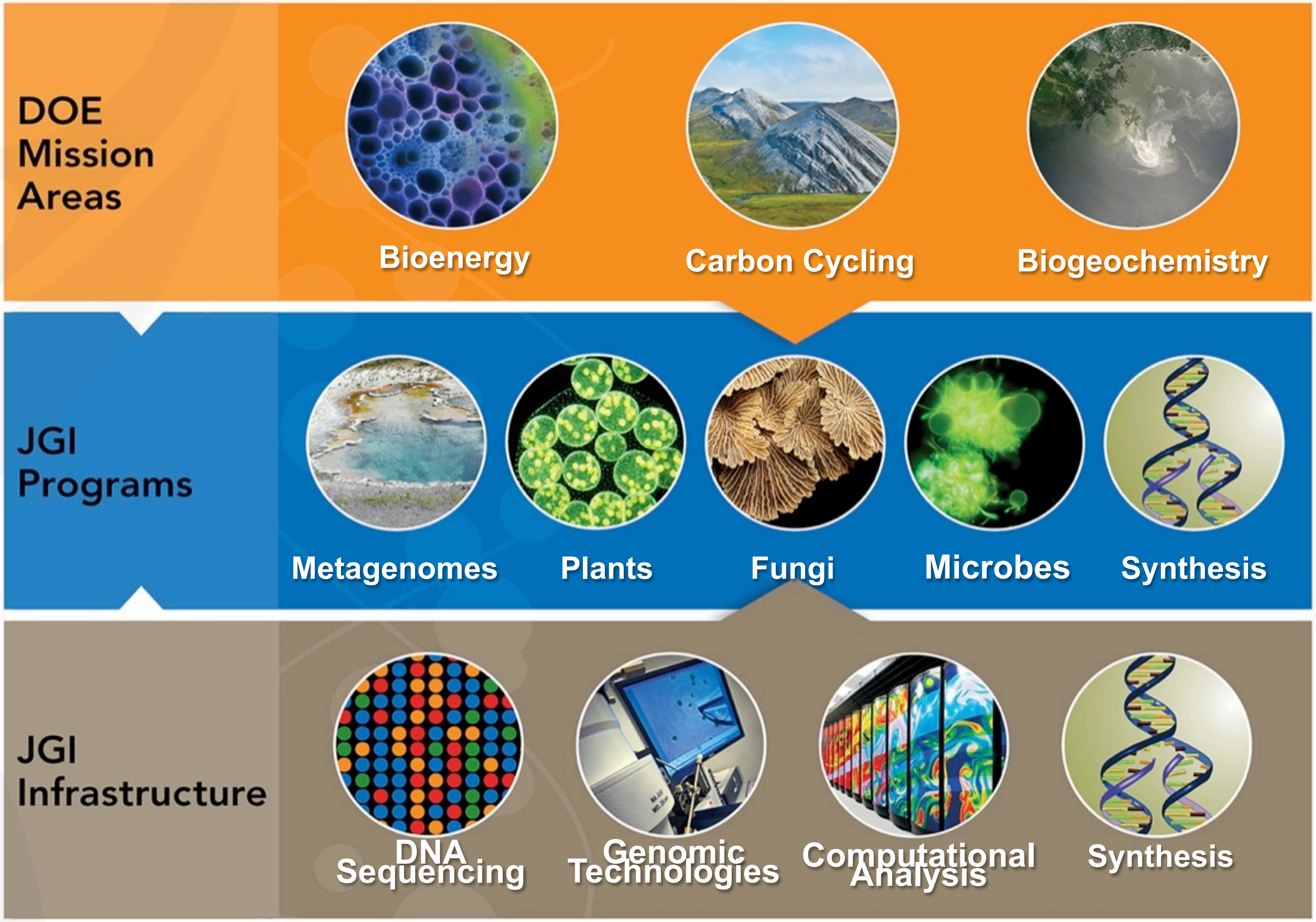
- **Single-cell science vignettes**
  - from symbionts to microbial dark matter
- **Crystal ball**
  - what the future may bring

DOE Joint Genome Institute - a National User Facility

- **DOE runs 50+ national user facilities**
- **Incl. JGI, a high-throughput sequencing & analysis facility**
    - Walnut Creek
    - 250 employees
    - Research
    - New technologies
    - Collaborative research
    - Building user communities

Serving as a genomic user facility in support of the DOE missions . . .

JGI Programs & Infrastructure



DOE Mission Areas

Bioenergy

Carbon Cycling

Biogeochemistry

JGI Programs

Metagenomes

Plants

Fungi

Microbes

Synthesis

JGI Infrastructure

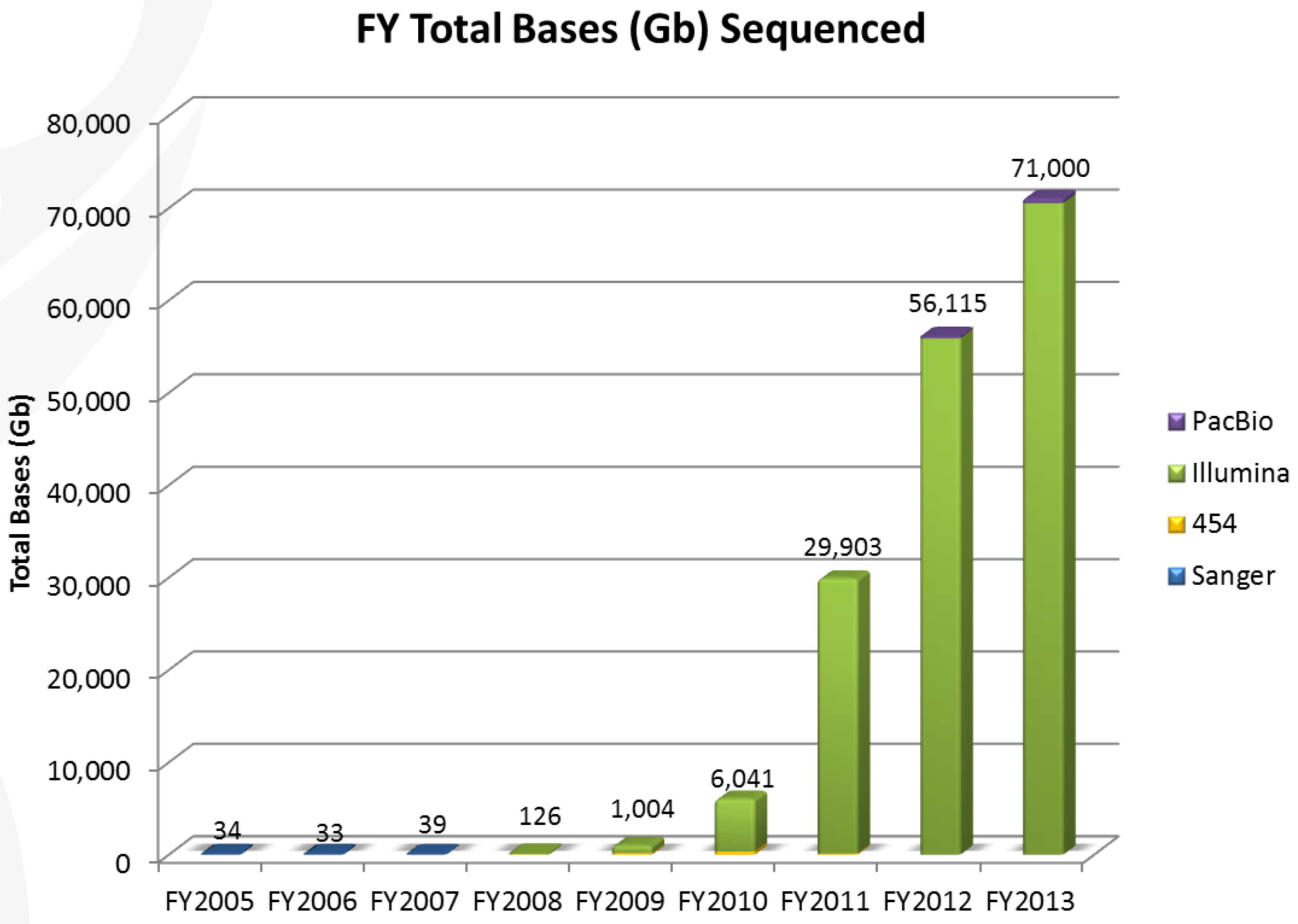
DNA Sequencing

Genomic Technologies

Computational Analysis

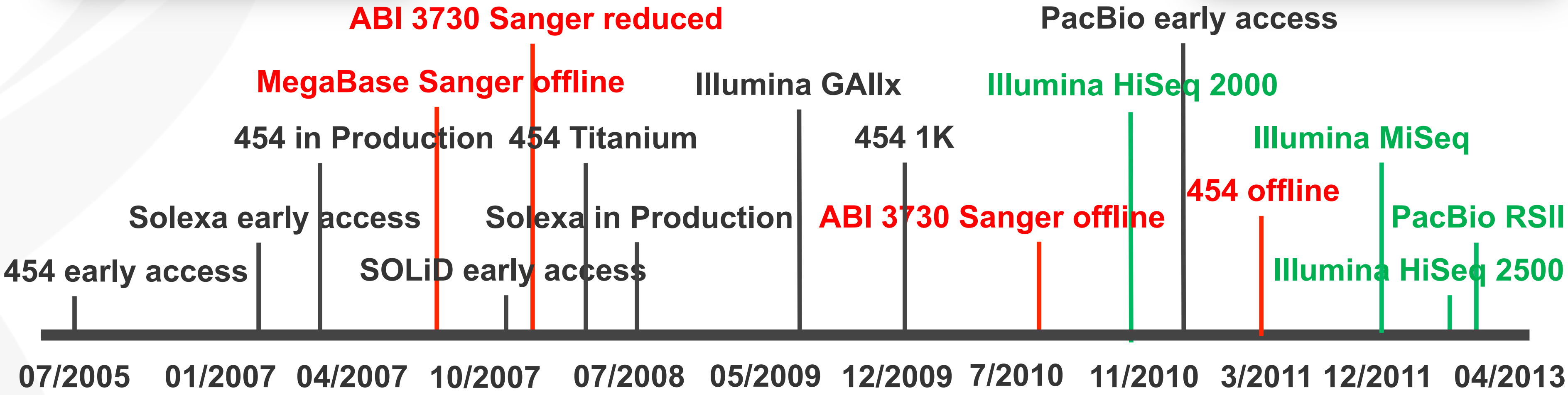
Synthesis

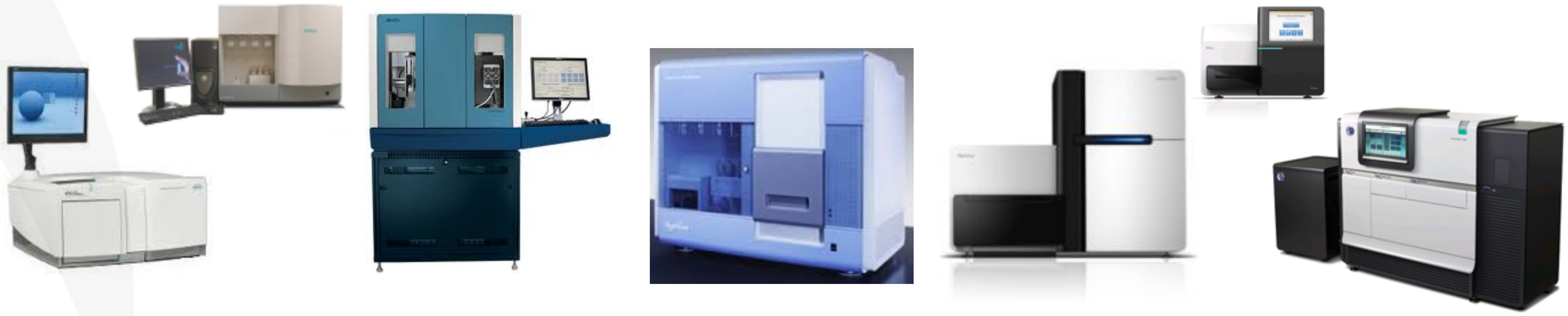
JGI Sequencing Output



Staying State of the Art

Sanger Sequencing to Next-Gen Sequencing by Synthesis





JGI Sequencing Platforms

	Illumina HiSeq 2000	Illumina HiSeq 2500	Illumina MiSeq	Pacific Biosciences RSII
Units	5	3	5	2
Applications	Primary Sequence Generator at JGI	Rapid output HiSeq	Fos Frags, Library QC, R&D	Assembly improvement, de novo, SynBio validation, methylation/ epigenome

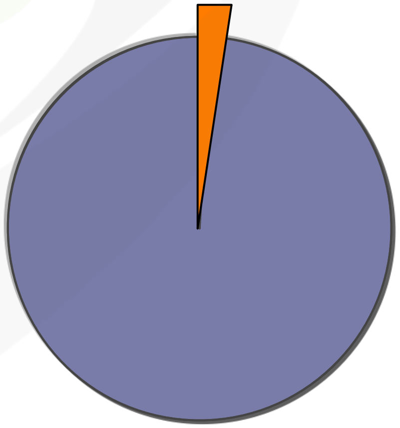
*Not supported by Illumina

Most published single cells have been sequenced on the Illumina platform: HiSeq, (MiSeq)

- **A few words about the JGI**
  - who we are & what we do
- **Single-cell genomics**
  - the why, the how & what to expect from it

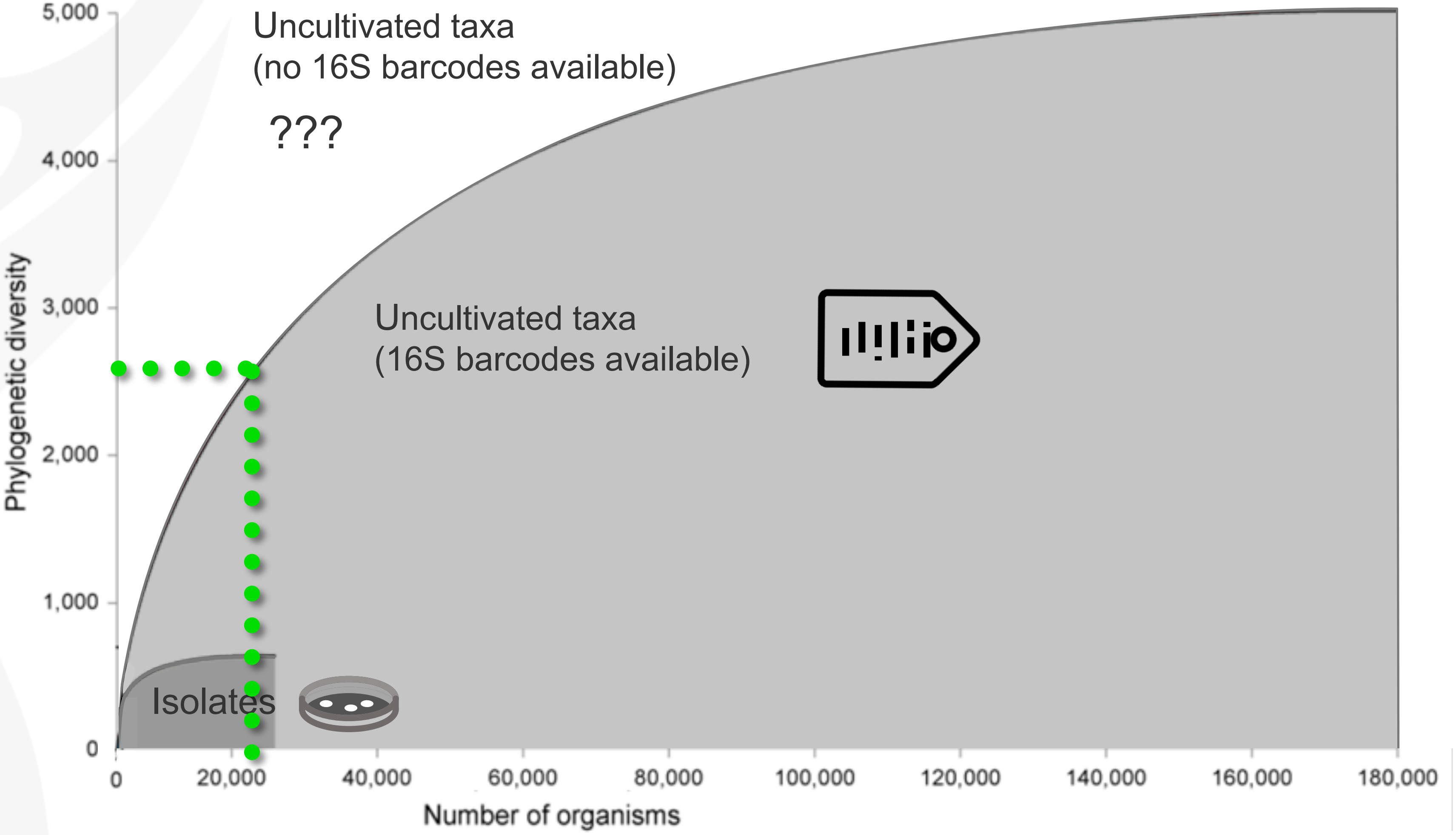
- **Single-cell science vignettes**
  - from symbionts to microbial dark matter
- **Crystal ball**
  - what the future may bring

The cultivated minority

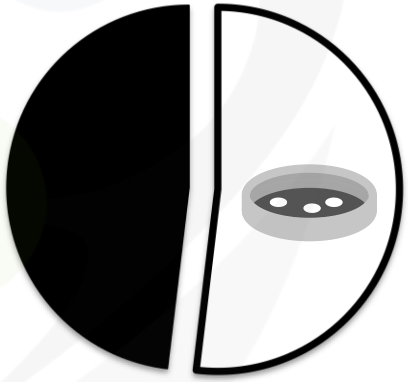


ecological functions & metabolic capabilities largely unknown

In the context of phylogenetic diversity

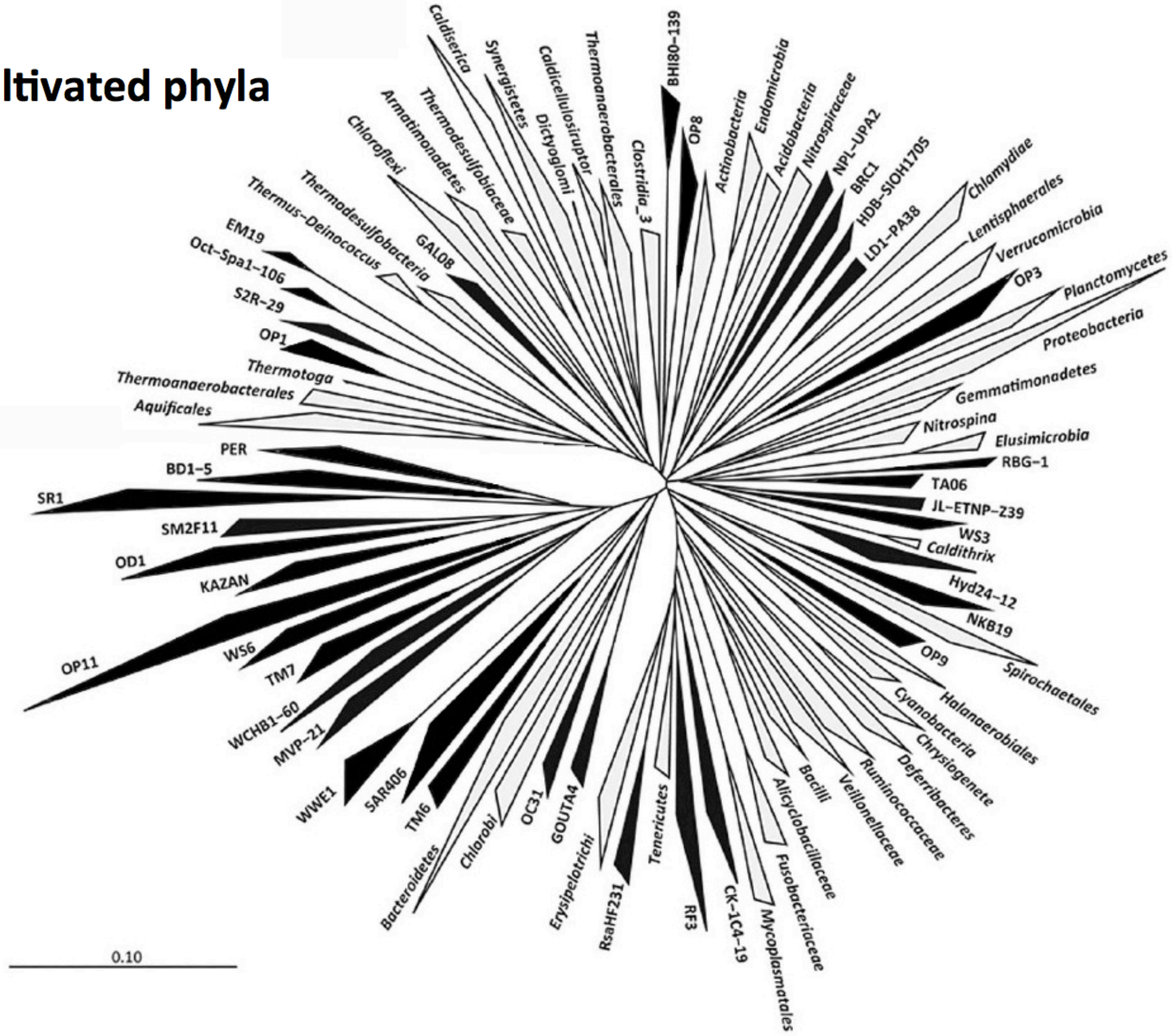


Major uncultivated branches in the bacterial ToL



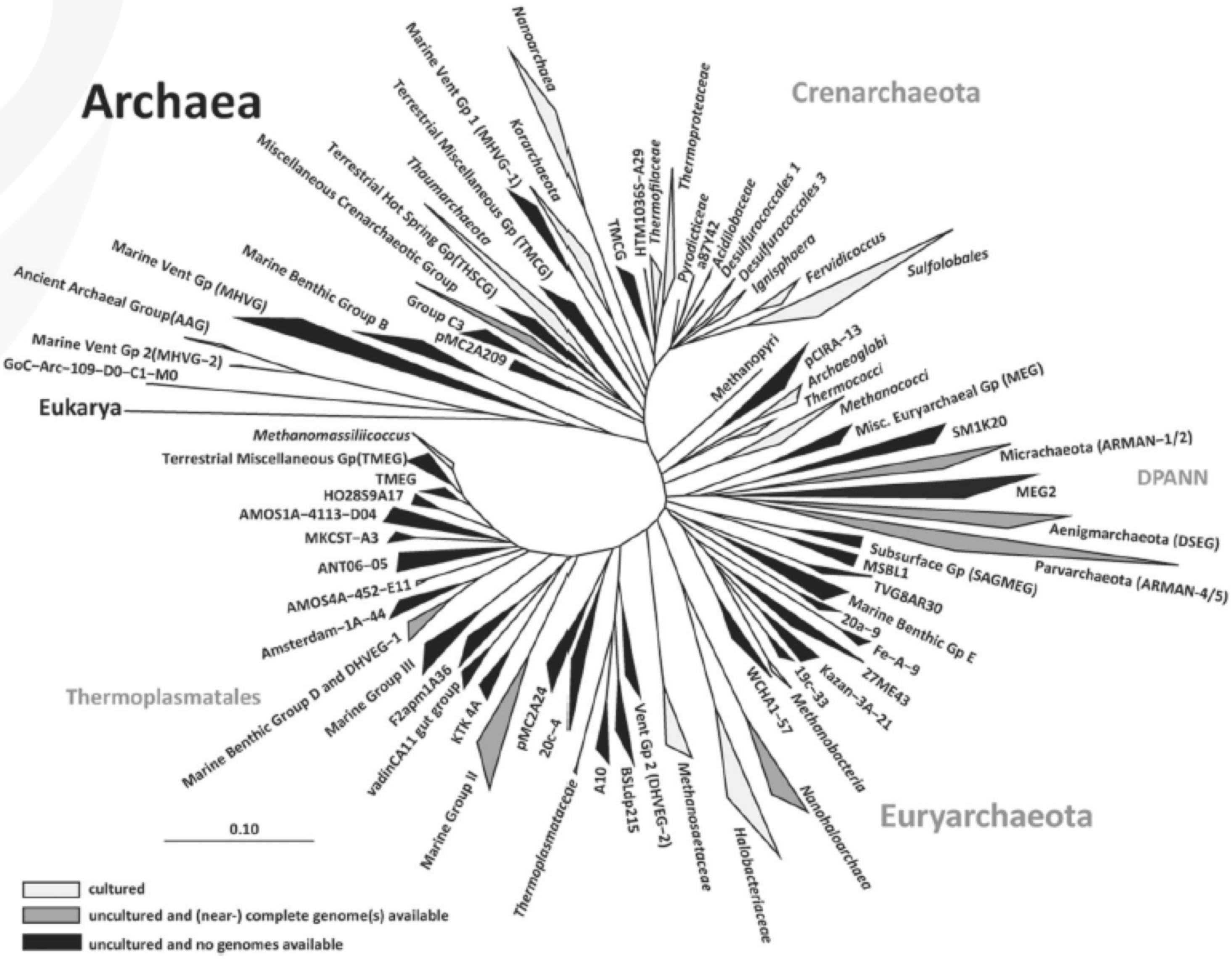
☐ **cultivated phyla**

■ **uncultivated phyla**



Modified from Baker et al 2013 (Microbe)

The situation is similar for archaea



Baker et al 2013 (Microbe)

EM3

GN01-GN05, GN09, GN13, GN14

BRC1

“guerrero negro"

“bacterial rice cluster”

NKB19

AC1

OD1

Marine group A = SAR406

“OP11-derived 1”

NC10

OP1-OP4, OP6–OP9, OP11-12

“Nullarbor cave”

EM19

“obsidian pool”

CD12

WYO

SPAM

SR1

TM6, TM7,

“spring alpine meadow”

“sulphur river”

ZB3

WS1-5, WS6

WWE1, WWE3

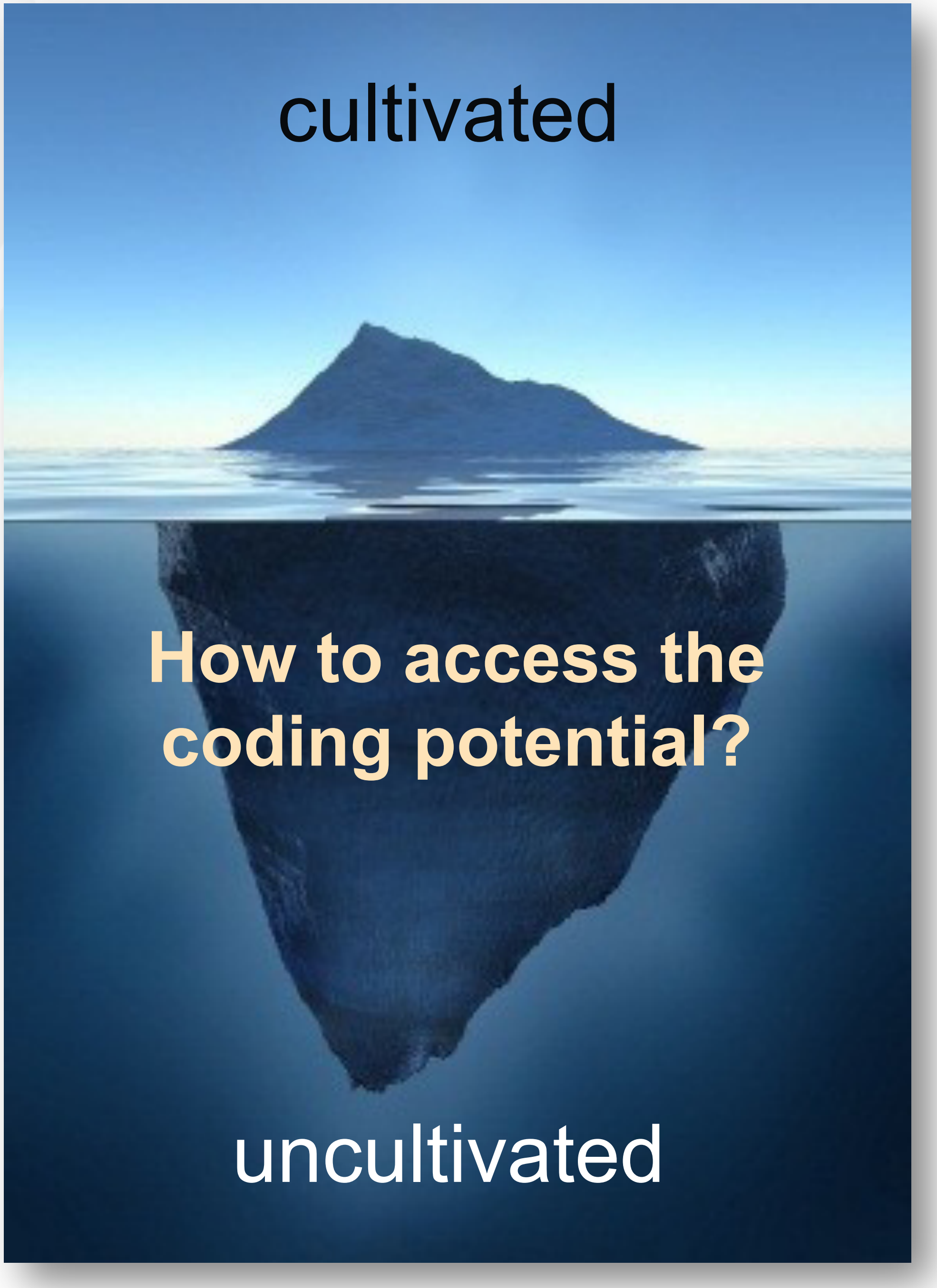
Sediment-1 – Sediment-4

KSB2, KSB3

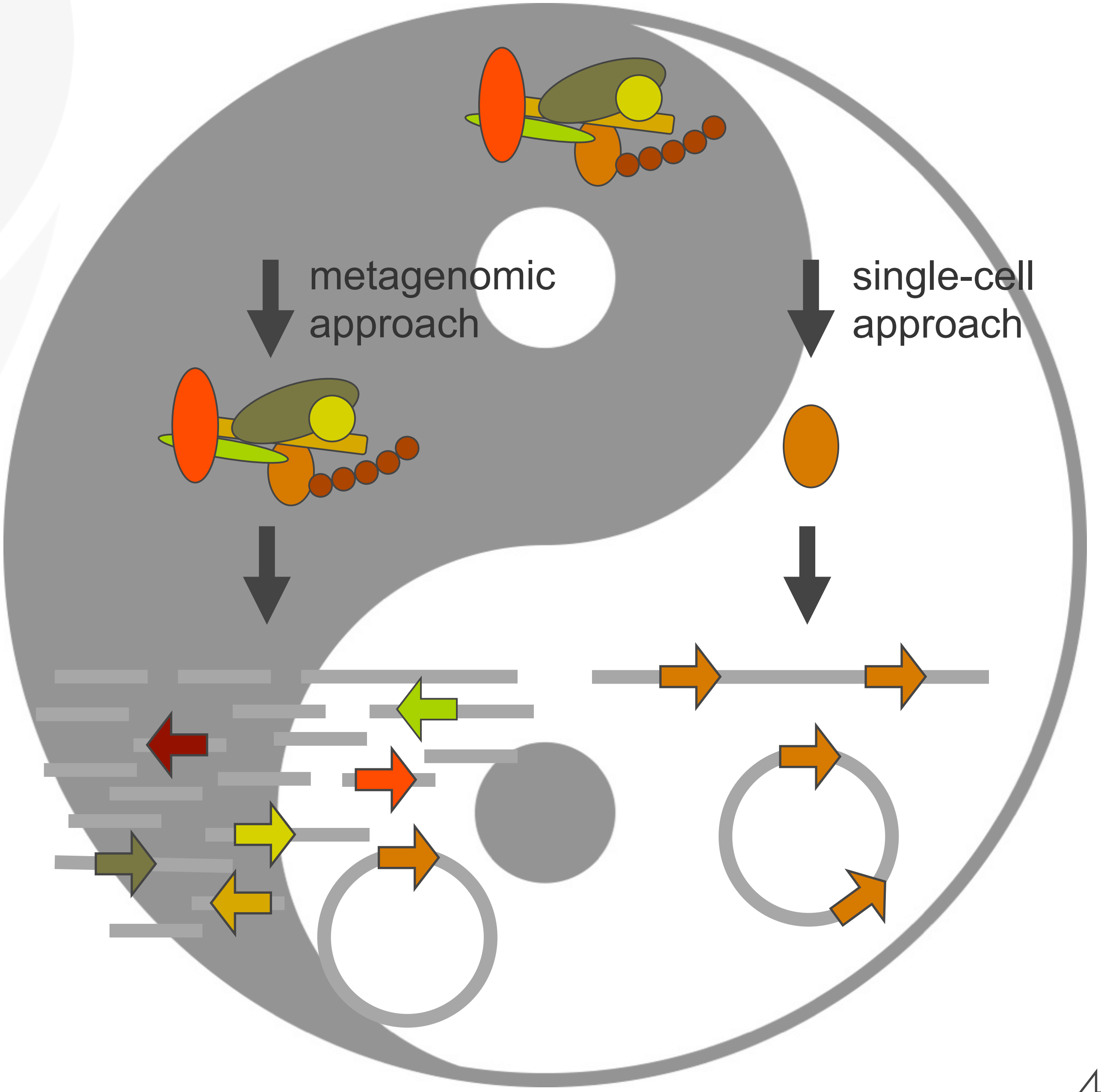
KB1 group

MSBL2, MSBL3, MSBL5, MSBL6

TG2, TG3



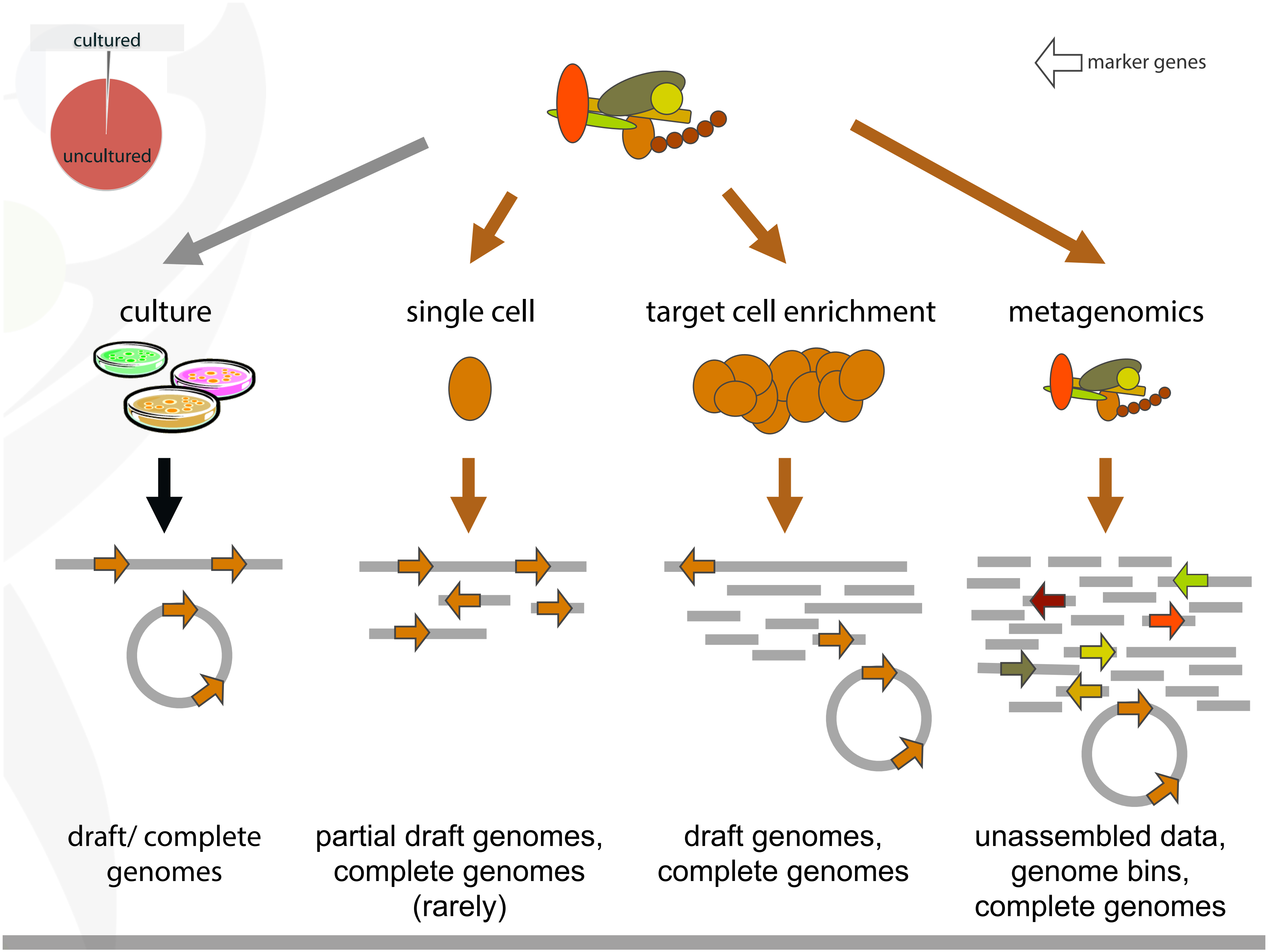
How to access the coding potential?



metagenomic
approach

single-cell
approach

⇦ marker genes



culture

single cell

target cell enrichment

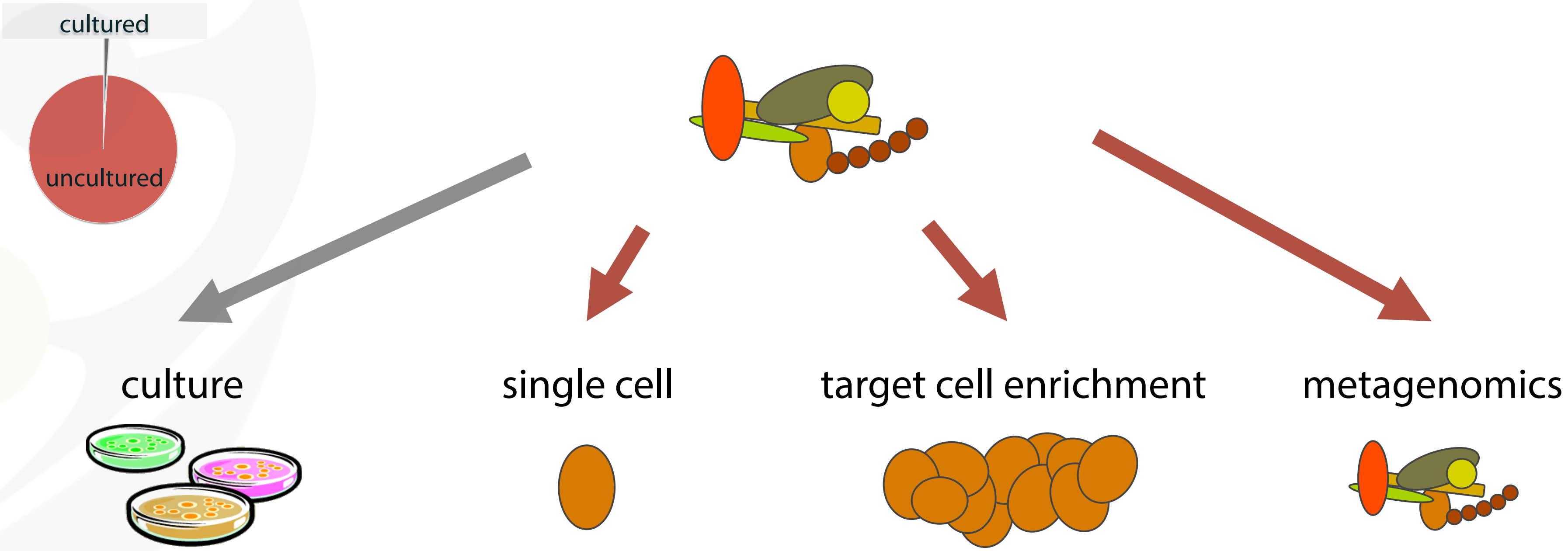
metagenomics

draft/ complete
genomes

partial draft genomes,
complete genomes
(rarely)

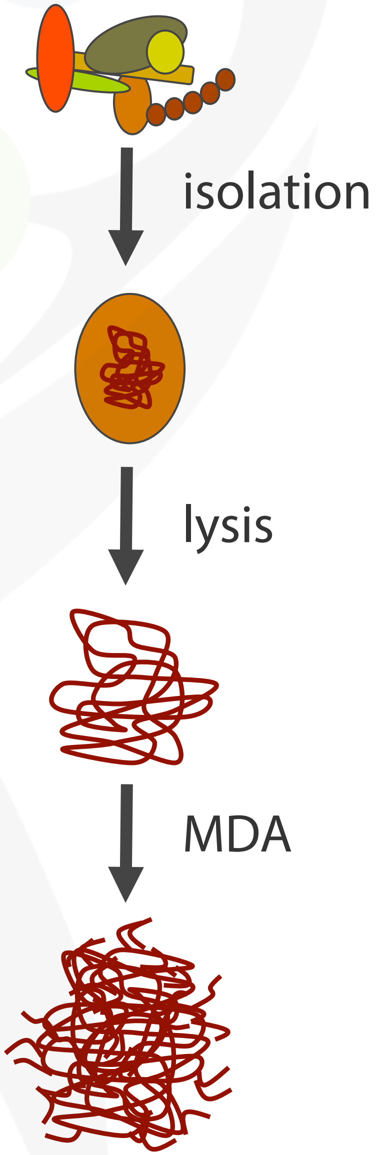
draft genomes,
complete genomes

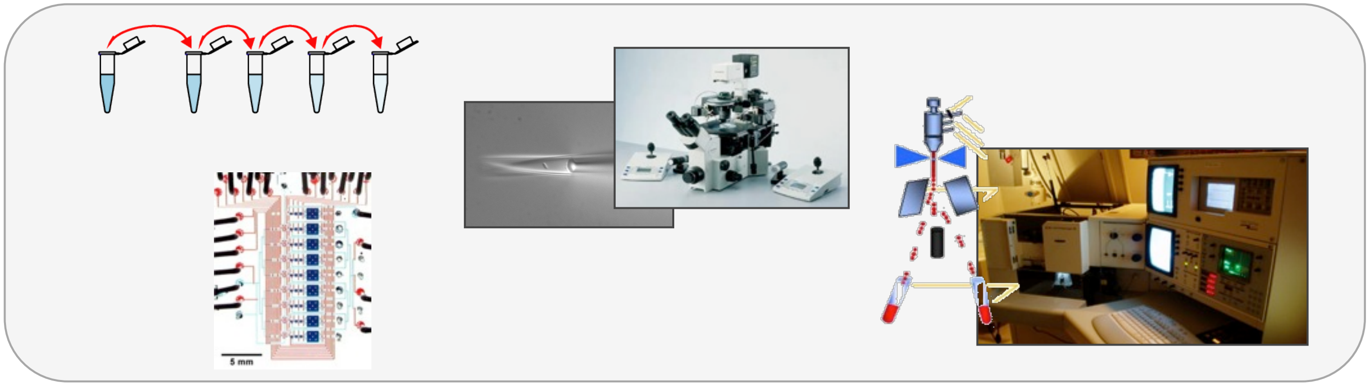
unassembled data,
genome bins,
complete genomes

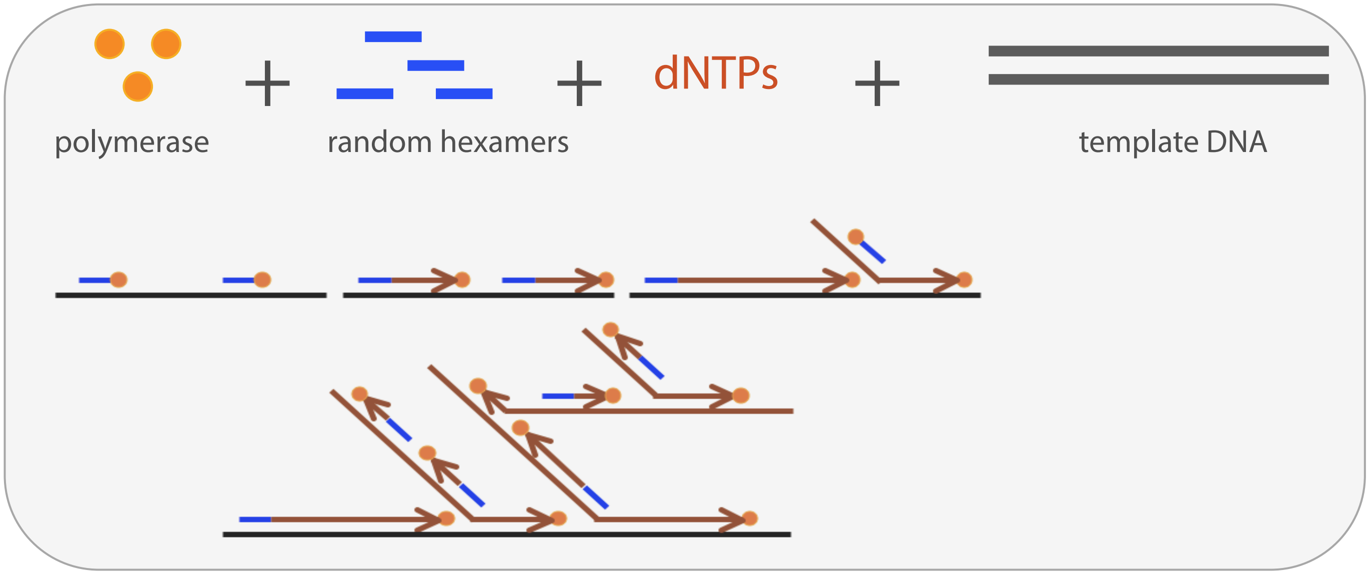


culture	single cell	target cell enrichment	metagenomics
biased: only some players can be cultured	biased: limited lysis & cell isolation	biased: only some organisms susceptible to enrichments	biased: assemblies largely limited to abundant players
applicability: +	applicability: +++	applicability: +	applicability: ++++
culturing can introduce genotype changes	genome snapshot at point in time	genome snapshot at point in time	genome snapshot at point in time
axenic: no assembly challenges	heterogenous populations may be dissected	heterogenous populations: assembly challenge	heterogenous populations: assembly challenge
phenotypic characterization / metadata: extensive	phenotypic characterization / metadata: limited to non-existing	phenotypic characterization / metadata: limited to non-existing	phenotypic characterization / metadata: limited to non-existing
generally less expensive	requires more specialized equipment, costly, technical challenges (bias, chimera..)	possibly requires more specialized equipment, costly	generally less expensive

The single-cell approach: how it works

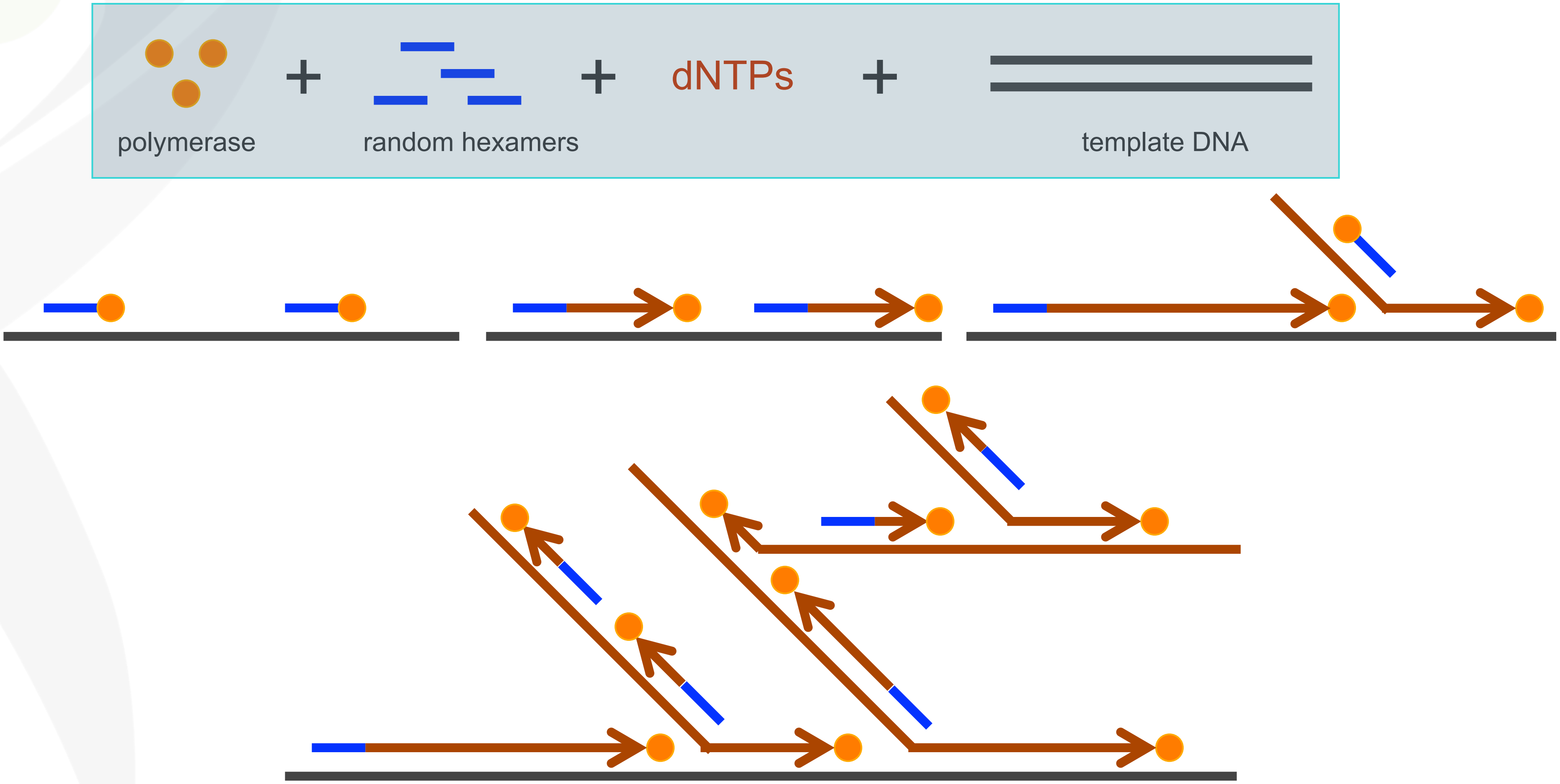




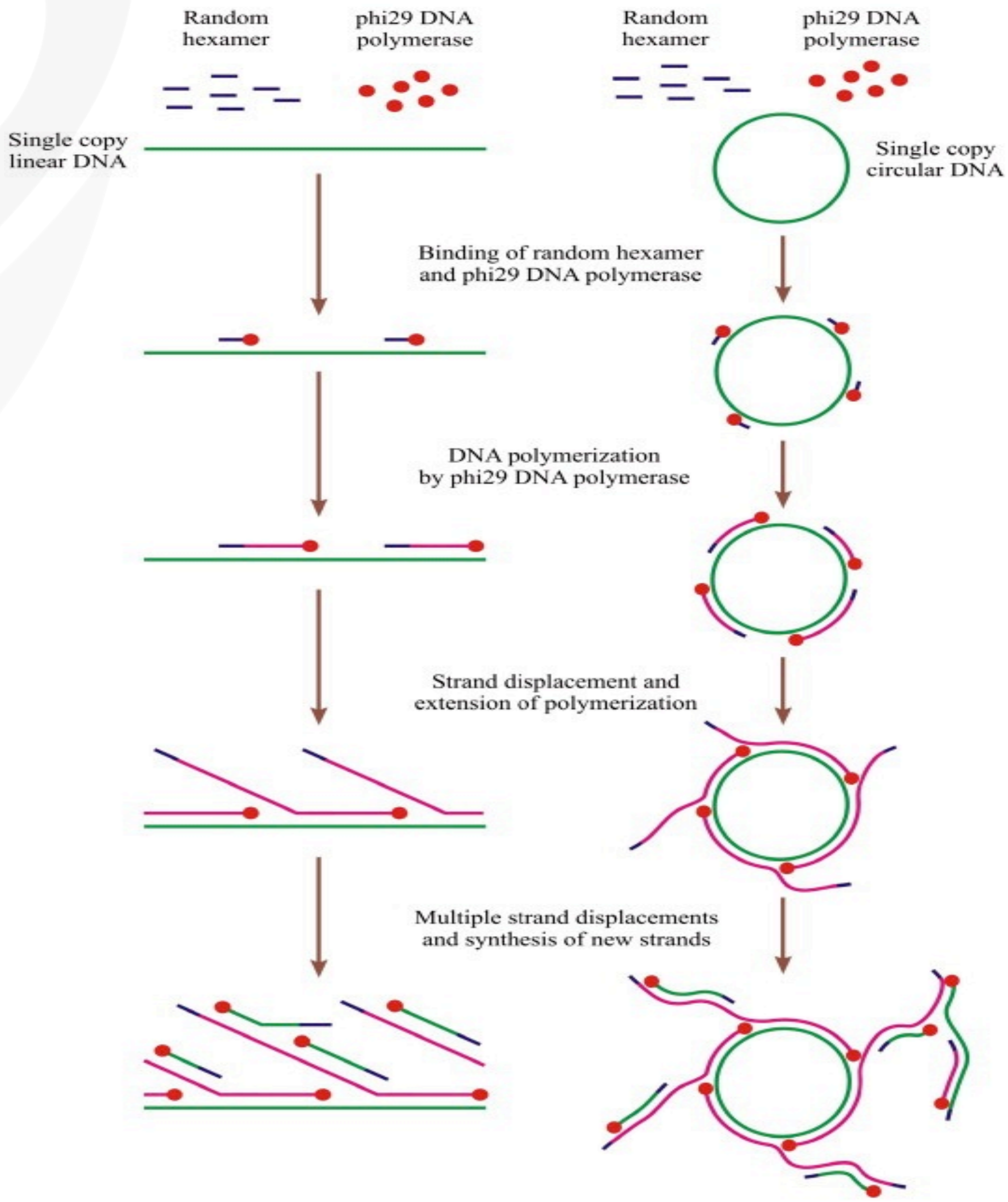


Multiple displacement amplification (MDA)

- isothermal amplification process
- requires polymerase + random primers + dNTPs



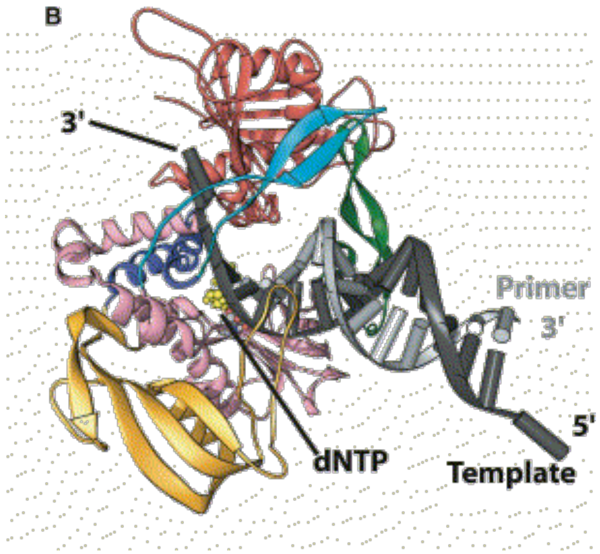
MDA on linear & circular templates



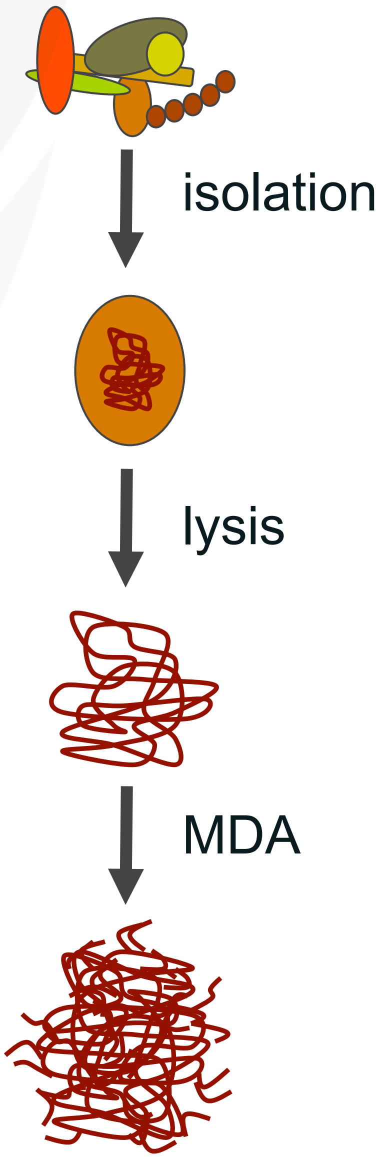
J. Rajendhran et al, 2008
(Biotechnology Advances)

Phi29

- DNA polymerase from phage phi29 (Φ29)
- exceptional strand displacement properties
- --> 100 kb amplicons(!)
- replication at moderate temperatures
- extreme processivity
- 3´→5´ exonuclease activity
- high-fidelity (error rate ~5x10^{-6})



Key challenges



CHALLENGE

Sample contamination
('hitchhiker' DNA)

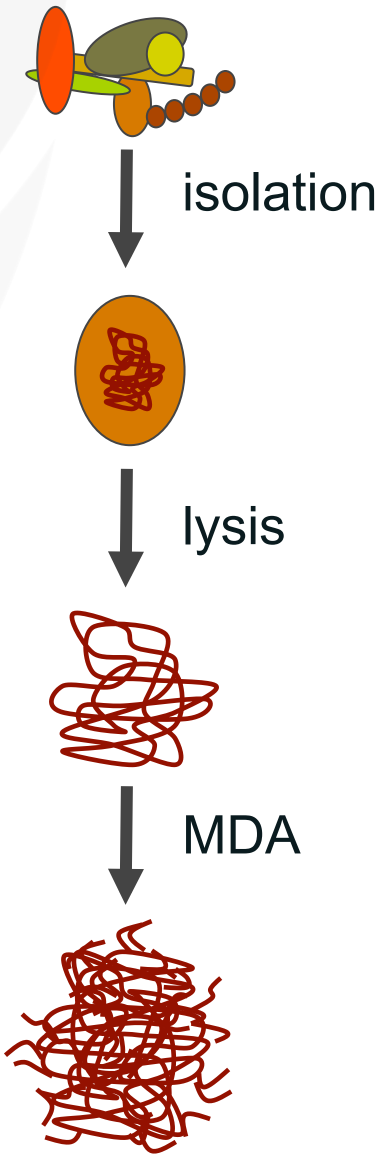
No universal lysis for all taxa

Chimerism

Reagent contamination

MDA bias

Key challenges



isolation

lysis

MDA

CHALLENGE

Sample contamination
(‘hitchhiker’ DNA)

No universal lysis for all taxa

Chimerism

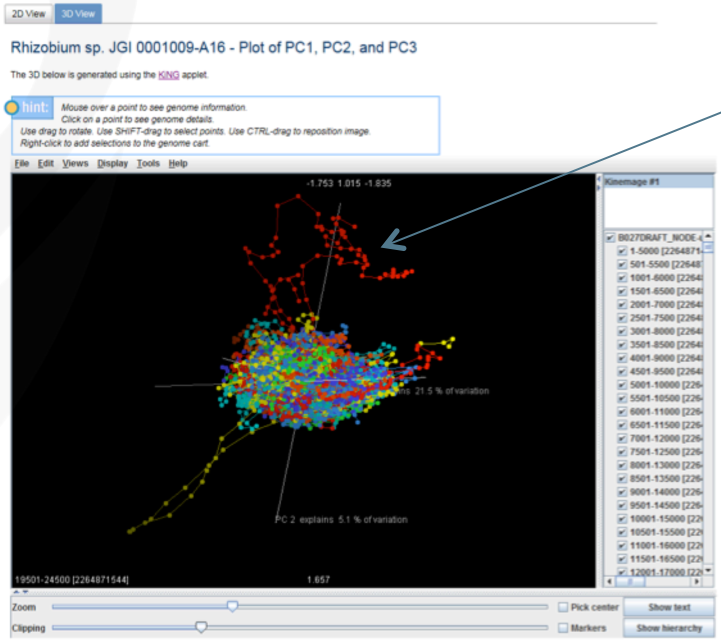
Reagent contamination

MDA bias

QC of single cell data is critical

- 16S rRNA gene and marker gene phylogenetic analyzes
- Tetramer analysis (Kmer Frequency Analysis) and GC contents
- BLAST analyzes and IMG's Phylogenetic profiler
- Removal of contamination via binning and data reload

Tetramer analysis



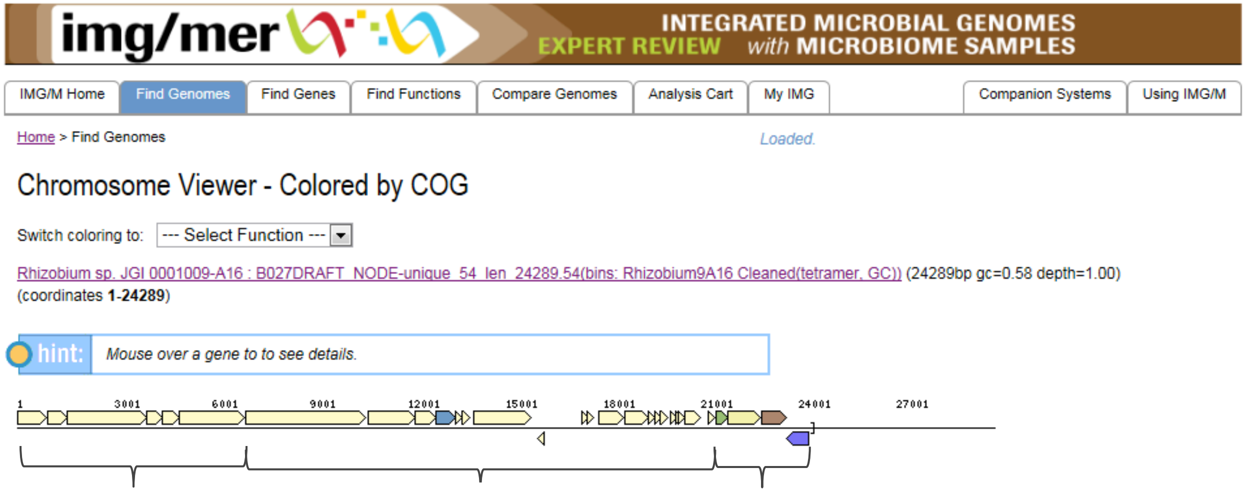
Rhizobium sp. JGI 0001009-A16 - Plot of PC1, PC2, and PC3

The 3D below is generated using the KiNG applet.

hint: *Mouse over a point to see genome information.*
Click on a point to see genome details.
Use drag to rotate. Use SHIFT-drag to select points. Use CTRL-drag to reposition image.
Right-click to add selections to the genome cart.

This red scaffold has points in the main cloud but extends well out.

Clicking on the points in this scaffold opens a separate window with more detail on the scaffold shown below.



Home > Find Genomes

Loaded.

Chromosome Viewer - Colored by COG

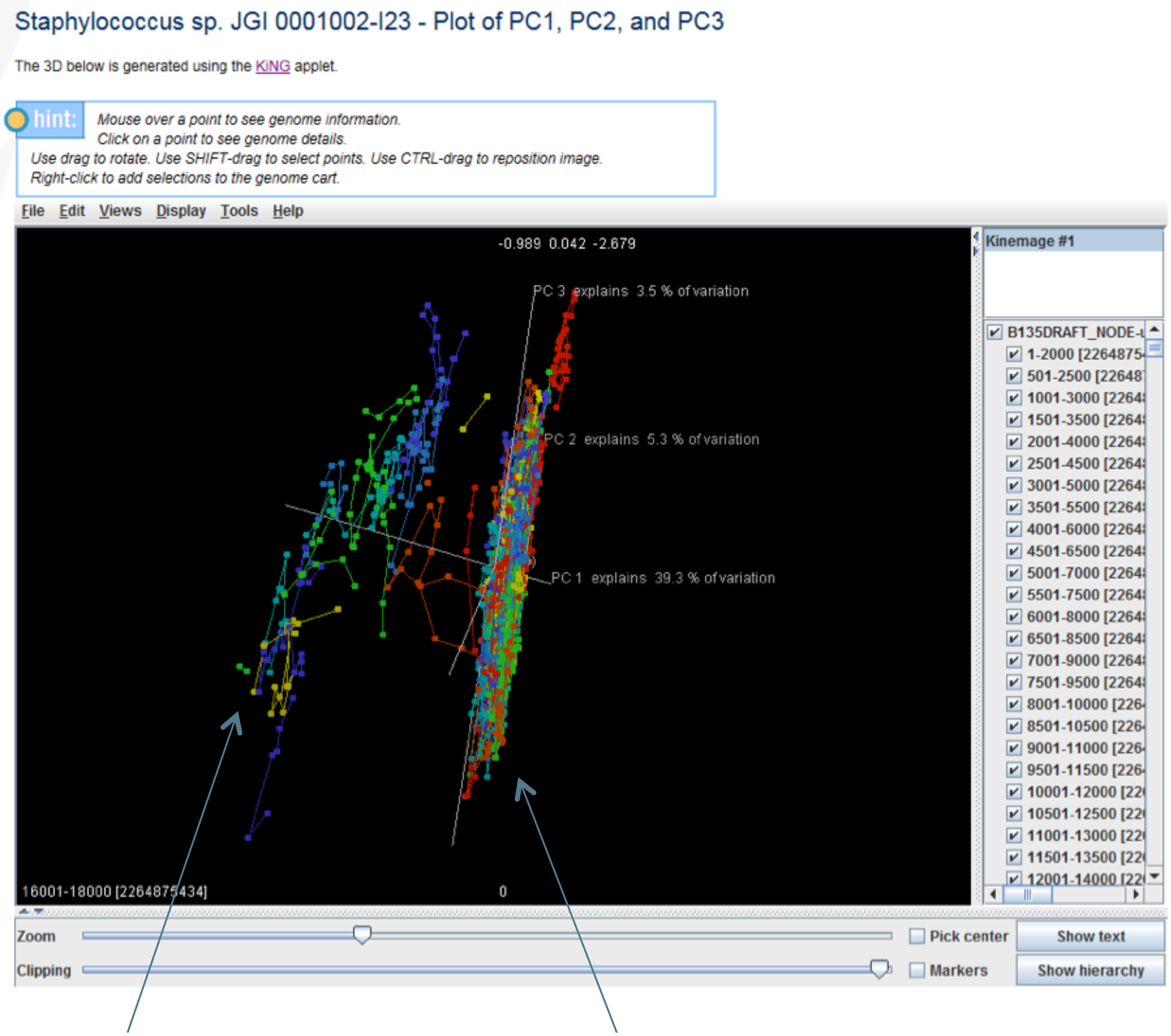
Switch coloring to: --- Select Function ---

Rhizobium sp. JGI 0001009-A16 : B027DRAFT_NODE-unique_54_len_24289.54(bins: Rhizobium9A16 Cleaned(tetramer, GC)) (24289bp gc=0.58 depth=1.00)
(coordinates **1-24289**)

hint: *Mouse over a gene to to see details.*

Rhizobium **Phage** **Rhizobium**

Tetramer analysis



Staphylococcus sp. JGI 0001002-I23 - Plot of PC1, PC2, and PC3

The 3D below is generated using the KiNG applet.

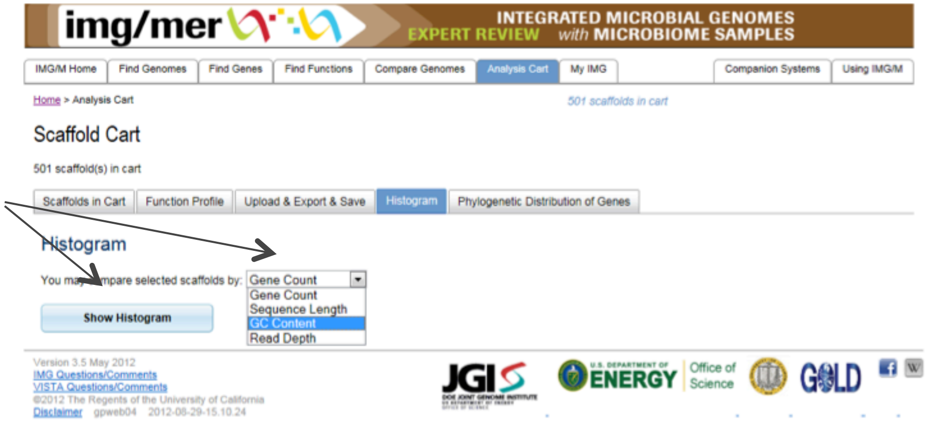
hint: *Mouse over a point to see genome information.*
Click on a point to see genome details.
Use drag to rotate. Use SHIFT-drag to select points. Use CTRL-drag to reposition image.
Right-click to add selections to the genome cart.

Potential Contaminants

Target Organism

GC contents

Select GC Content from drop down menu then click Show Histogram.



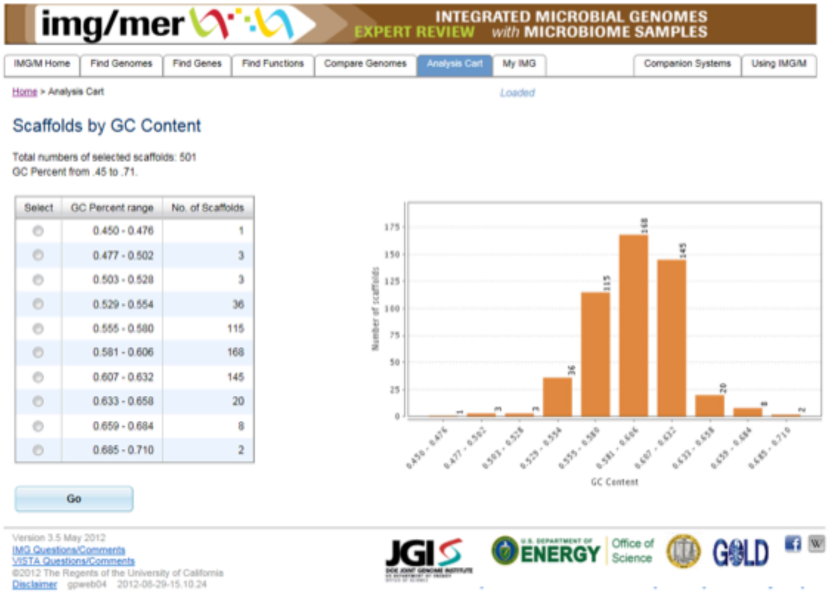
Scaffold Cart

501 scaffold(s) in cart

Histogram

You may compare selected scaffolds by: Gene Count

Version 3.5 May 2012
IMG Questions/Comments
VISTA Questions/Comments
©2012 The Regents of the University of California
Disclaimer gpweb04 2012-08-29-15.10.24



Scaffolds by GC Content

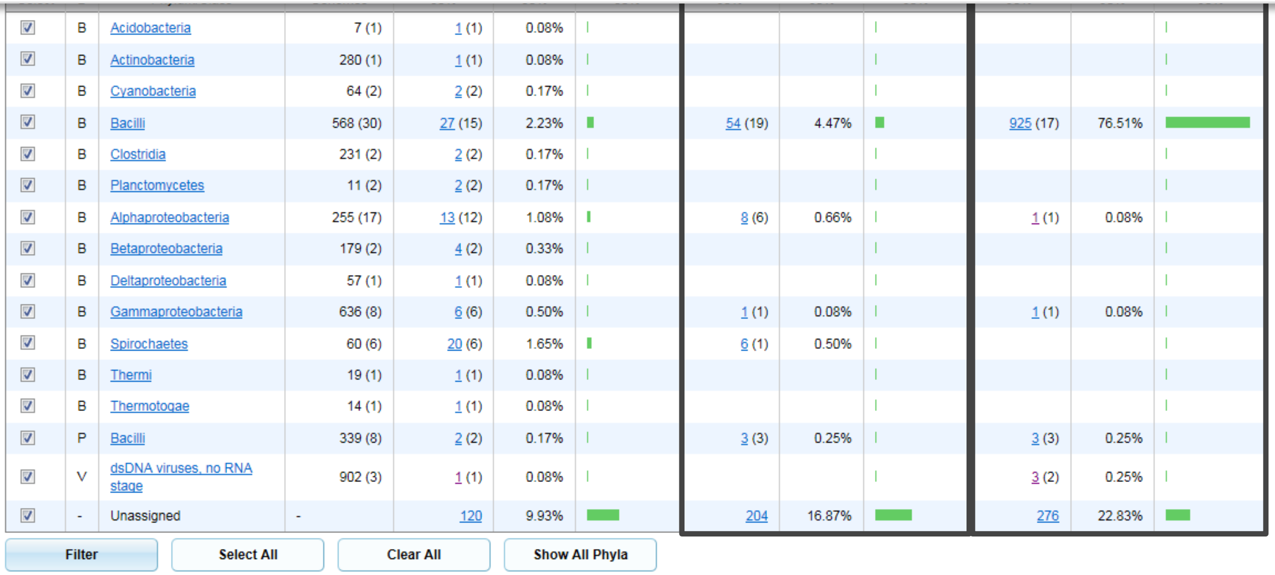
Total numbers of selected scaffolds: 501
GC Percent from .45 to .71.

Select	GC Percent range	No. of Scaffolds
○	0.450 - 0.476	1
○	0.477 - 0.502	3
○	0.503 - 0.528	3
○	0.529 - 0.554	36
○	0.555 - 0.580	115
○	0.581 - 0.606	168
○	0.607 - 0.632	145
○	0.633 - 0.658	20
○	0.659 - 0.684	8
○	0.685 - 0.710	2

Version 3.5 May 2012
IMG Questions/Comments
VISTA Questions/Comments
©2012 The Regents of the University of California
Disclaimer gpweb04 2012-08-29-15.10.24

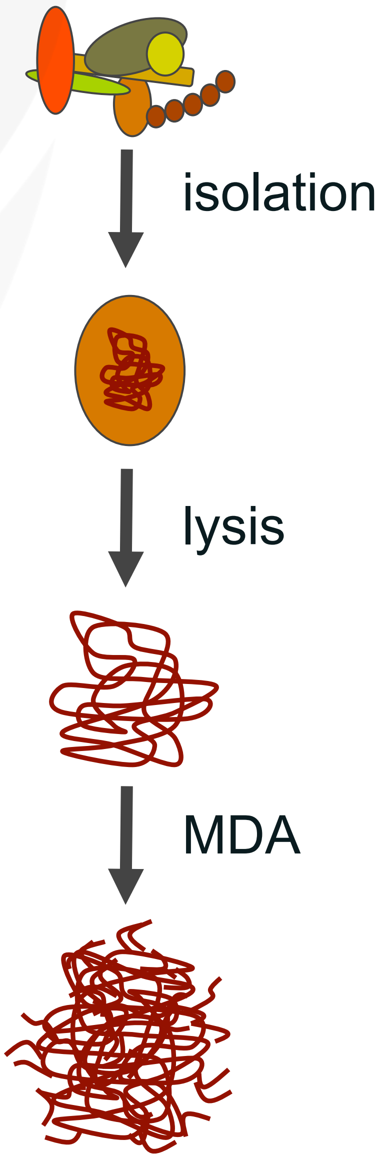
Phylogenetic profiler

Guideline by Scott Clingenpeel
(https://img.jgi.doe.gov)



☑	B	Acidobacteria	7 (1)	1 (1)	0.08%							
☑	B	Actinobacteria	280 (1)	1 (1)	0.08%							
☑	B	Cyanobacteria	64 (2)	2 (2)	0.17%							
☑	B	Bacilli	568 (30)	27 (15)	2.23%	🟩	54 (19)	4.47%	🟩	925 (17)	76.51%	🟩
☑	B	Clostridia	231 (2)	2 (2)	0.17%							
☑	B	Planctomycetes	11 (2)	2 (2)	0.17%							
☑	B	Alphaproteobacteria	255 (17)	13 (12)	1.08%	🟩	8 (6)	0.66%		1 (1)	0.08%	
☑	B	Betaproteobacteria	179 (2)	4 (2)	0.33%							
☑	B	Deltaproteobacteria	57 (1)	1 (1)	0.08%							
☑	B	Gammaproteobacteria	636 (8)	6 (6)	0.50%		1 (1)	0.08%		1 (1)	0.08%	
☑	B	Spirochaetes	60 (6)	20 (6)	1.65%	🟩	6 (1)	0.50%				
☑	B	Thermi	19 (1)	1 (1)	0.08%							
☑	B	Thermotogae	14 (1)	1 (1)	0.08%							
☑	P	Bacilli	339 (8)	2 (2)	0.17%		3 (3)	0.25%		3 (3)	0.25%	
☑	V	dsDNA viruses, no RNA stage	902 (3)	1 (1)	0.08%					3 (2)	0.25%	
☑	-	Unassigned	-	120	9.93%	🟩	204	16.87%	🟩	276	22.83%	🟩

Key challenges



CHALLENGE

Sample contamination
(‘hitchhiker’ DNA)

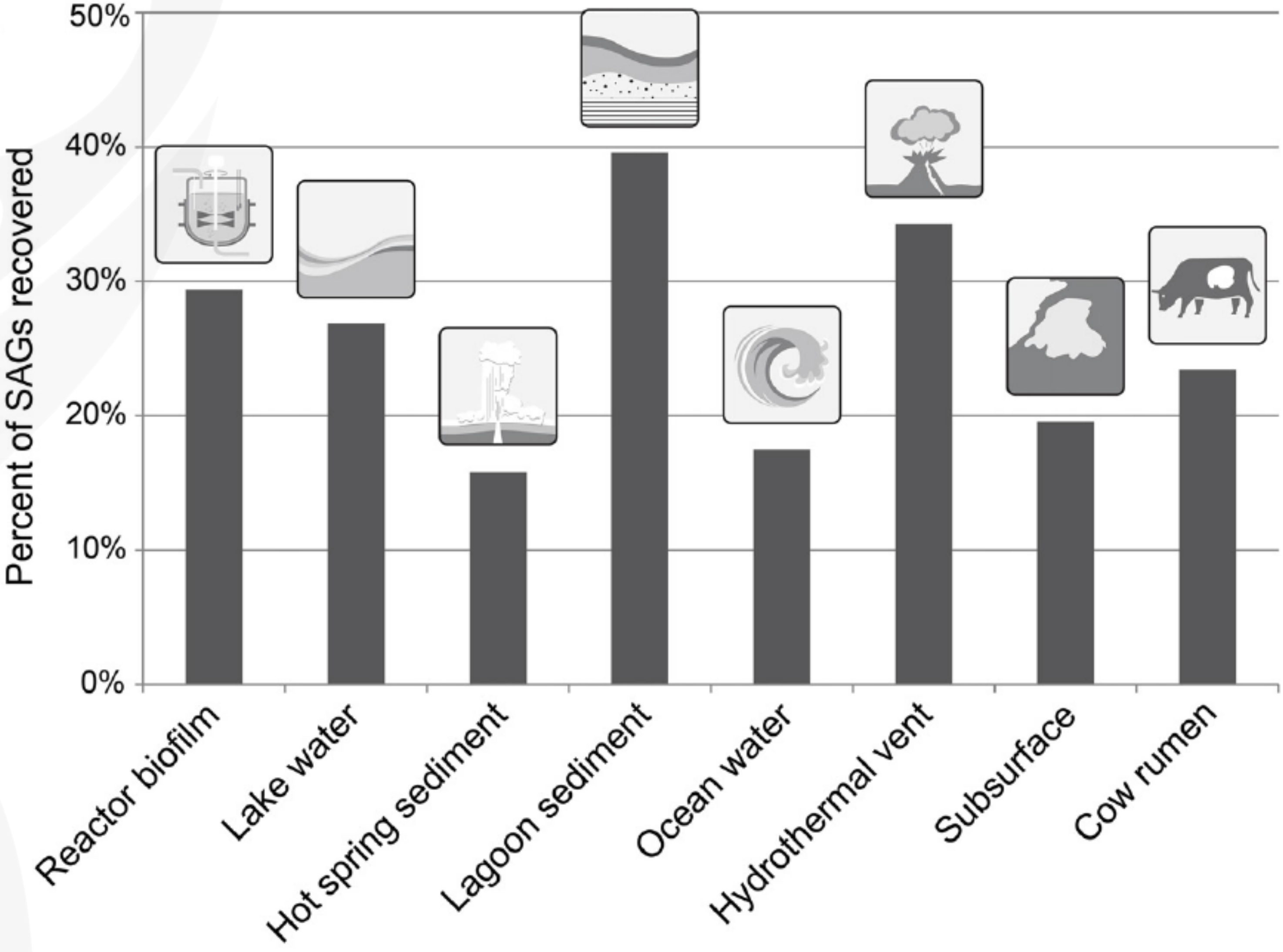
No universal lysis for all taxa

Chimerism

Reagent contamination

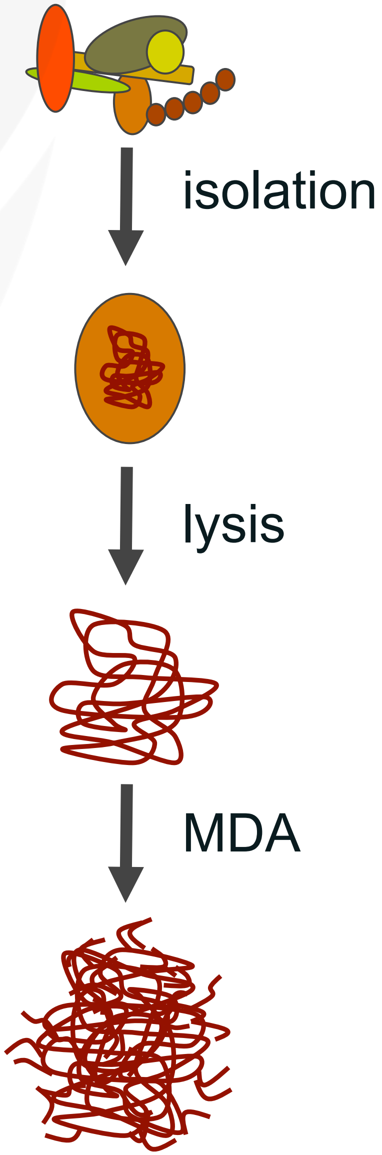
MDA bias

A fraction of single cells can be recovered



Clingenpeel et al 2014 (Frontiers in Microbiol)

Key challenges



CHALLENGE

Sample contamination
(‘hitchhiker’ DNA)

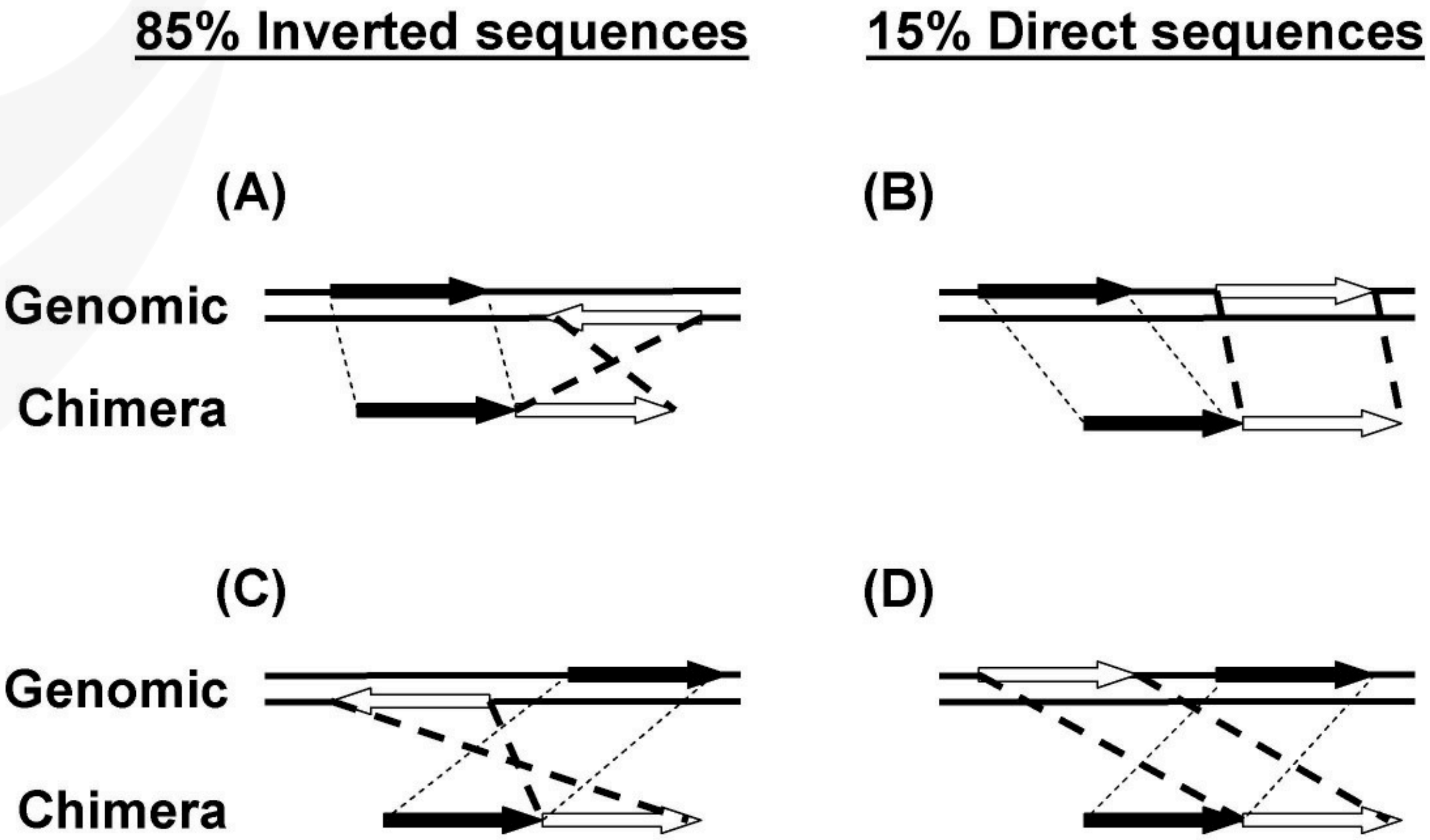
No universal lysis for all taxa

Chimerism

Reagent contamination

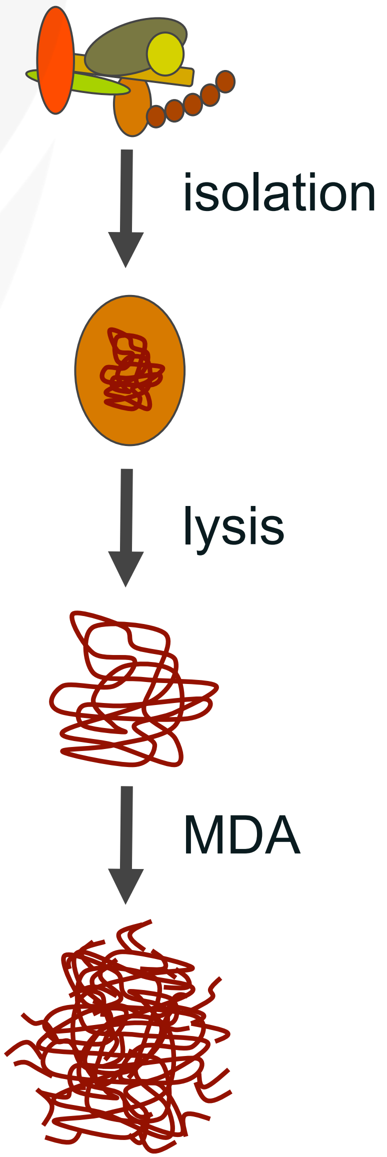
MDA bias

Chimeric rearrangements



Lasken et al 2007 (BMC Biotechnology)

Key challenges



isolation

lysis

MDA

CHALLENGE

Sample contamination
(‘hitchhiker’ DNA)

No universal lysis for all taxa

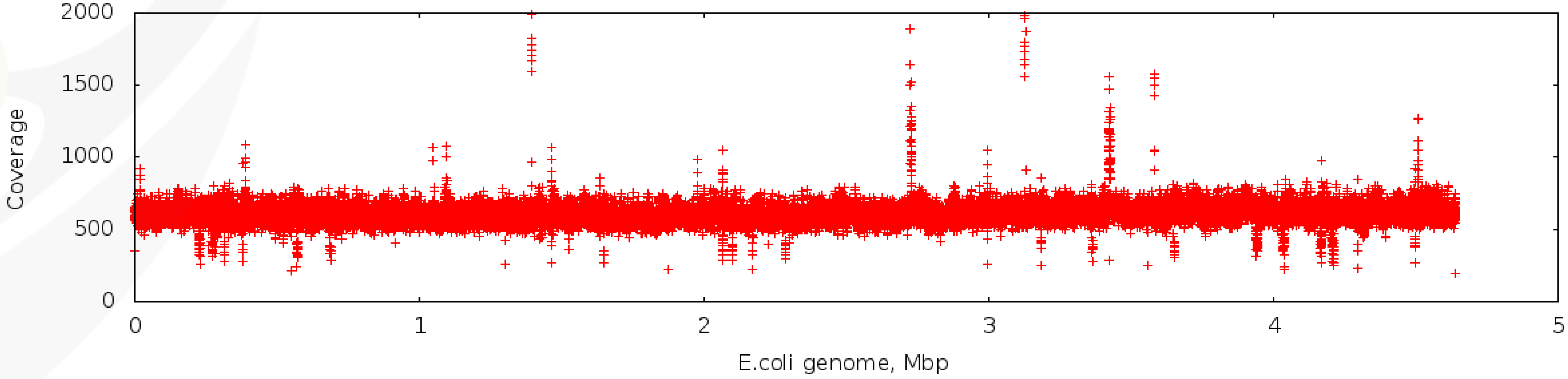
Chimerism

Reagent contamination

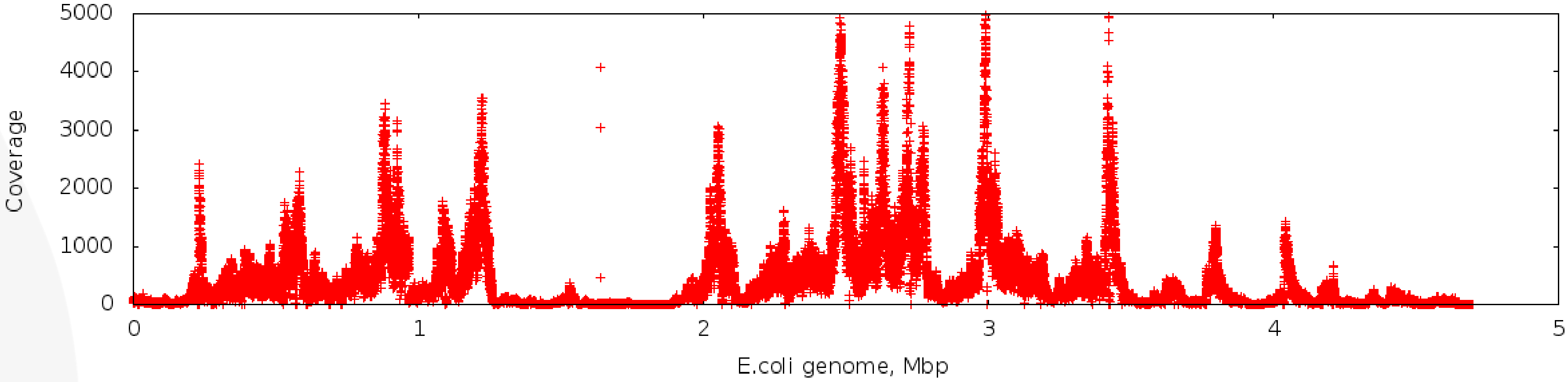
MDA bias

MDA – coverage bias

- *E. coli* isolate dataset



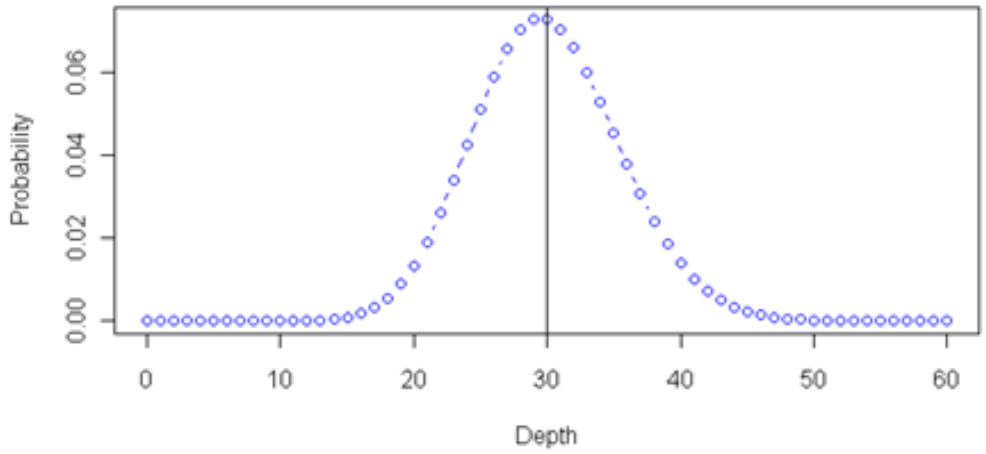
- *E.coli* single-cell dataset

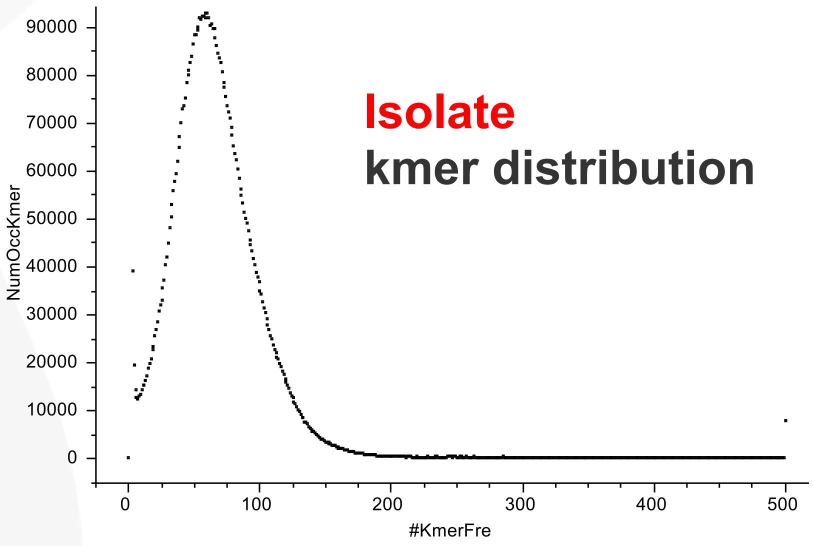


A Copeland

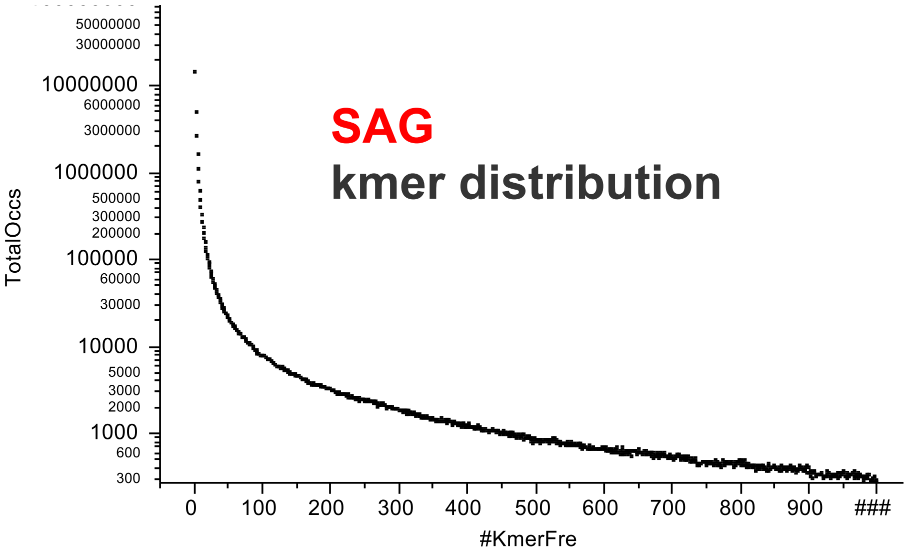
MDA coverage bias

Shotgun sequencing **theoretical** kmer distribution





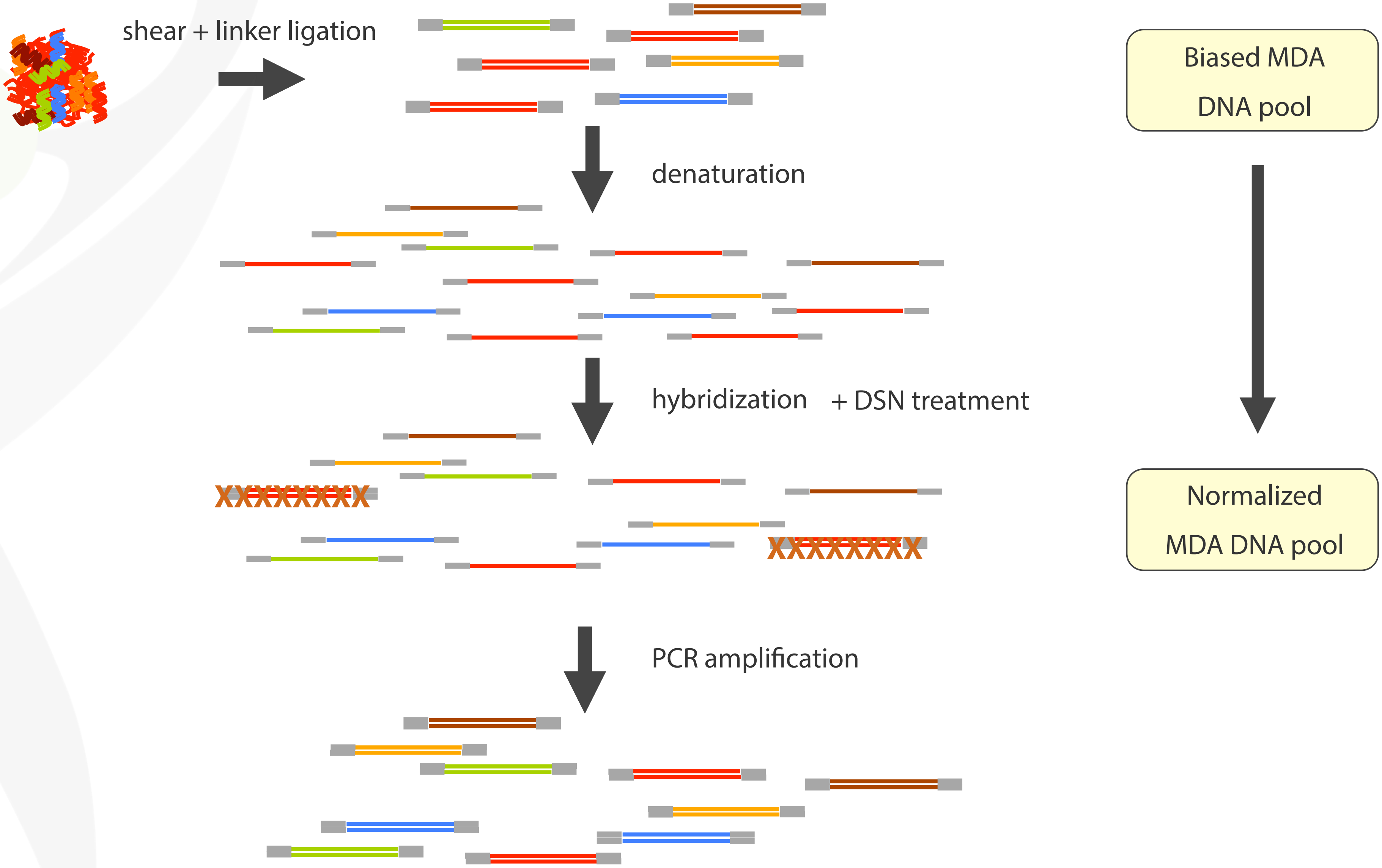
Isolate
kmer distribution



SAG
kmer distribution

A Copeland

MDA DNA normalization (wet lab)



shear + linker ligation

denaturation

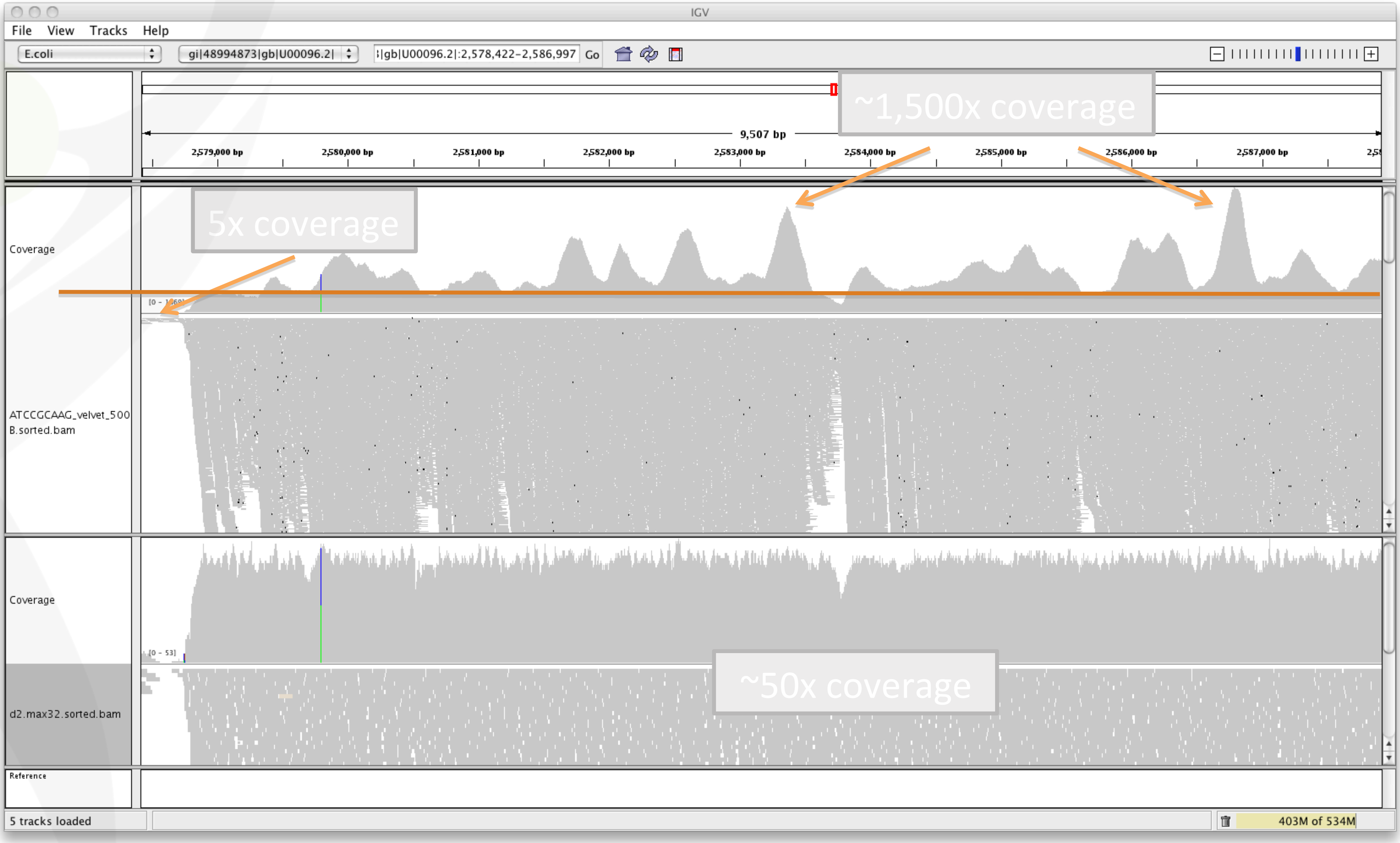
hybridization + DSN treatment

PCR amplification

Biased MDA DNA pool

Normalized MDA DNA pool

Coverage normalization (computational)

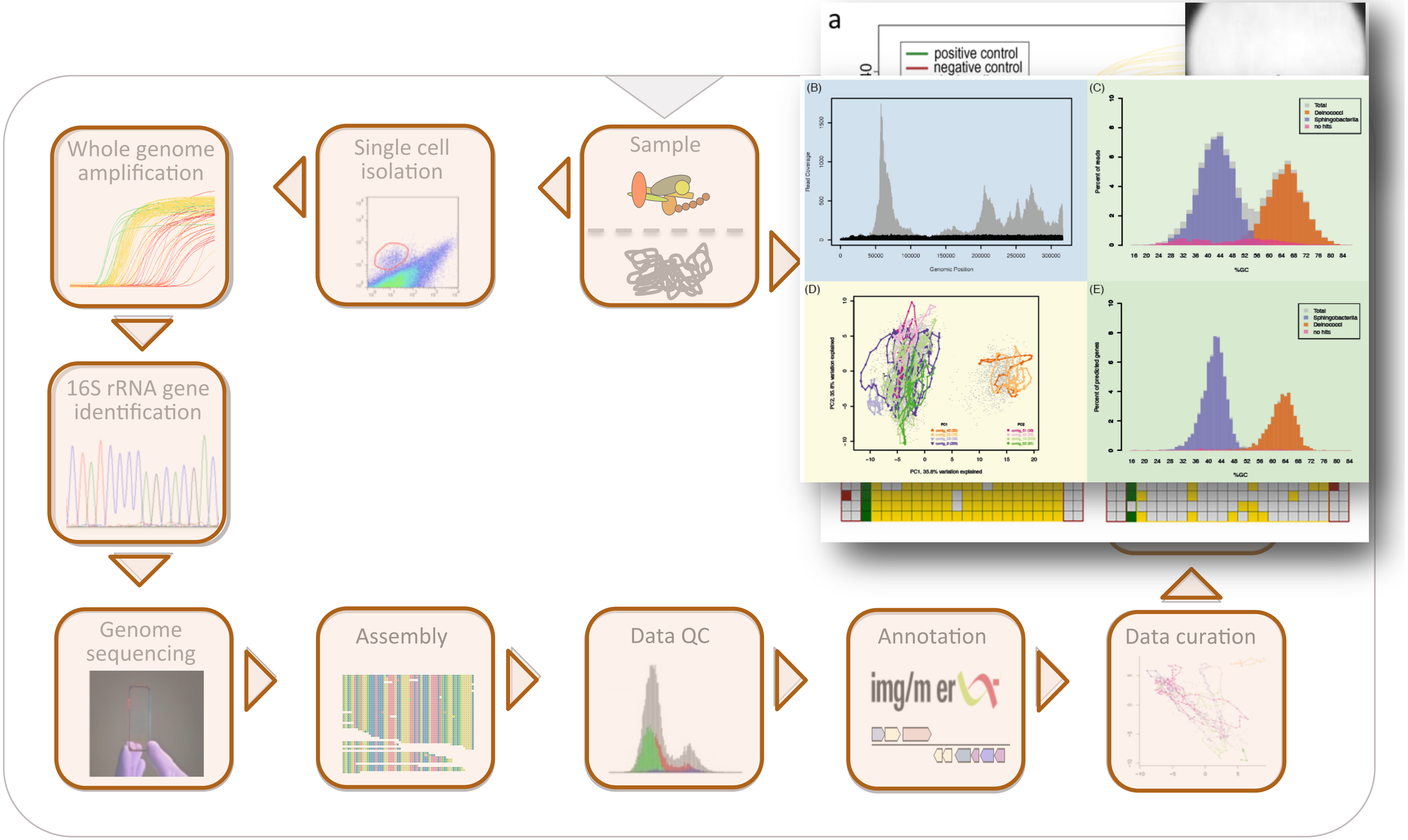


~1,500x coverage

5x coverage

~50x coverage

JGI single-cell sequencing pipeline



Sample

Single cell isolation

Whole genome amplification

16S rRNA gene identification

Genome sequencing

Assembly

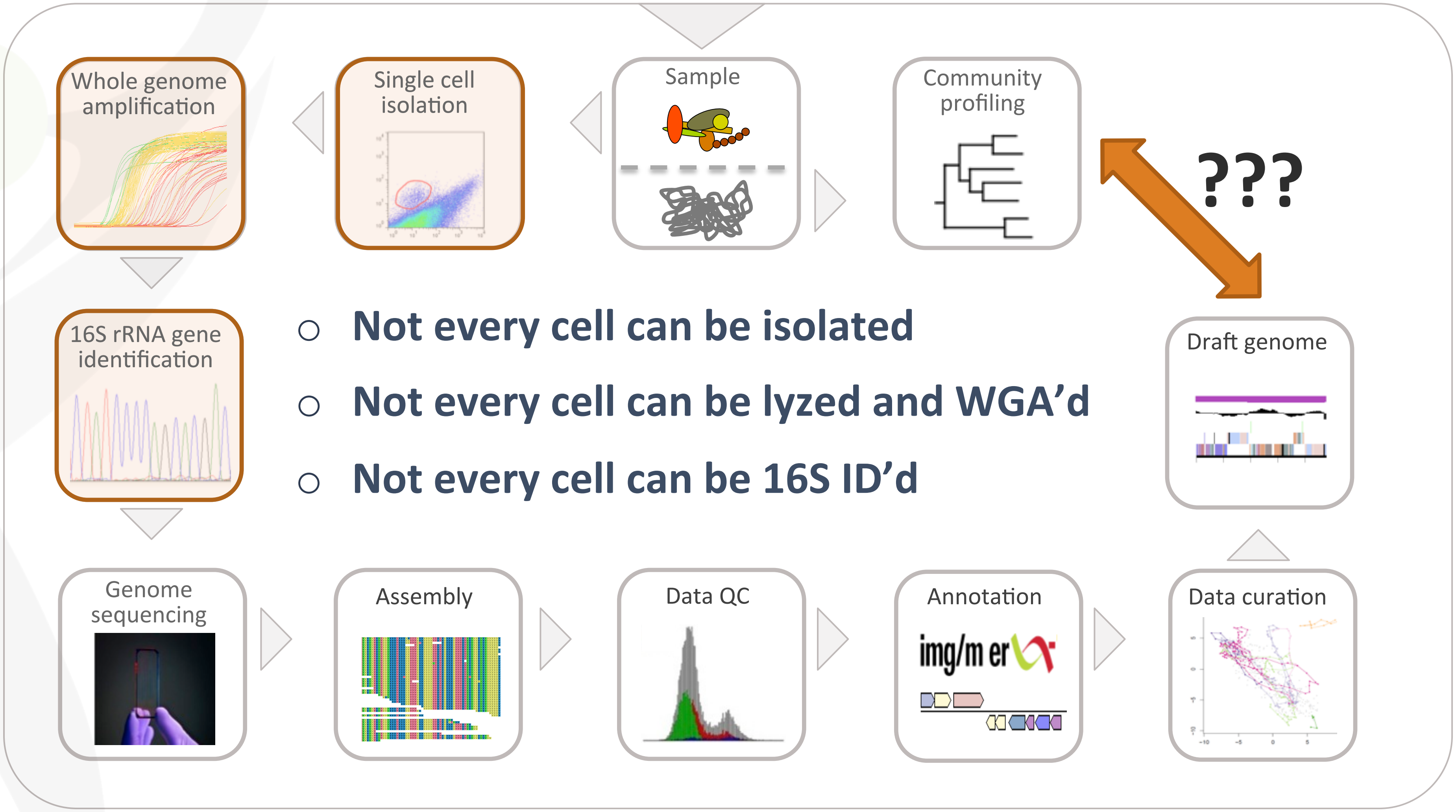
Data QC

Annotation

Data curation

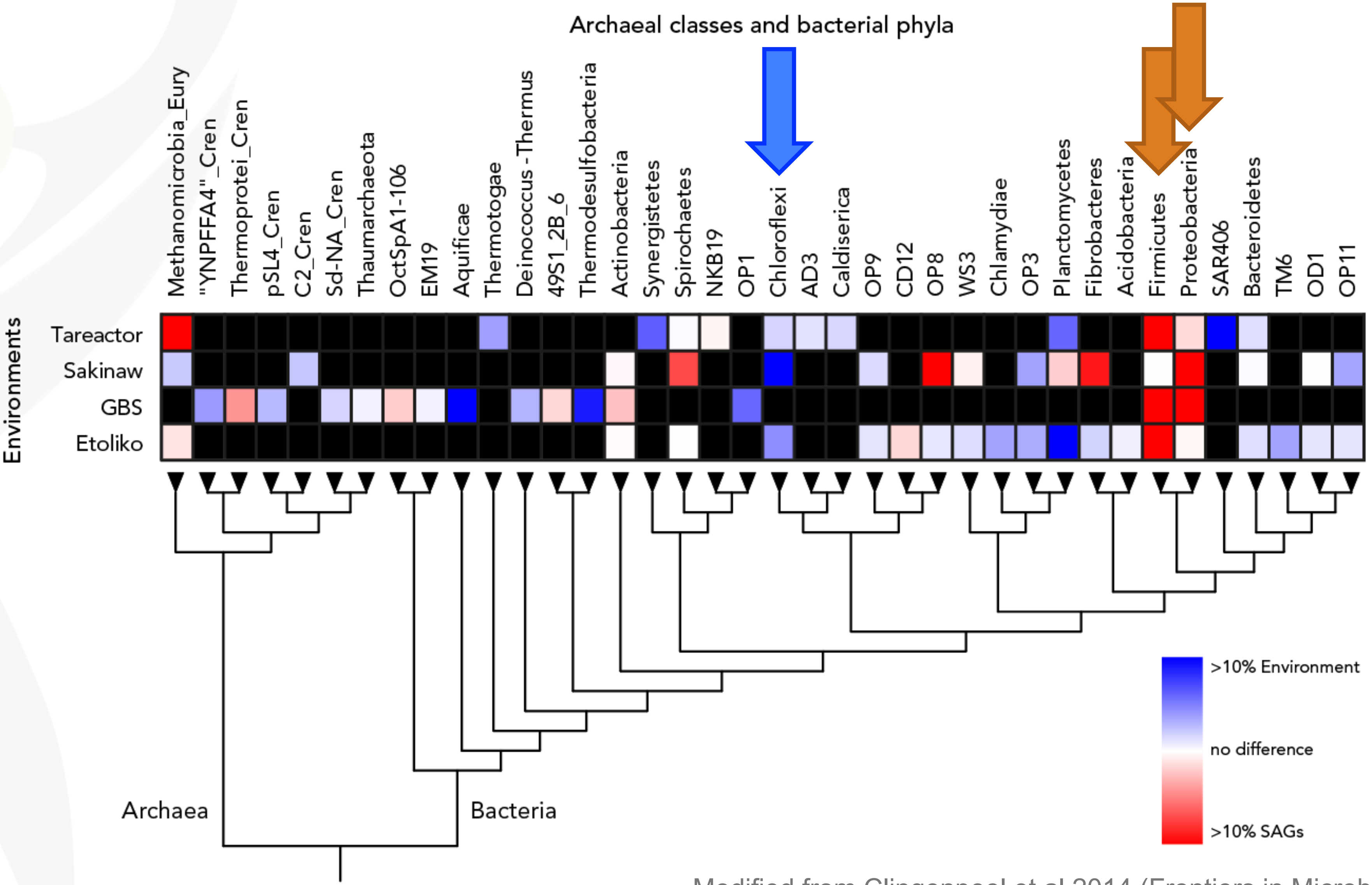
1-2 ul MDA reactions possible with Echo LH

Current limitations: 100 cells ≠ 100 SAGs

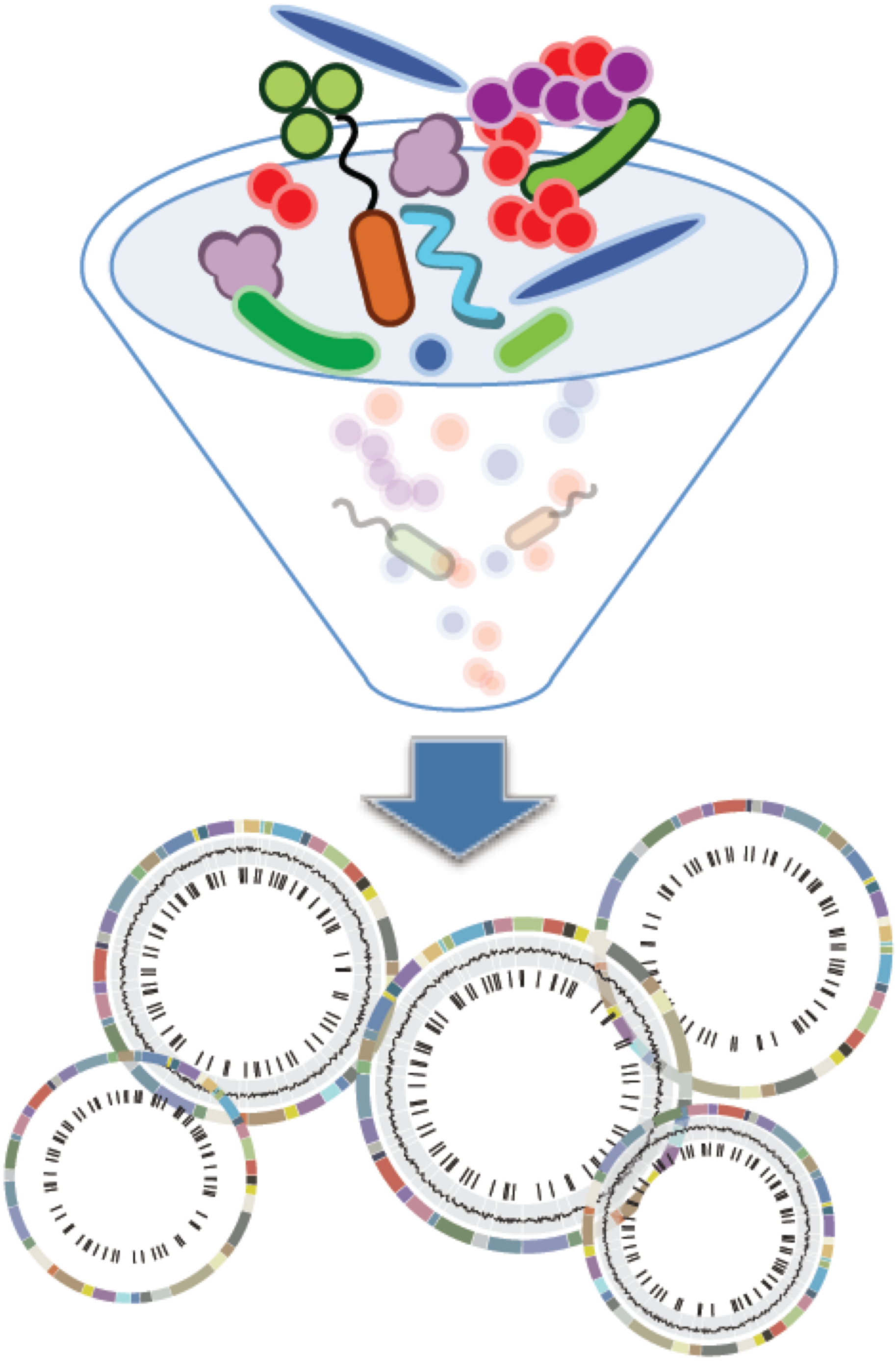


- ○ **Not every cell can be isolated**
- ○ **Not every cell can be lyzed and WGA'd**
- ○ **Not every cell can be 16S ID'd**

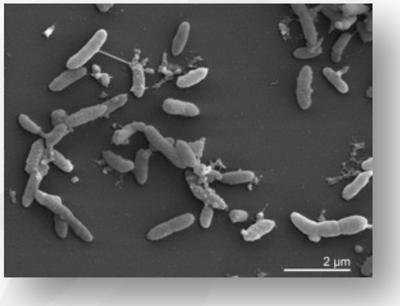
Recovered diversity: 16S tags vs SAGs



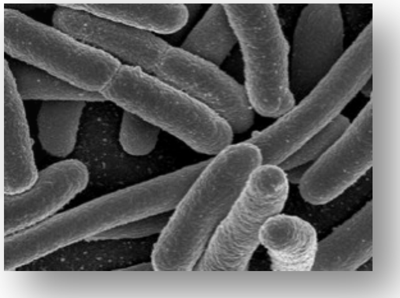
Modified from Clingenpeel et al 2014 (Frontiers in Microbiol)



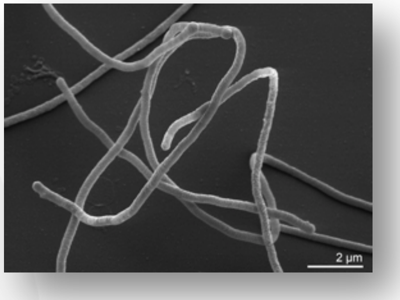
Benchmark experiment: what to expect?



P: *Pedobacter heparinus*
Size: 5,167,383 bp
GC-content: 42%



E: *Escherichia coli* str. K-12
Size: 4,639,675 bp
GC-content: 51%

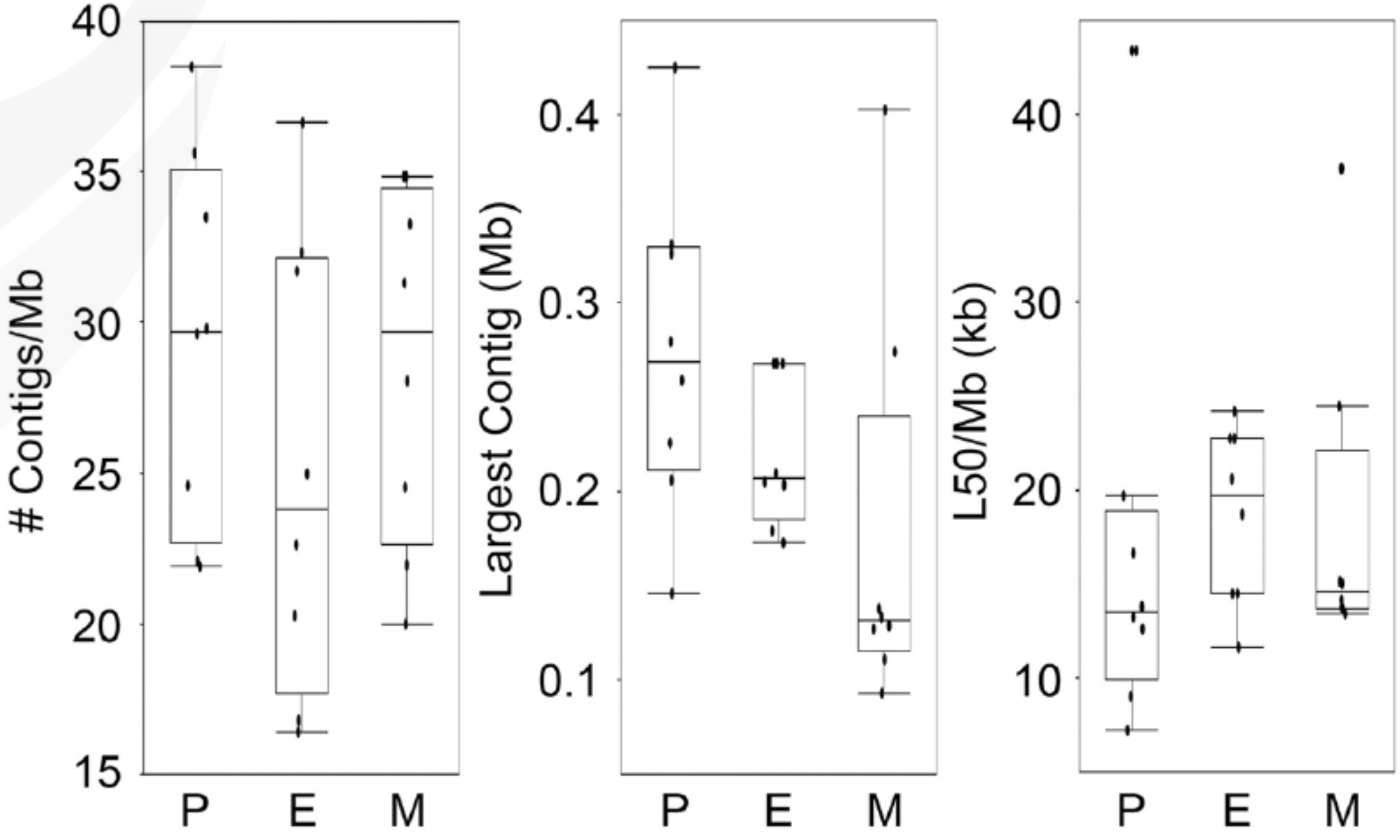


M: *Meiothermus ruber*
Size: 3,097,457 bp
GC-content: 63%

Sequenced 8 single-cell genomes/each

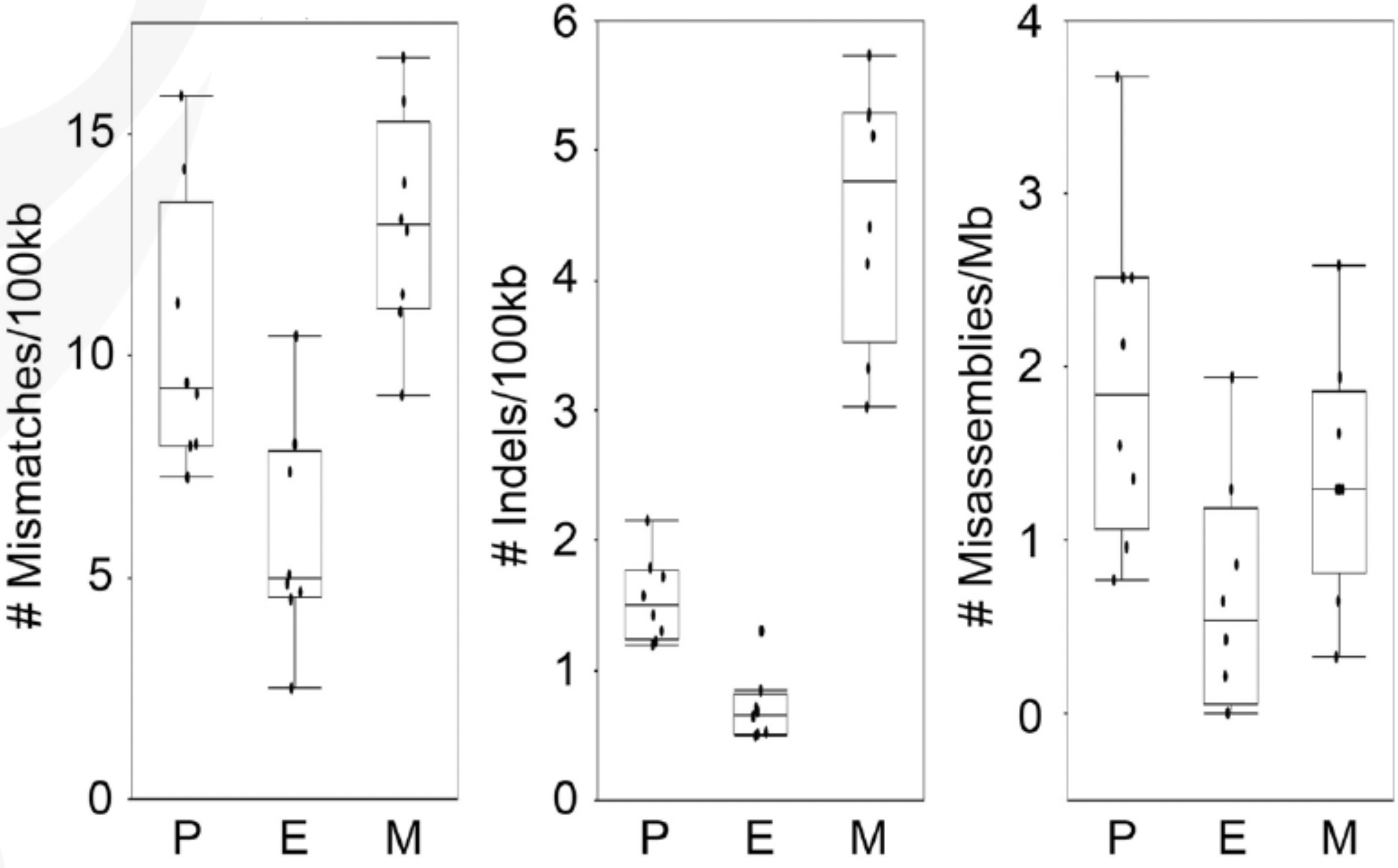
Image source: http://www.standardsingenomics.org

Assemblies are draft quality



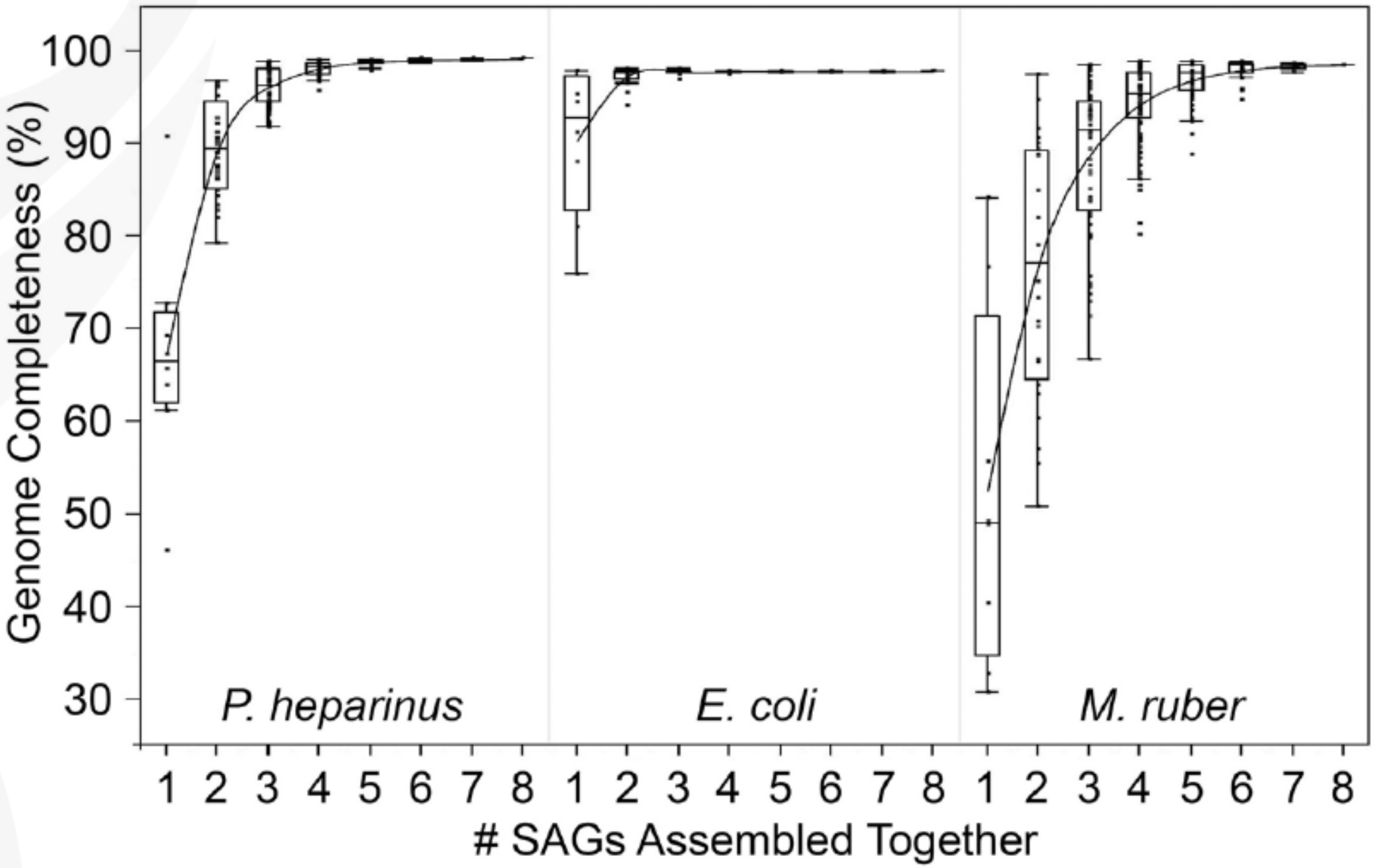
Clingenpeel et al 2014 (Frontiers in Microbiol)

With minimal errors



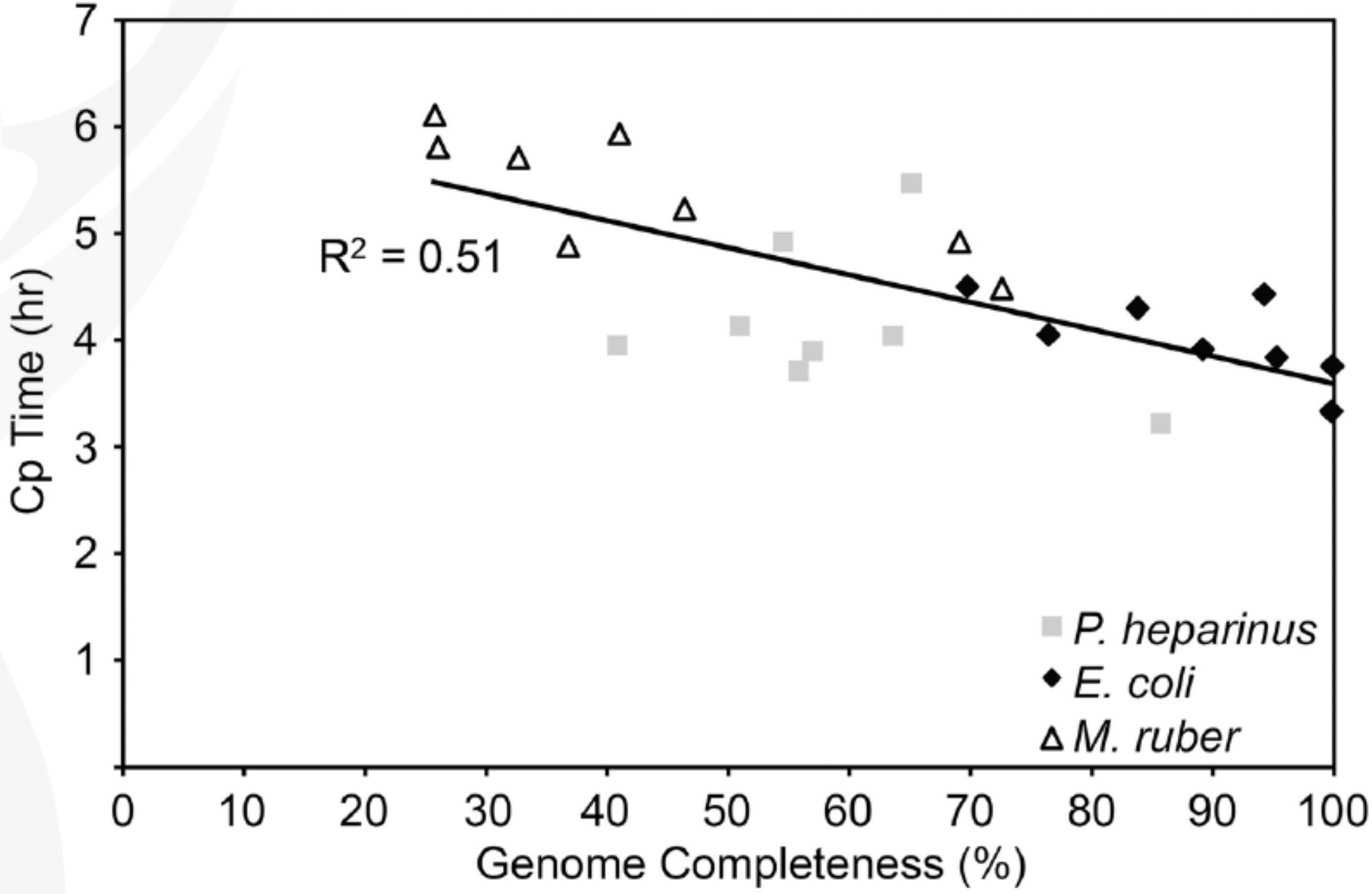
Clingenpeel et al 2014 (Frontiers in Microbiol)

SAG assemblies complement each other



Clingenpeel et al 2014 (Frontiers in Microbiol)

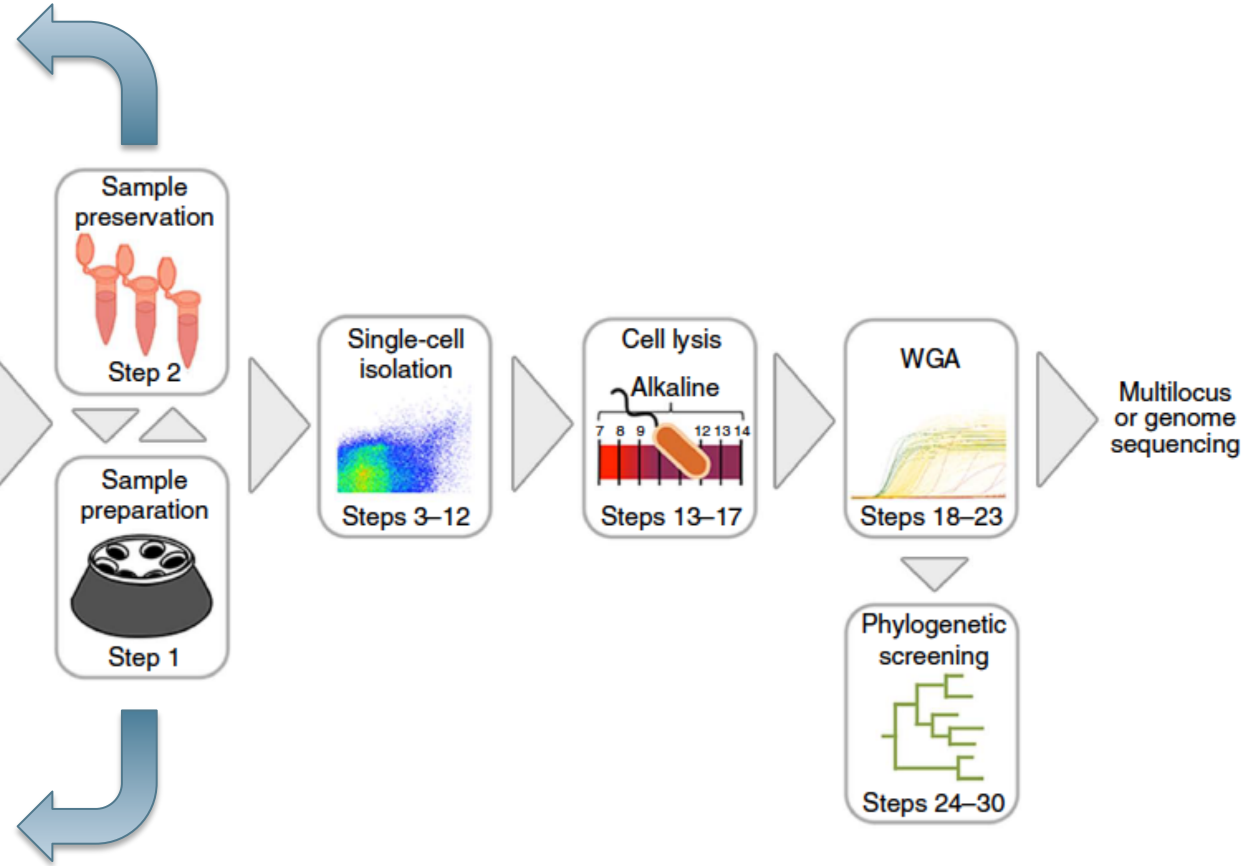
Low CP values lead to larger assemblies

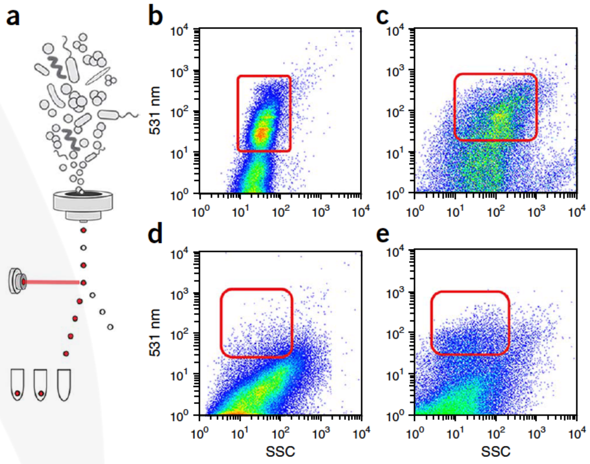


Clingenpeel et al 2014 (Frontiers in Microbiol)

A few words on sample preservation & preparation

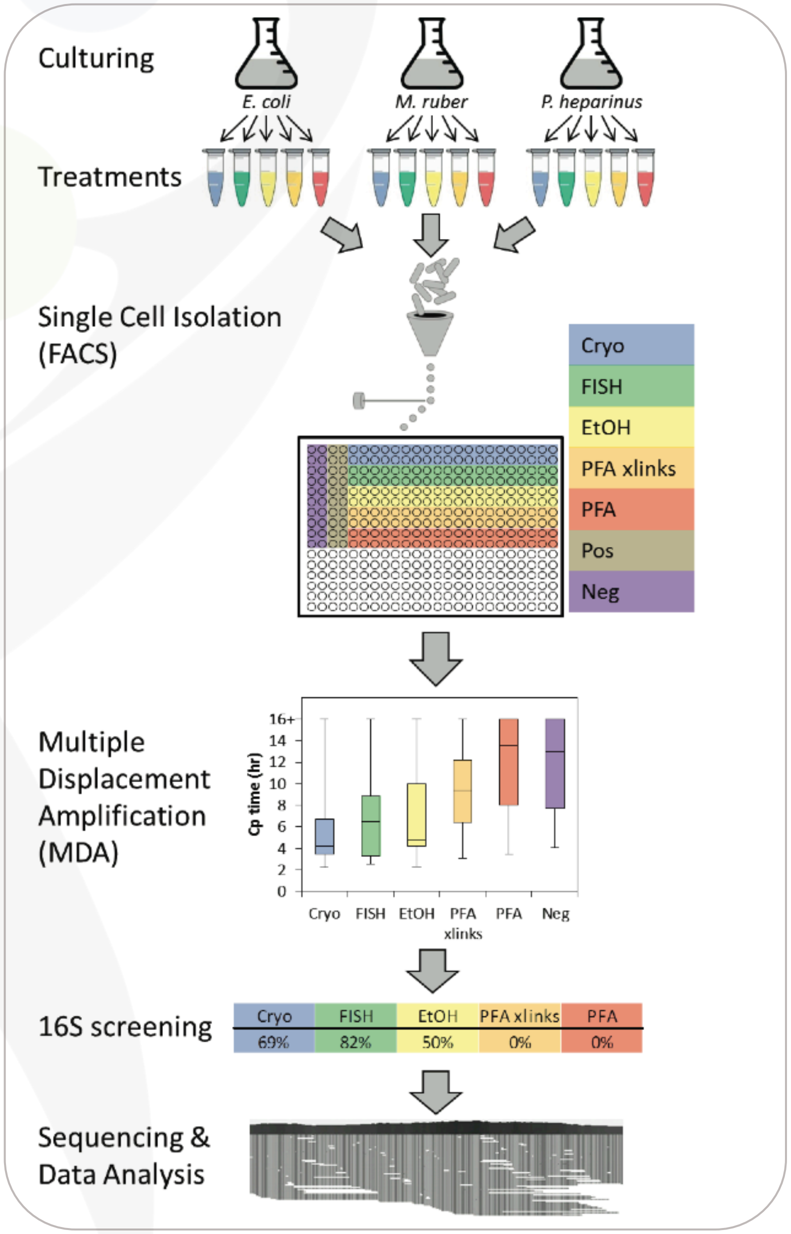
glycerol, betaine, DMSO, ..
(https://scgc.bigelow.org/PDFs/Sample_cryopreservation_glyTE.pdf)



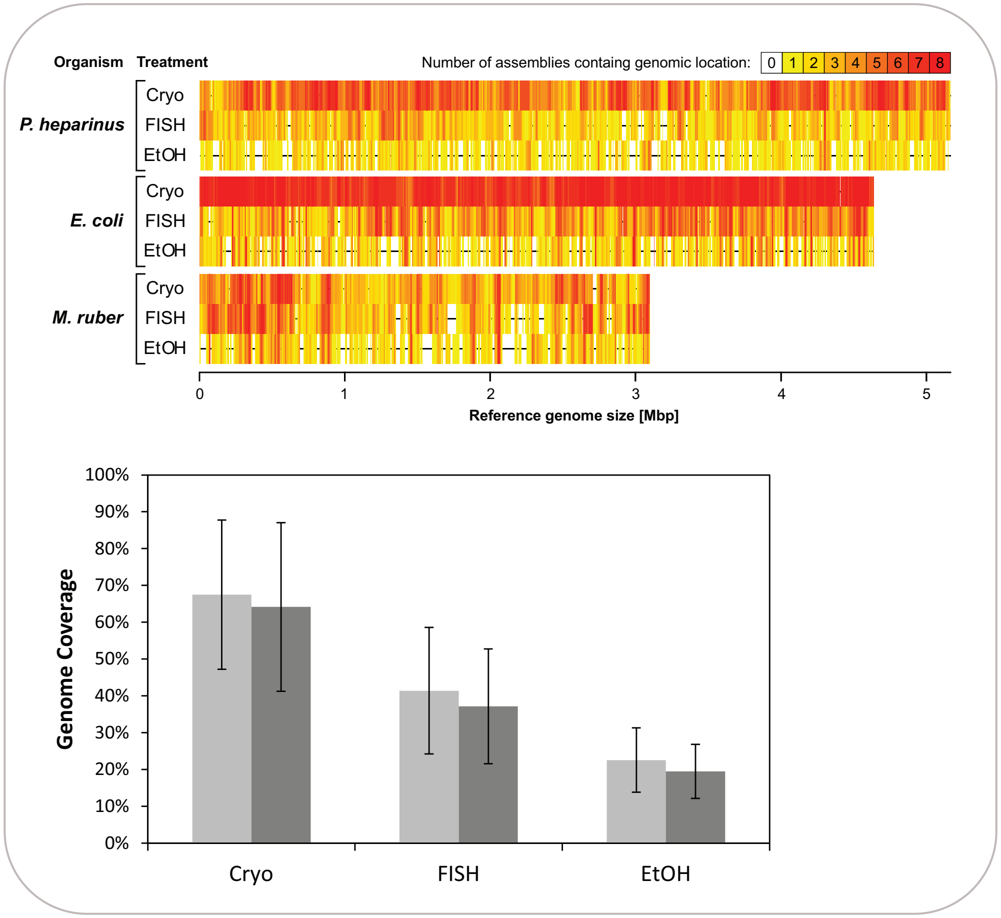


Rinke et al 2014 (Nature Protocols)

Does sample pre-treatment affect genome recovery?



Cryo	FISH	EtOH	PFA xlinks	PFA
69%	82%	50%	0%	0%



Clingenpeel et al 2014 (ISME J)

What to do with this single-cell data?

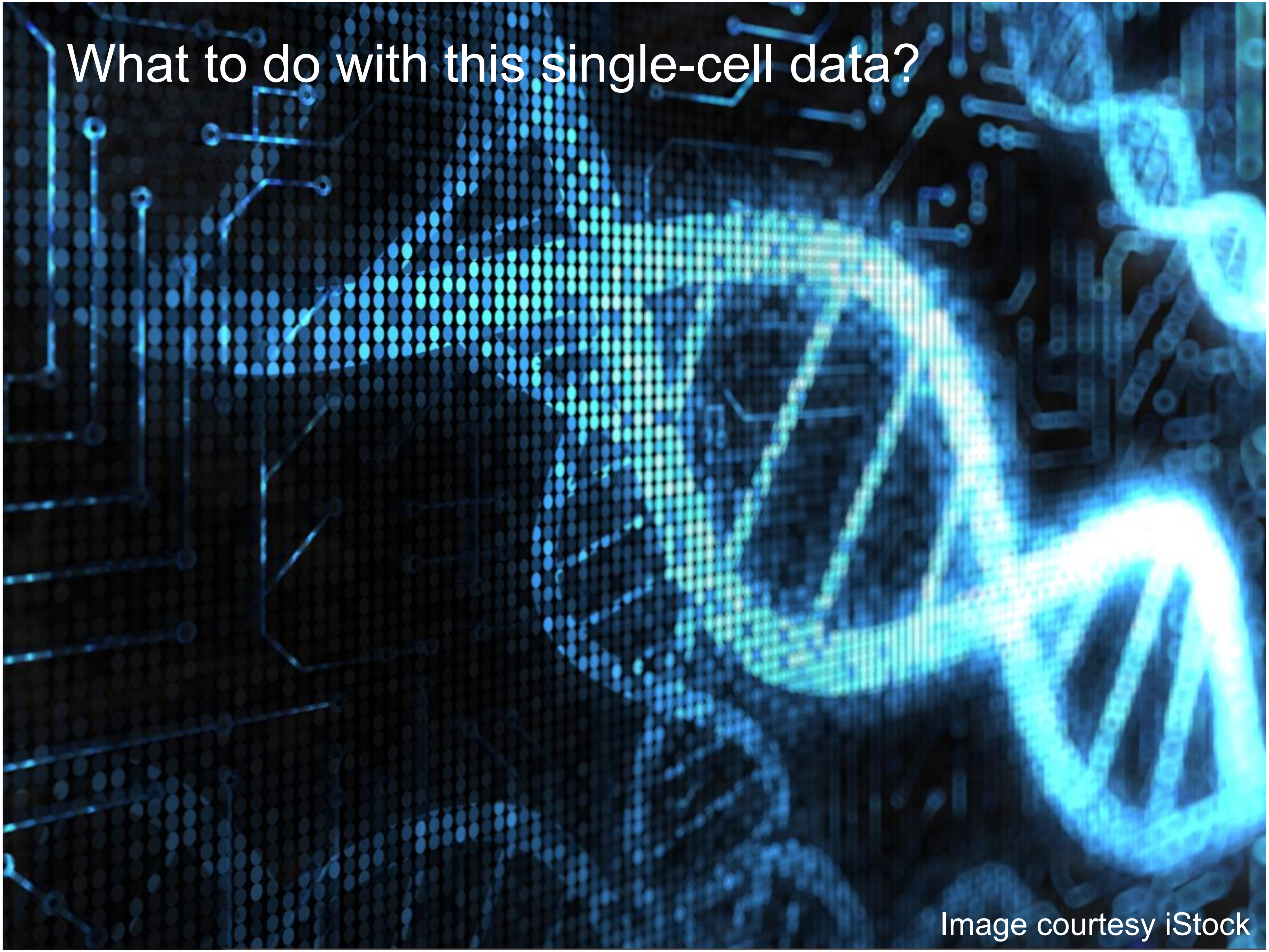
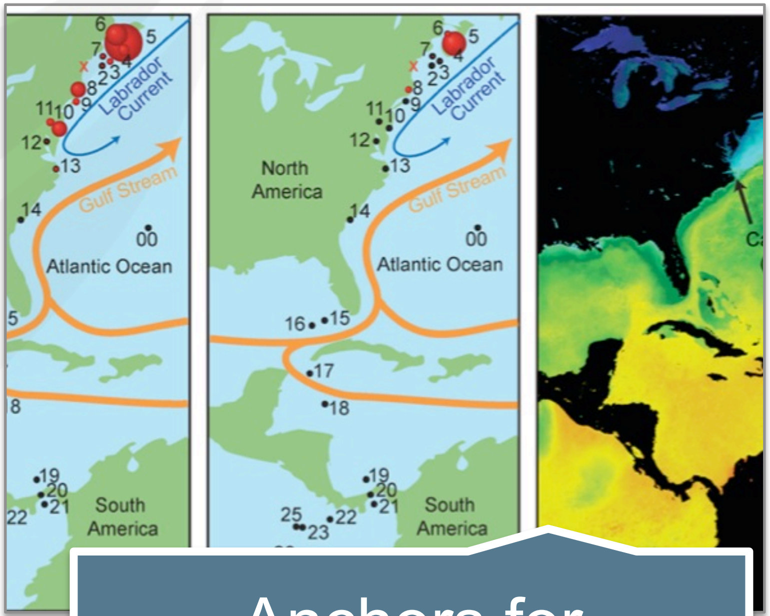
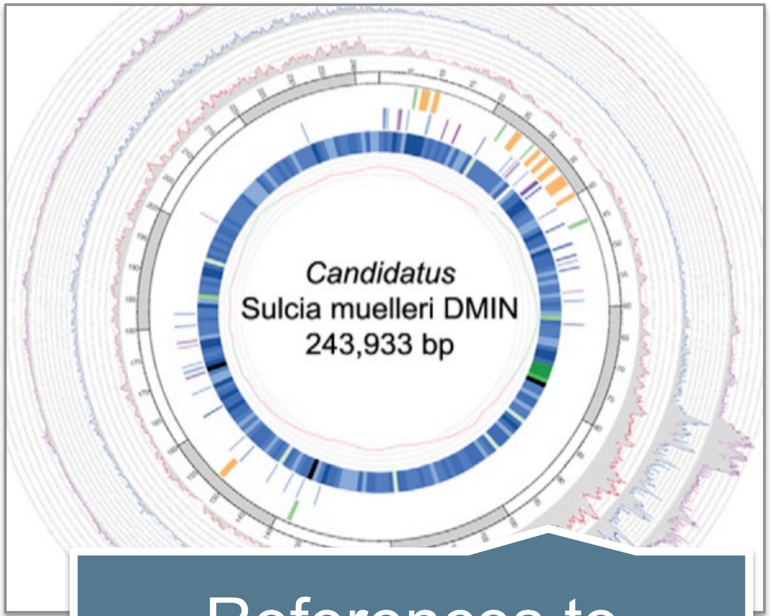


Image courtesy iStock

Utility of single-cell data is nearly endless..

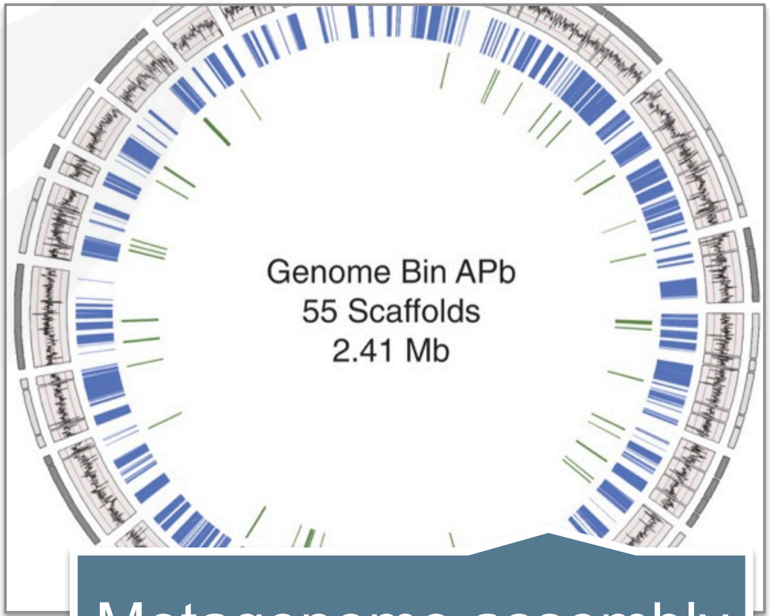


Anchors for biogeography studies

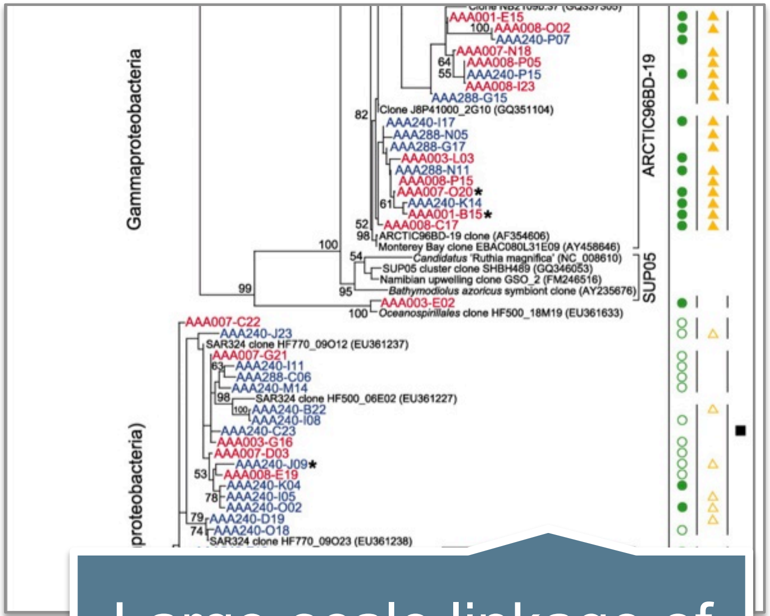


References to assess heterogeneity

Utility of single-cell data is nearly endless..

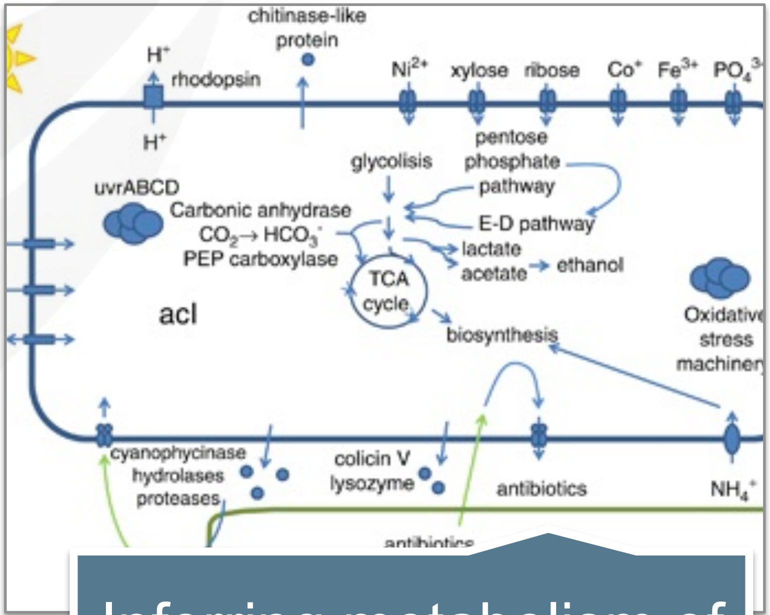


Metagenome assembly and binning validation

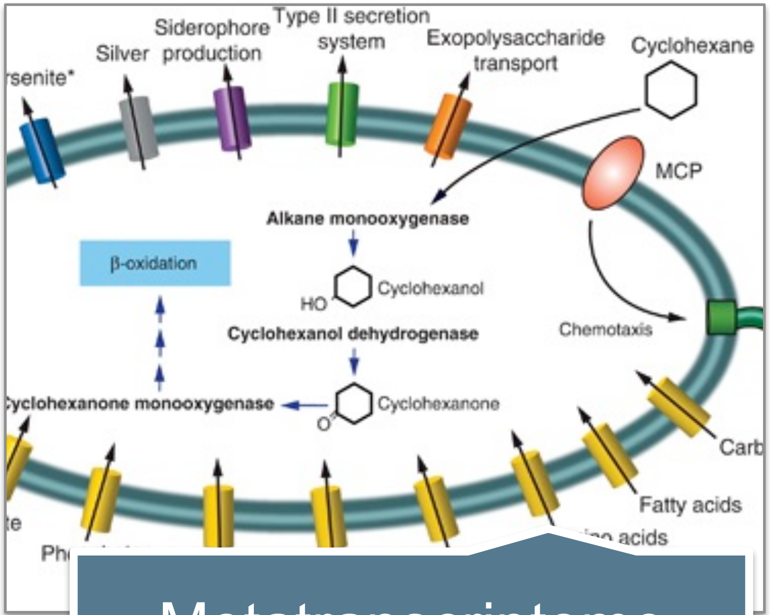


Large-scale linkage of phylogeny and function

Utility of single-cell data is nearly endless..



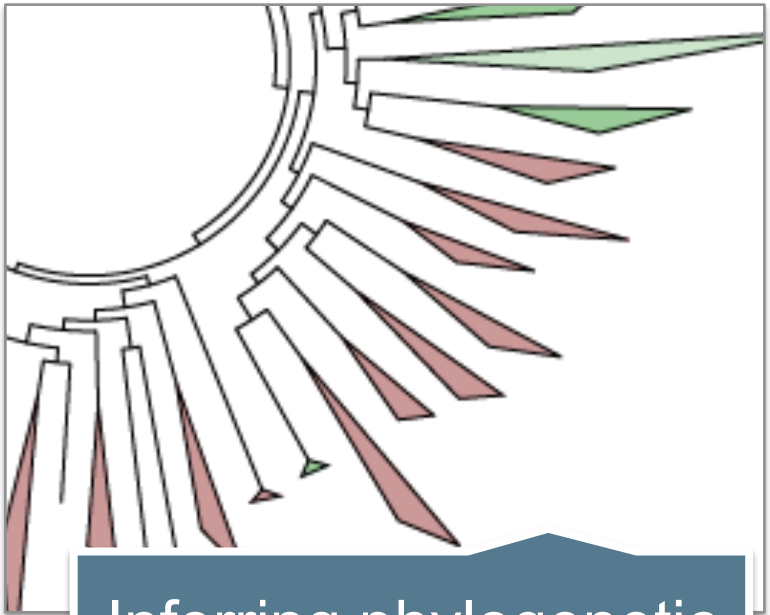
Inferring metabolism of uncultured microbes



Metatranscriptome data mapping

Utility of single-cell data is nearly endless..

Identifying novel genomic features



Inferring phylogenetic relationships

- **A few words about the JGI**
  - who we are & what we do
- **Single-cell genomics**
  - the why, the how & what to expect from it

- **Single-cell science vignettes**
  - from symbionts to microbial dark matter
- **Crystal ball**
  - what the future may bring

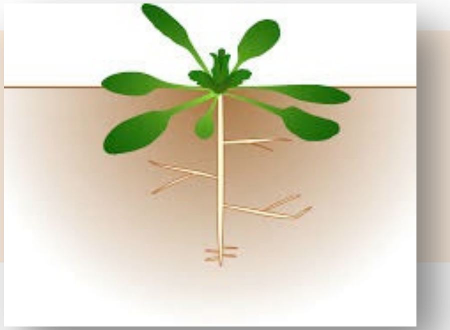
Marine Flavobacteria

R. Stepanauskas
(Bigelow Lab)

Sharpshooter symbiont

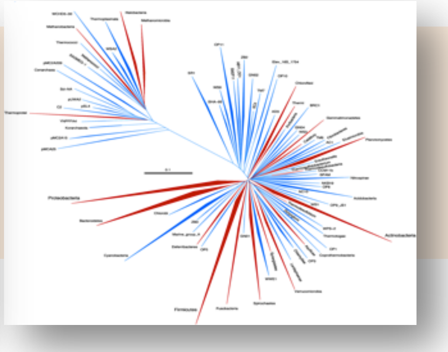
N. Moran
(U of Arizona)

***Arabidopsis* endophytes**



J. Dangl
(UNC)

Microbial Dark Matter

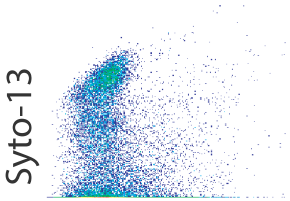


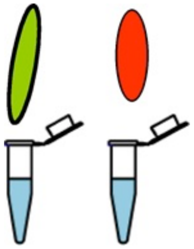
MDM
Community

Two draft single-cell genomes

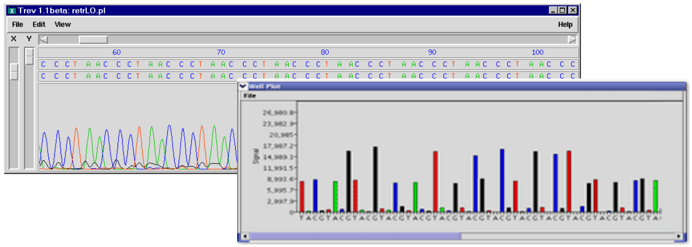
Marine Flavobacteria sp.

Ramunas Stepanauskas
(Bigelow Lab)

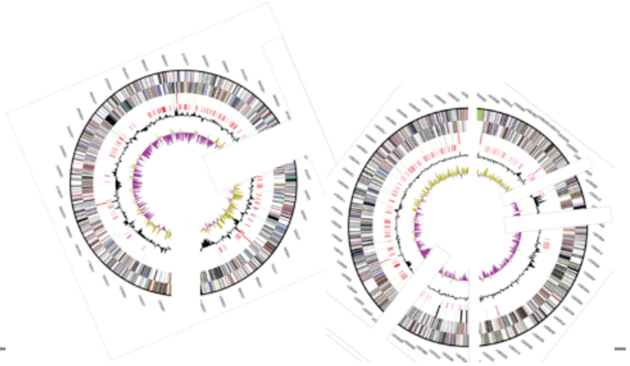




single cells



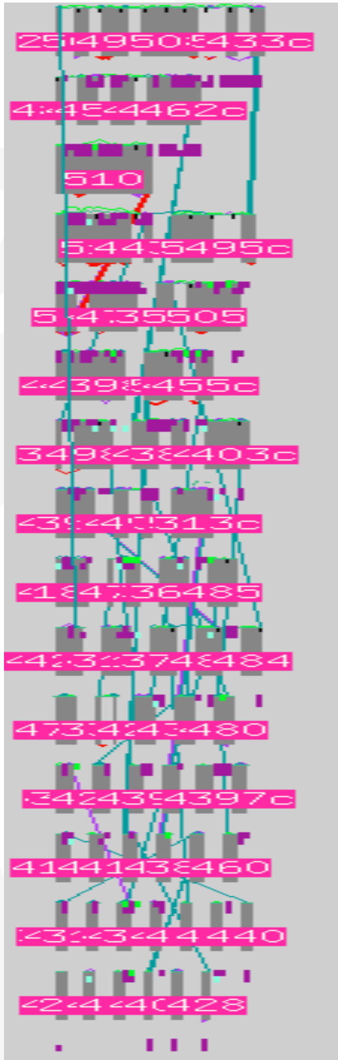
shotgun
sequencing



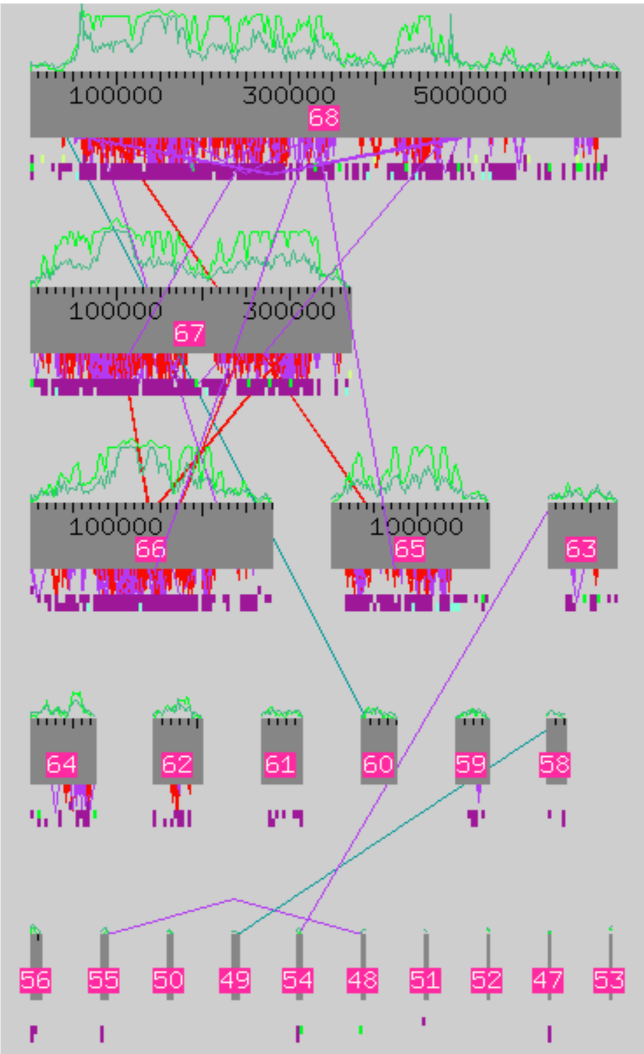
Assembly statistic	MS024-2A	MS024-3C
Assembly size [Mbp]	1.905	1.515
Estimated genome size [Mbp]	2.095	1.947
Estimated genome recovery [%]	91	78
Number of contigs	17	21
Largest contig [kbp]	684	549

Woyke et al, PLoS One (2009)

Extensive manual gap closures

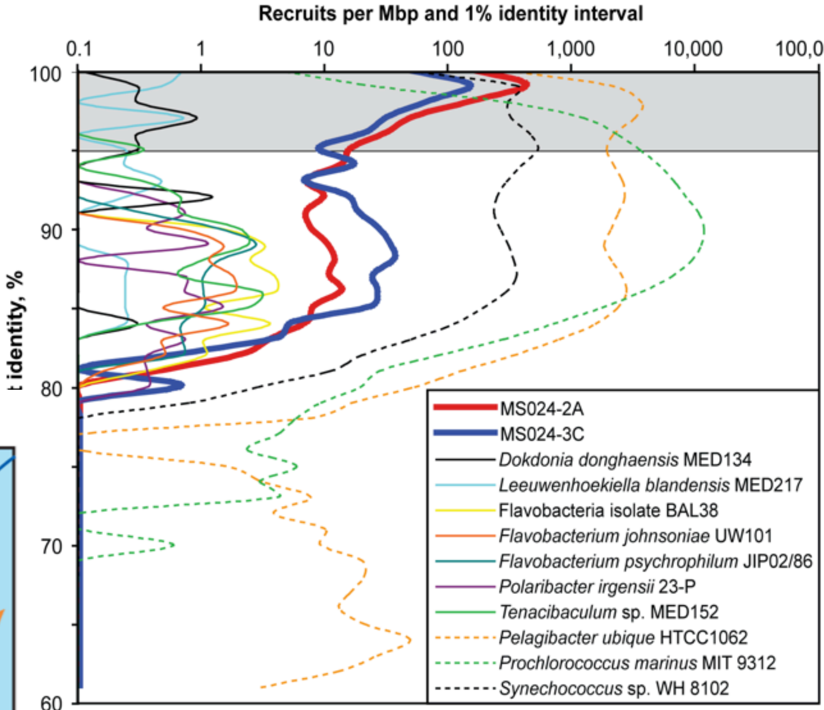


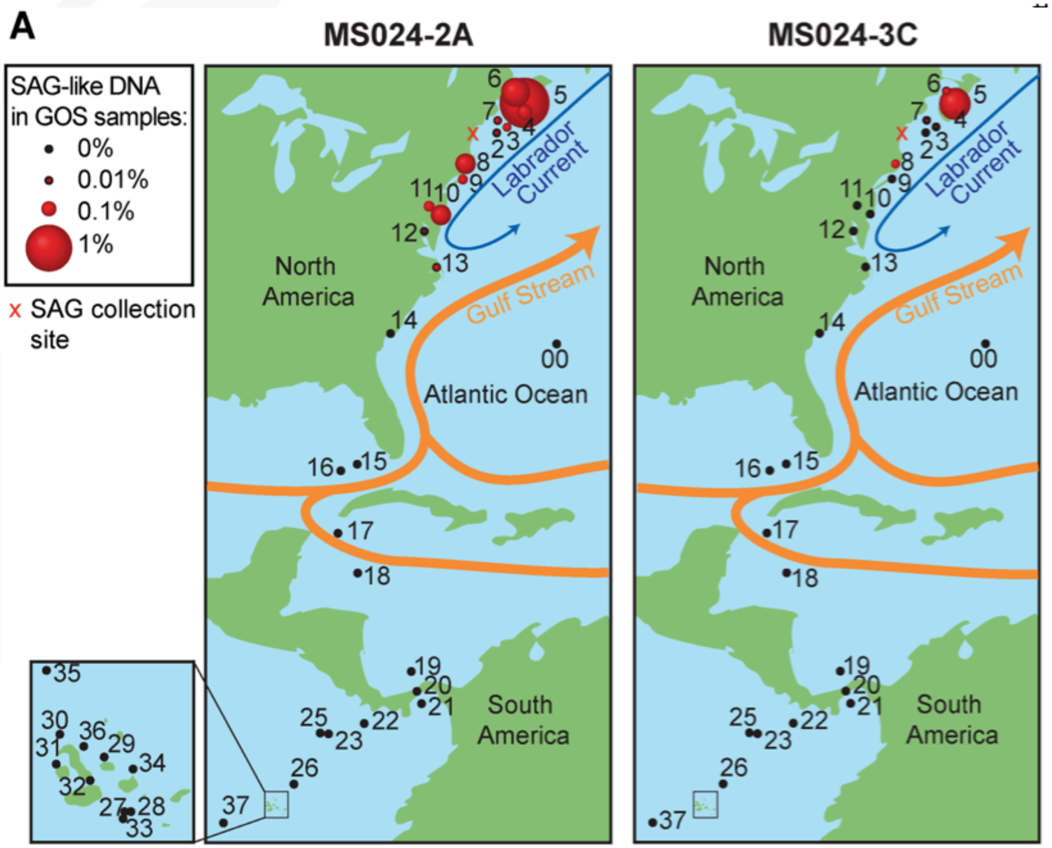
Ace #3
1.44 Mbp
47 contigs
Longest 56 kb



Ace #65
1.9 Mbp
17 contigs
Longest 680 kb

SAGs are representative of their environment





Woyke et al, PLoS One (2009)

Marine Flavobacteria

R. Stepanauskas
(Bigelow Lab)

Sharpshooter symbiont

N. Moran
(U of Arizona)

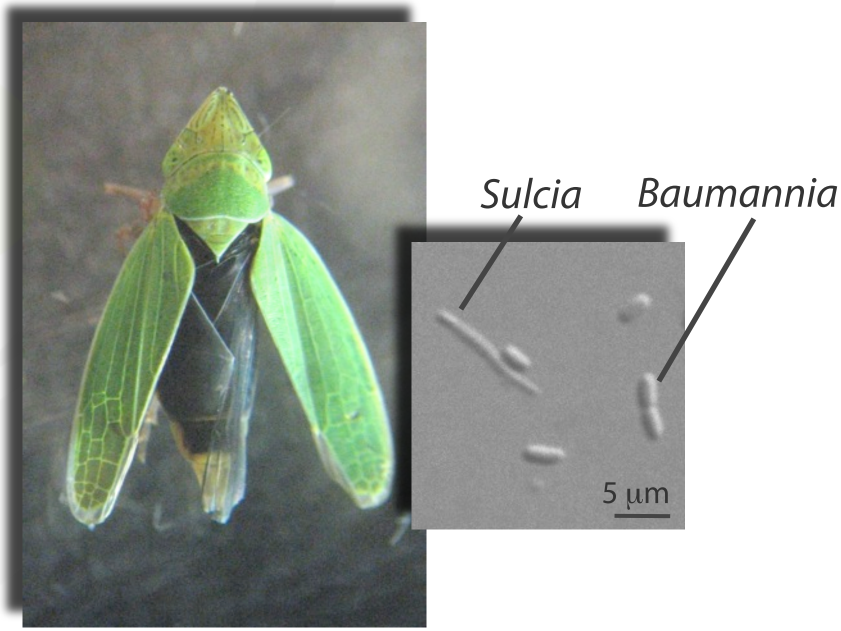
***Arabidopsis* endophytes**

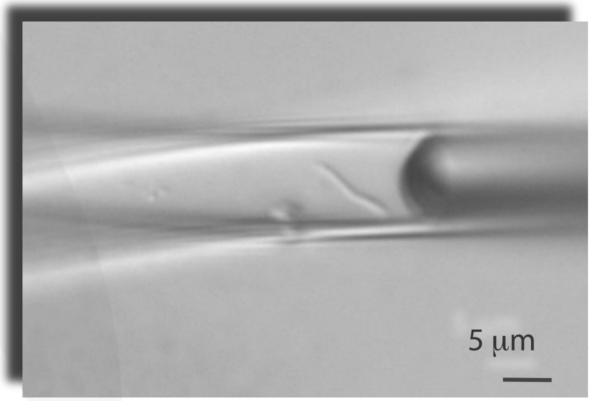
J. Dangl
(UNC)

Microbial Dark Matter

MDM
Community

Green sharpshooter symbionts





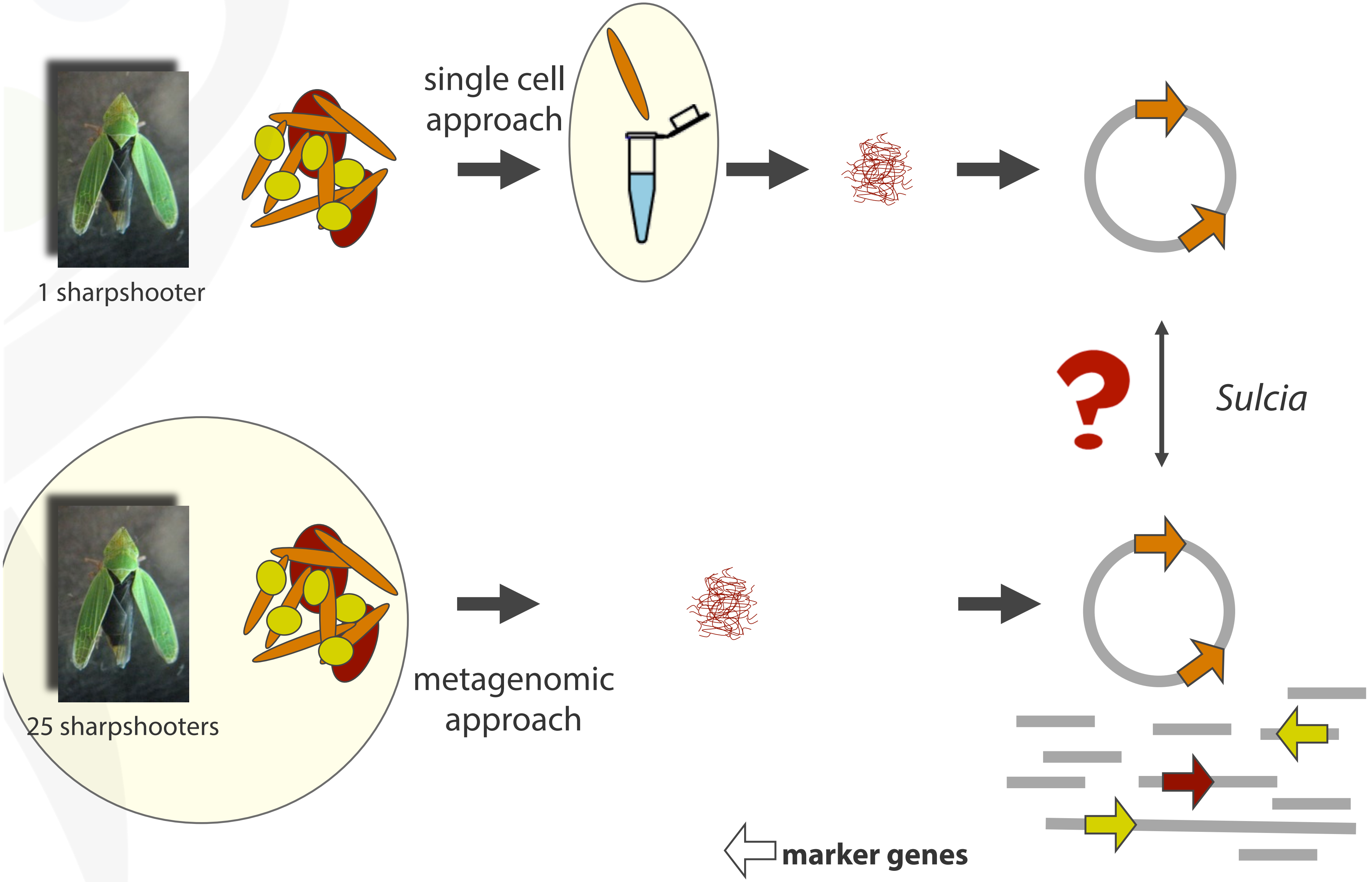
Green sharpshooter (*D. minerva*)

- sap-feeding leafhoppers
- harbors two symbionts
- involved in plant pathogen spread (*Xylella*)
- causes serious plant diseases

Sulcia as ideal candidate for single-cell finishing

- uncultured ✓
- small genome size ✓
- simple single cell isolation ✓
- fresh sample availability ✓
- polyploid ✓

Single cell & metagenome approach



1 sharpshooter

single cell approach

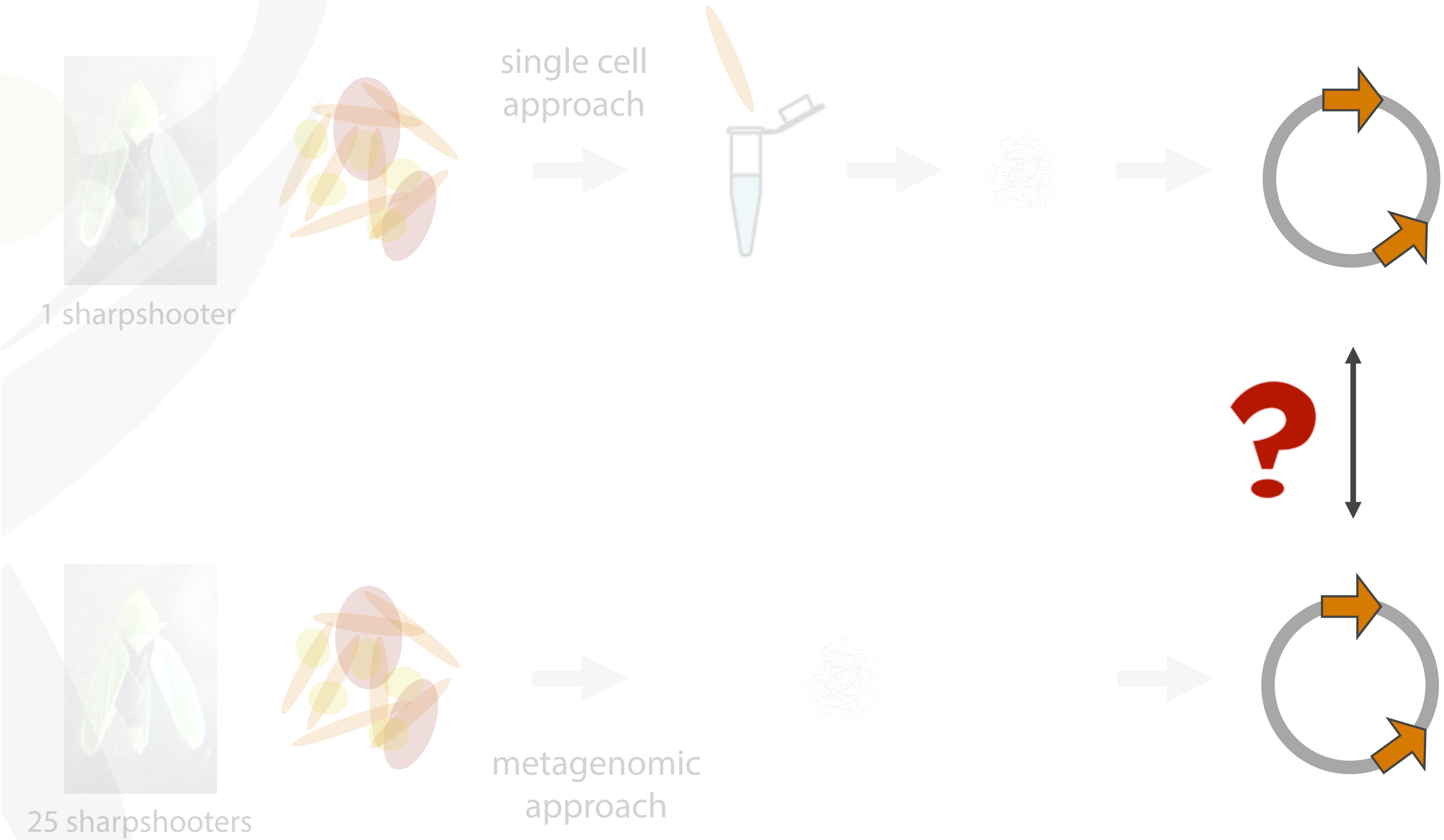
Sulcia

25 sharpshooters

metagenomic approach

marker genes

Are the genomes identical?



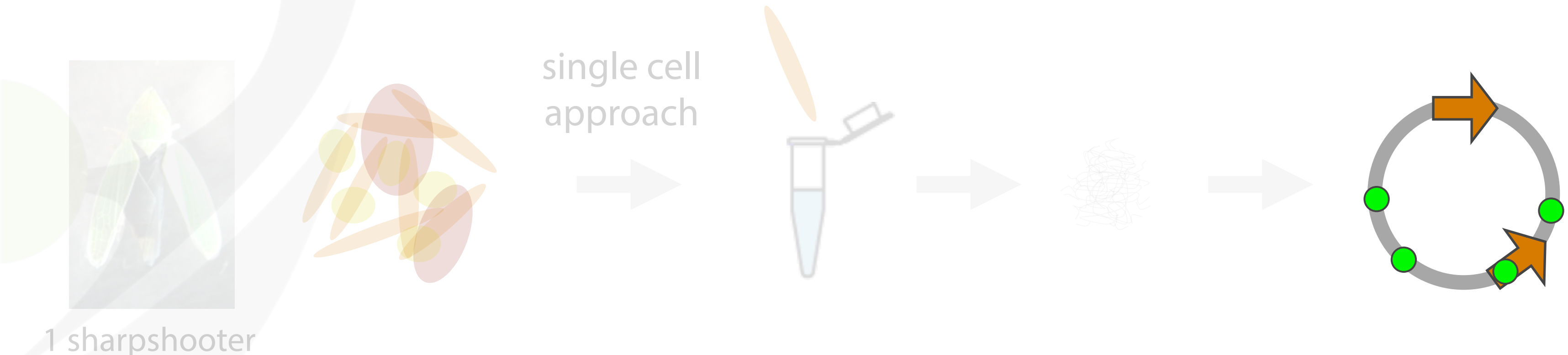
single cell approach

1 sharpshooter

metagenomic approach

25 sharpshooters

Are the genomes identical?



- ● TGAA[ttatca]$_n$TTAA
- ● CTAA[tgaagttaa]$_n$CCAT
- ● TAAA[agaaattgagaagttgcgaataataaa]$_n$TTTA
- ● CTTG[ctactta]$_n$ATAA

One complete single-cell genome

Sequence data

(Sanger/454)

↓

Assembly

↓

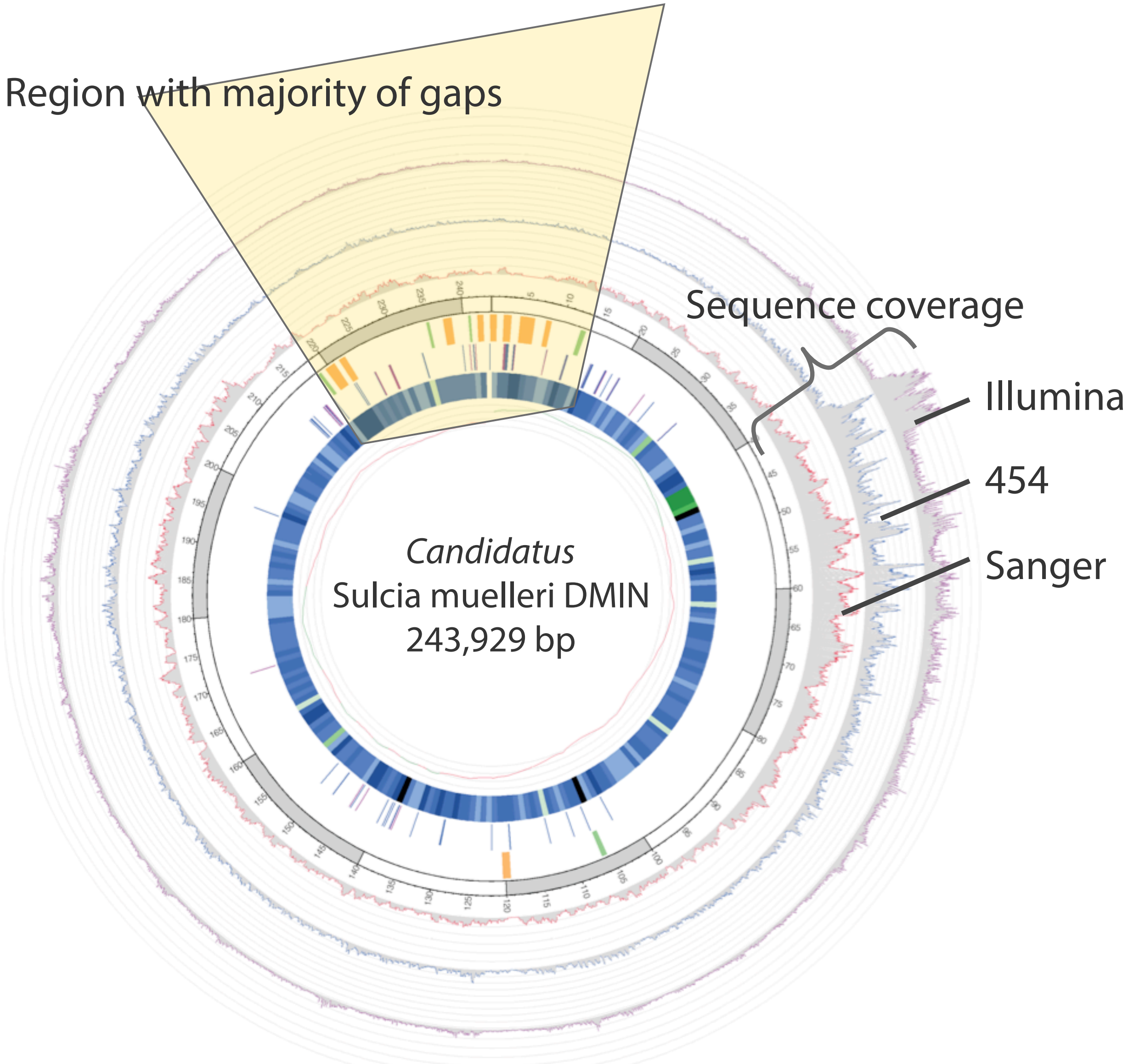
31 contigs

↓

Gap closure / polishing

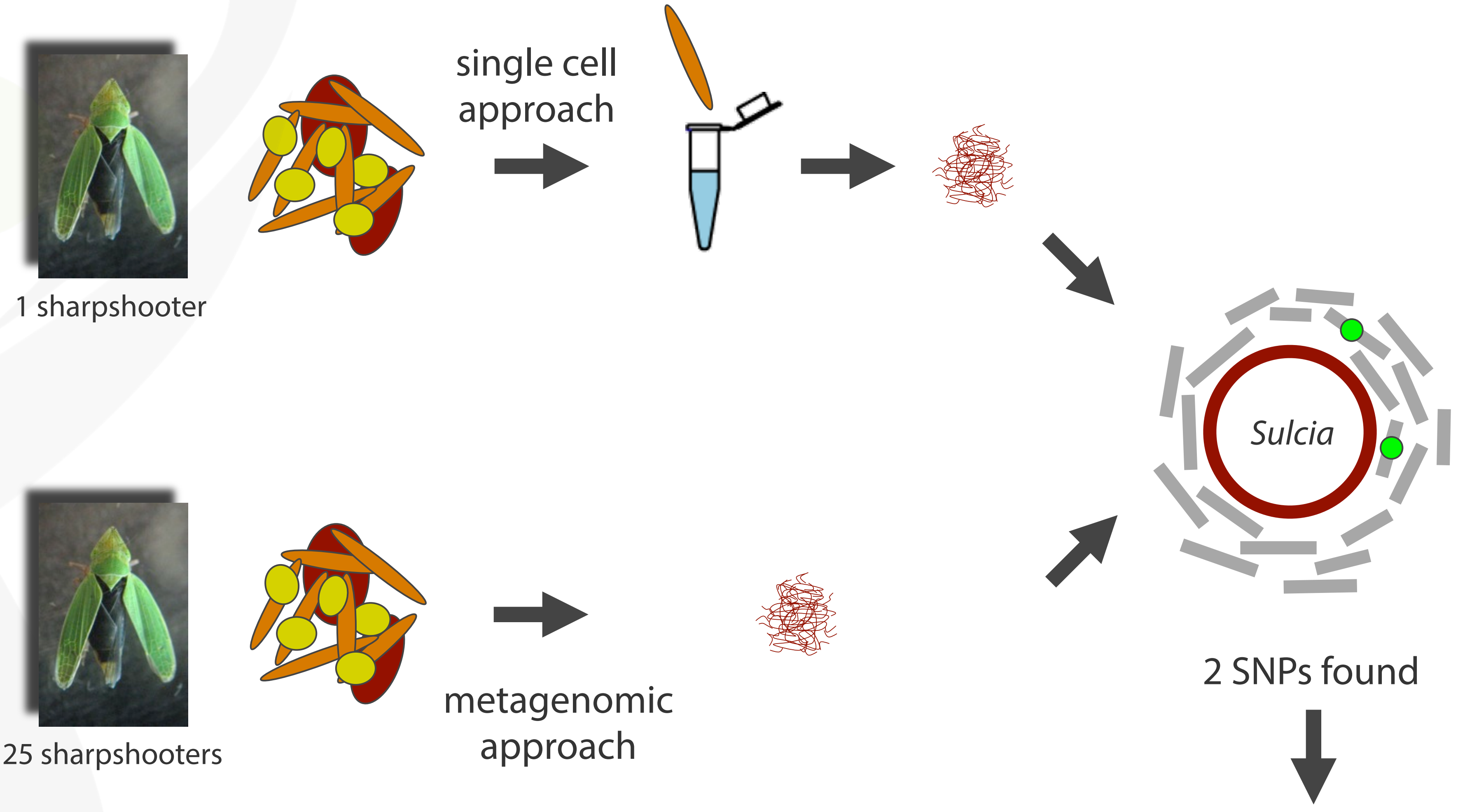
(Sanger/ Illumina)

Genome Properties	
Genome Size [bp]	243,929
DNA coding bases [bp]	236,976
GC contents [%]	22
Total genes	261



Woyke et al, PLoS One (2010)

Evaluation of population variation



1 sharpshooter

single cell approach

25 sharpshooters

metagenomic approach

Sulcia

2 SNPs found

low genetic diversity within

Sulcia population

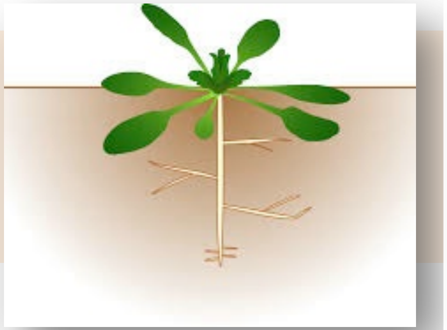
Marine Flavobacteria

R. Stepanauskas
(Bigelow Lab)

Sharpshooter symbiont

N. Moran
(U of Arizona)

***Arabidopsis* endophytes**

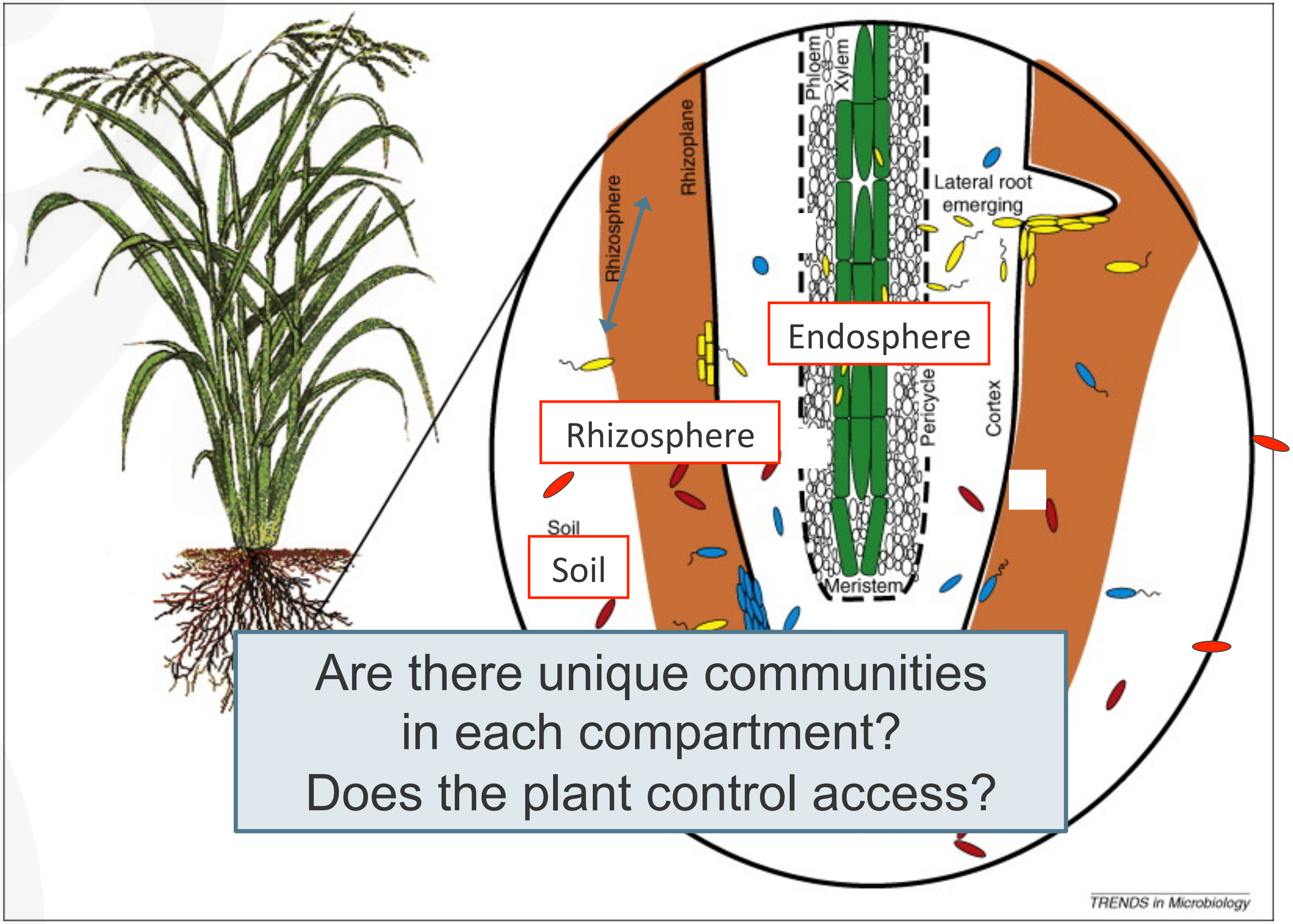


J. Dangl
(UNC)

Microbial Dark Matter

MDM
Community

Arabidopsis microbiome project



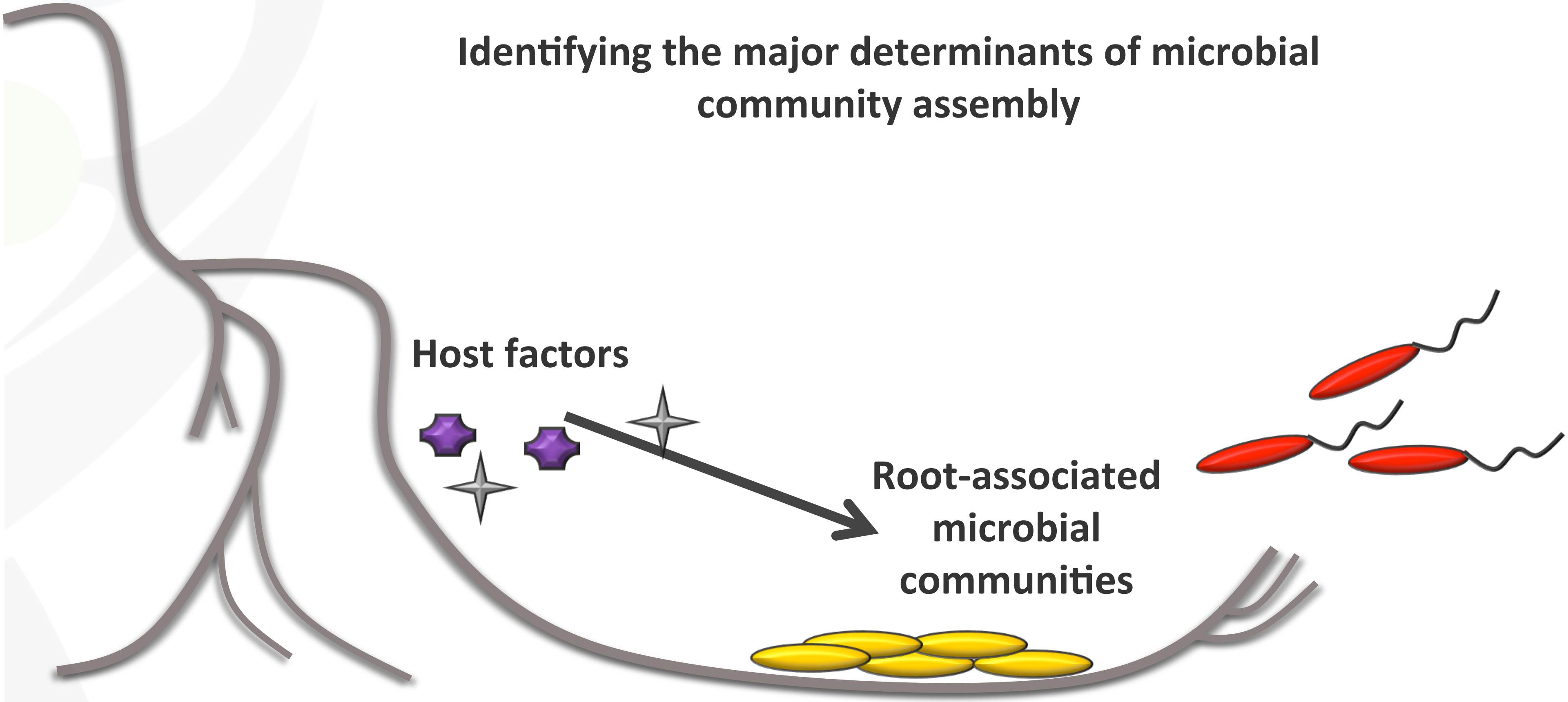
TRENDS in Microbiology

Are there unique communities in each compartment?
Does the plant control access?

S. Tringe

Arabidopsis microbiome project

Identifying the major determinants of microbial community assembly

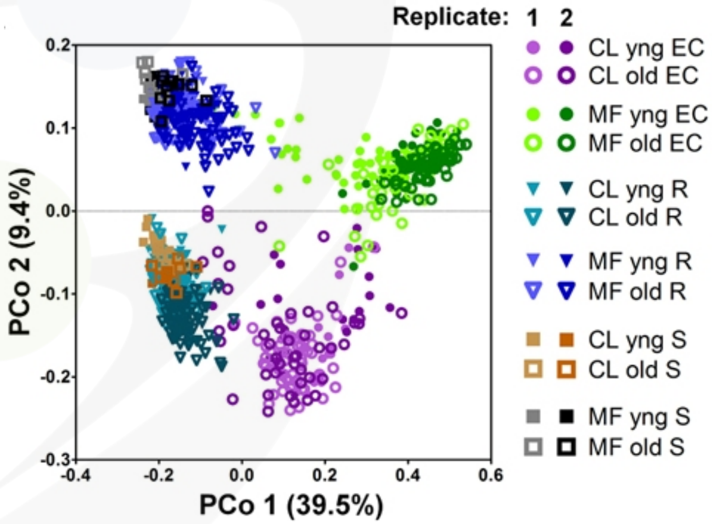


Full factorial design
1117 samples
16S pyrotag profiles

Variables investigating:
Soil type – Mason Farm vs. Clayton
Sample fraction – Bulk soil vs. rhizosphere vs. endophyte
Plant age – bolting (young) vs. senescent (old)
Genotype – 8 ecotypes
Individual – Aim for 10 individuals per condition

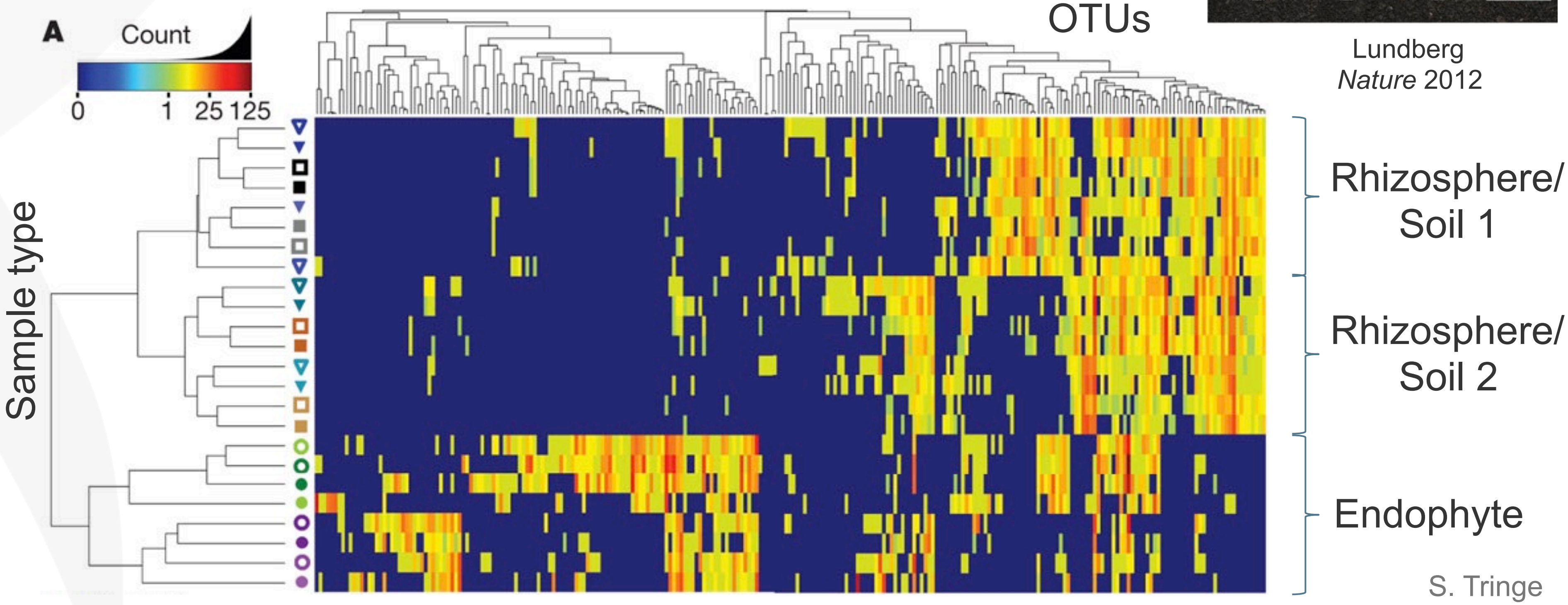
J. Dangl

The Arabidopsis microbiome



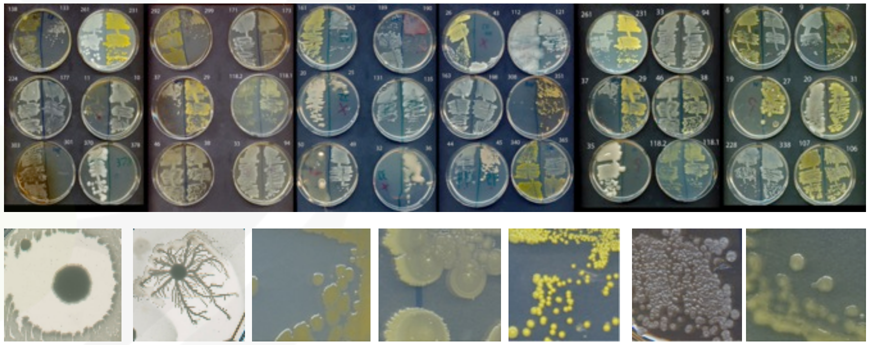
The endophyte community is unique and reproducible and similar across soil types

Lundberg
Nature 2012

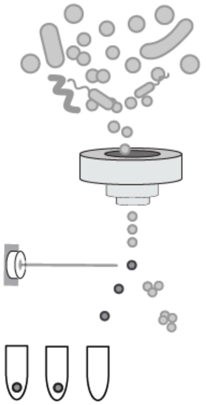


S. Tringe

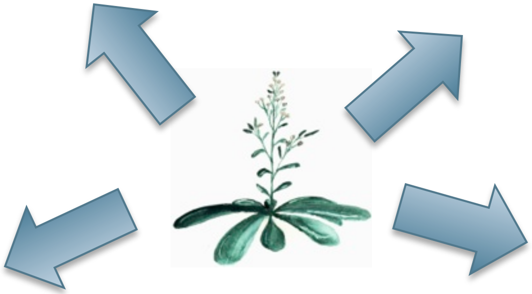
Isolates, single cells & enriched metagenomes



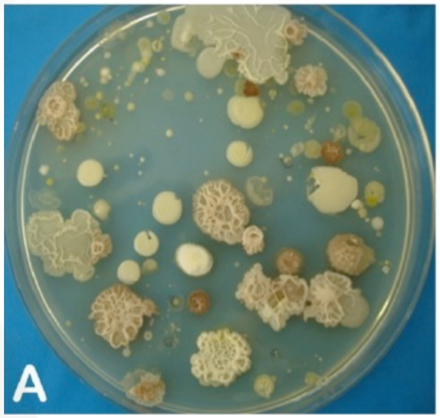
Cultured isolates



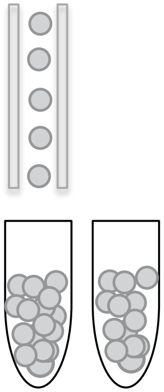
Single cells



“Plate scrape” metagenomes



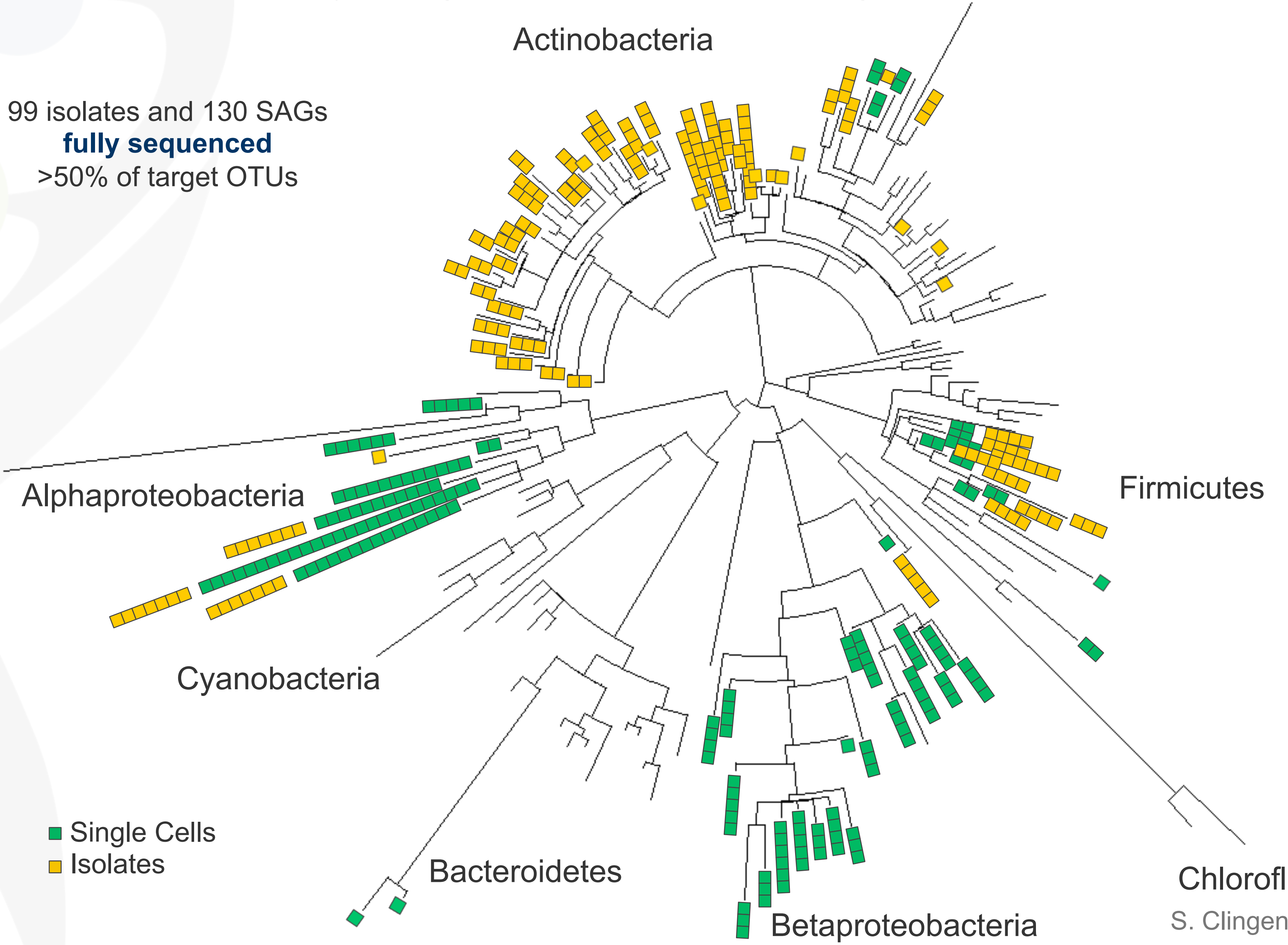
Flow-sorted “mini-metagenomes”



S. Tringe

An endophyte genome catalog

99 isolates and 130 SAGs
fully sequenced
>50% of target OTUs



S. Clingenpeel

Marine Flavobacteria

R. Stepanauskas
(Bigelow Lab)

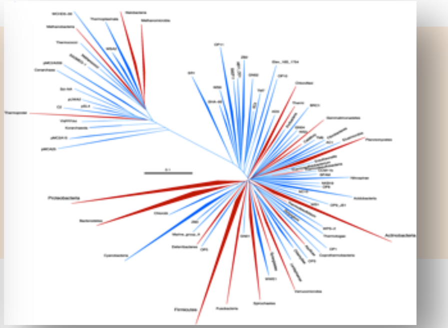
Sharpshooter symbiont

N. Moran
(U of Arizona)

***Arabidopsis* endophytes**

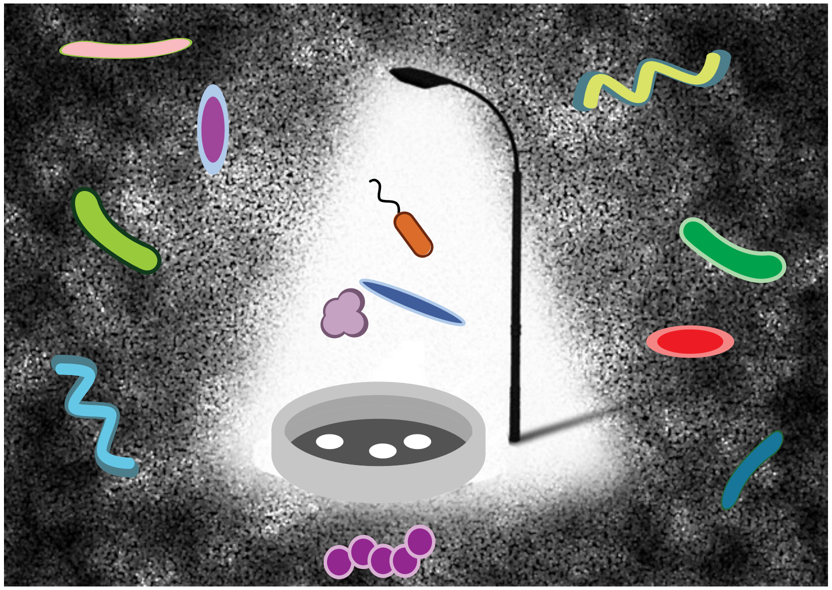
J. Dangl
(UNC)

Microbial Dark Matter



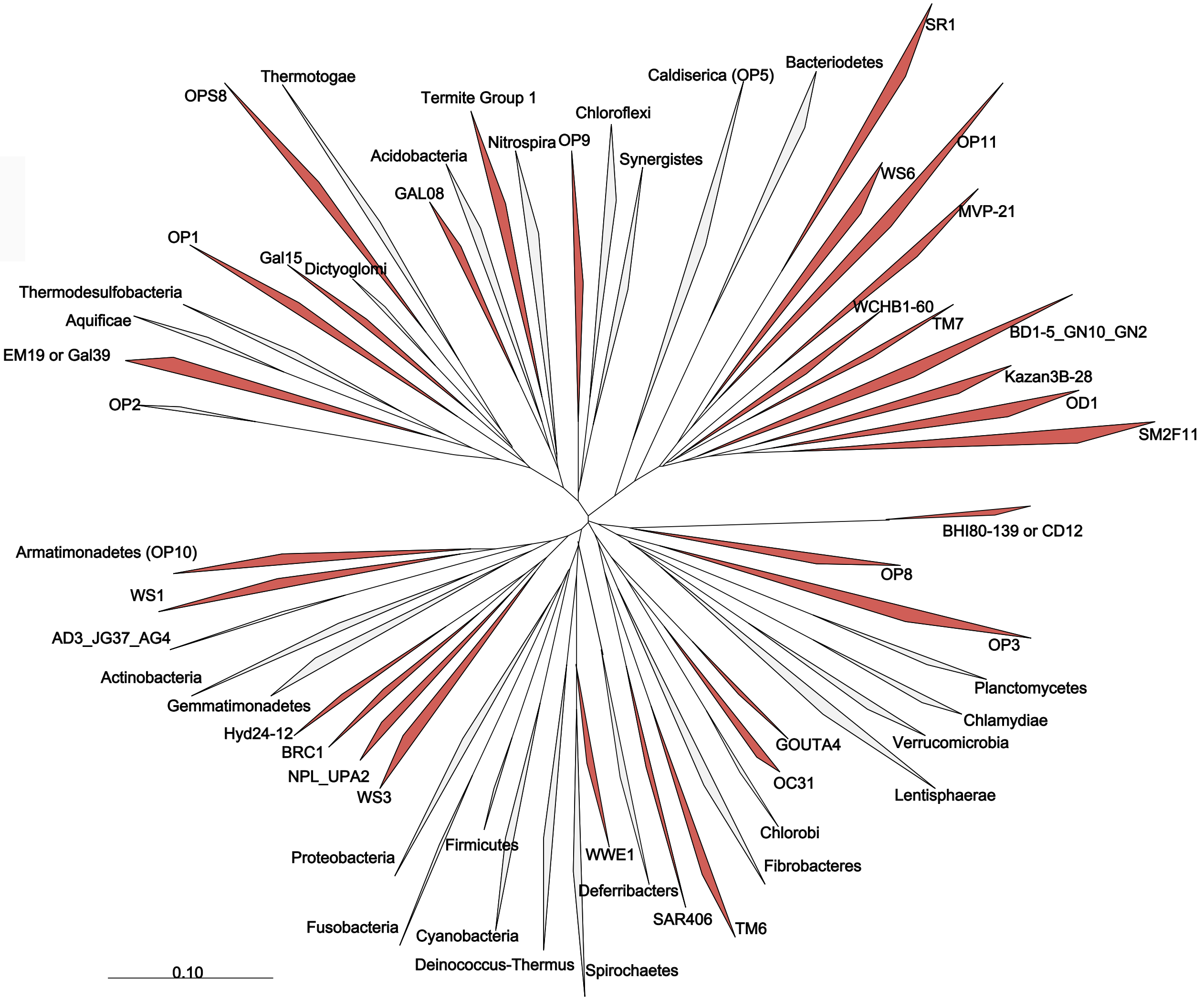
**MDM
Community**

Expansion of phylogenetic diversity



Using single-cell genomics to look outside the lamp post

Expansion of phylogenetic diversity



16S rRNA tree of known bacterial phyla

Phillip Hugenholz

Nikos Kyrpides

Natalia Ivanova

Jonathan Eisen

Ramunas Stepanauskas

Steven Hallam

Stefan Sievert

Wen-Tso Liu

George Tsiamis

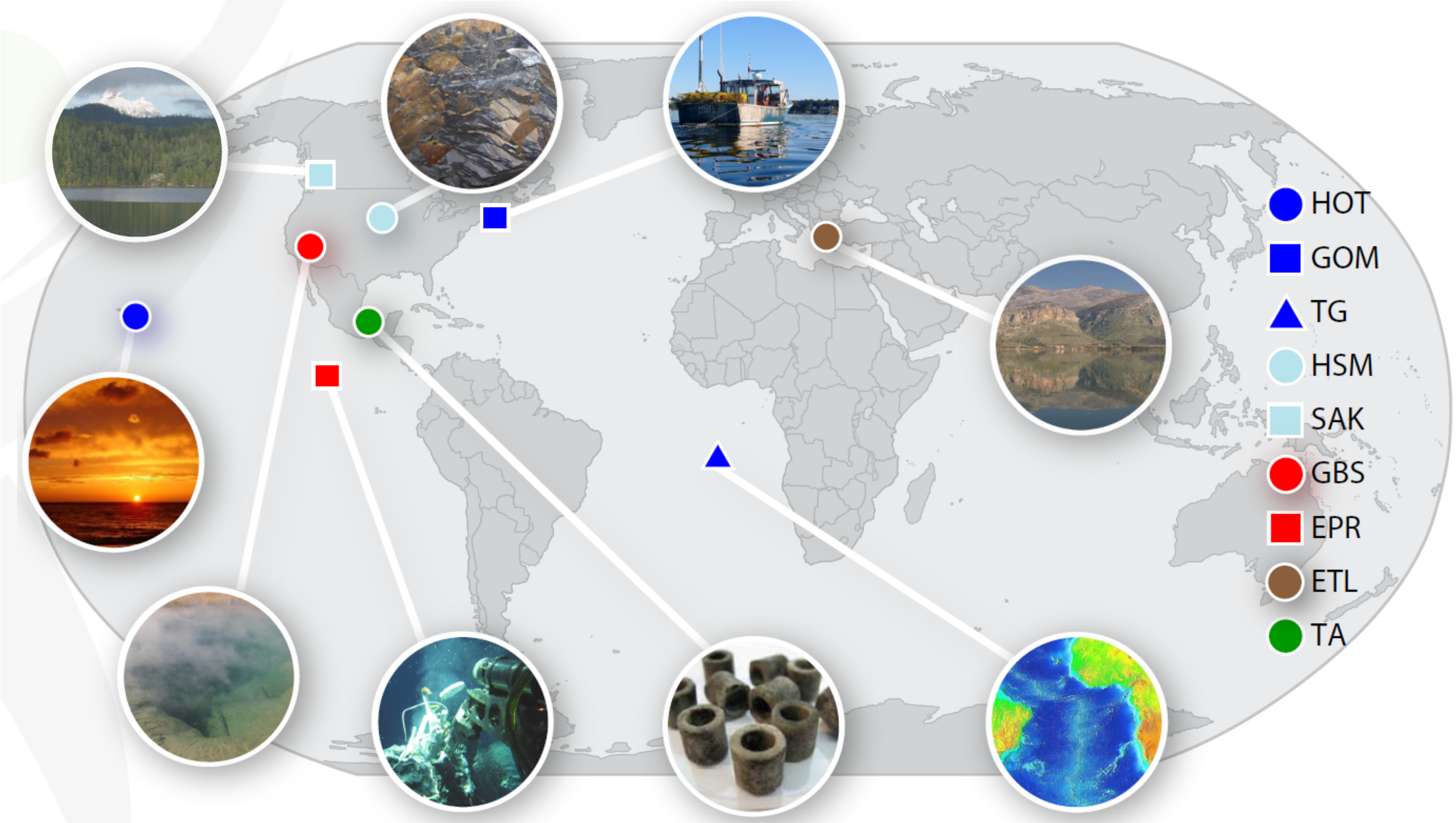
Brian Hedlund

Christian Rinke

Alex Sczyrba

Patrick Schwientek

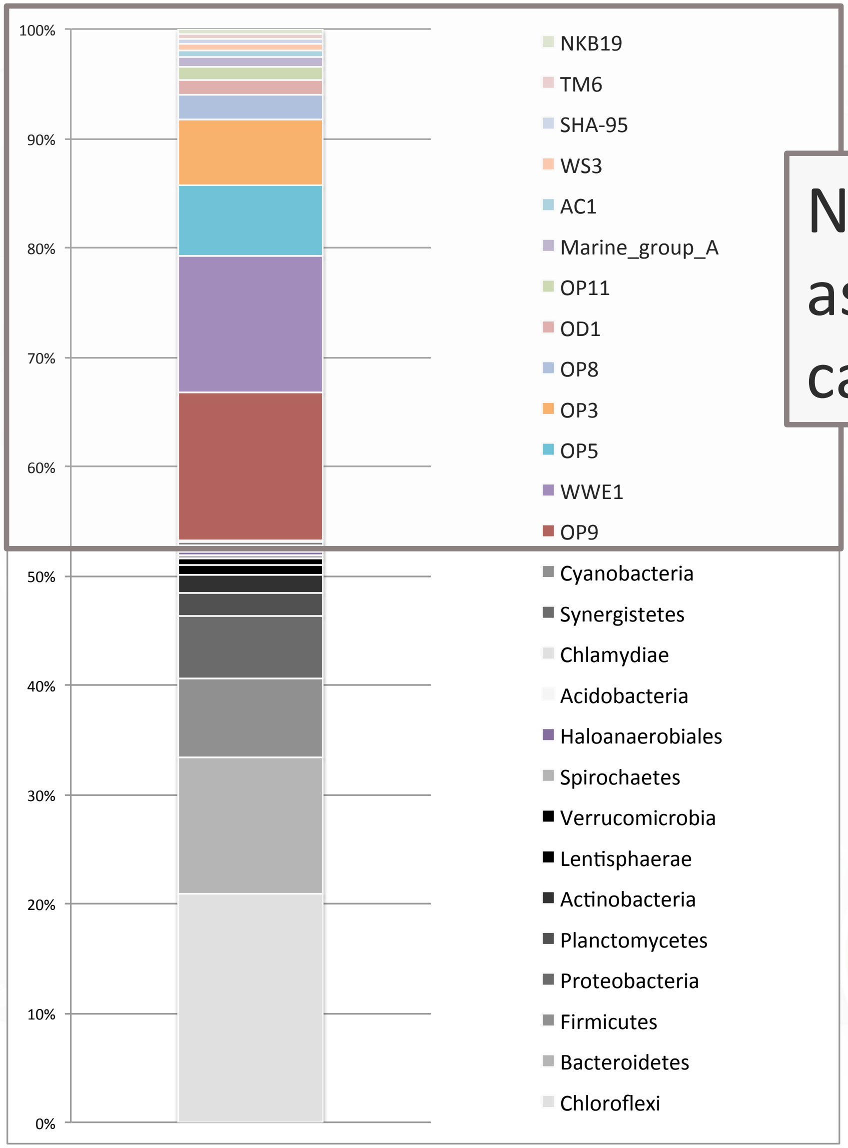
Samples from 9 sites were selected



Some sites of high underexplored diversity

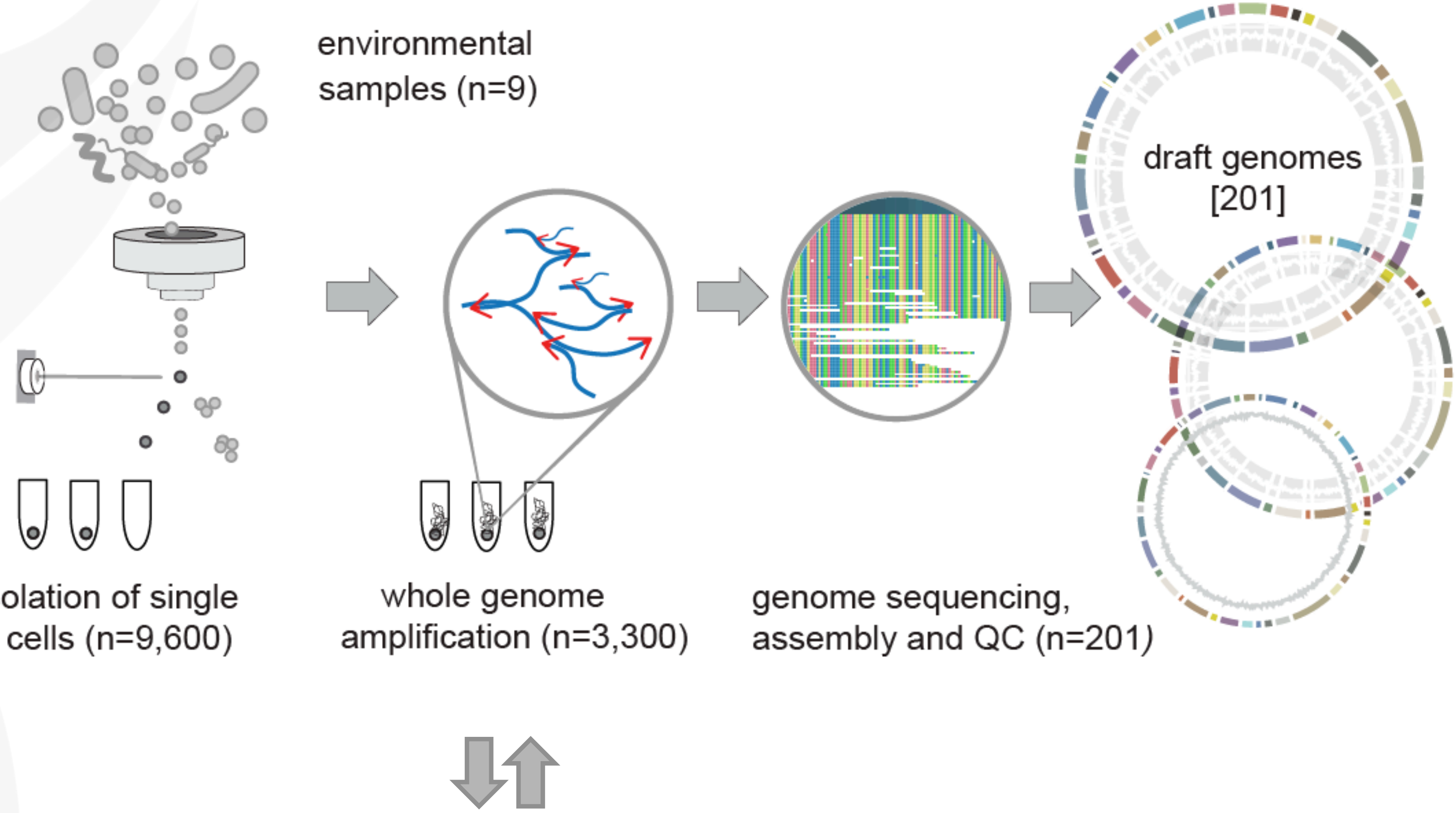
Sakinaw Lake

16S pyrotag data



Nearly 50% of tags assigned to candidate phyla

From nearly 10,000 cells to 200 draft genomes



environmental
samples (n=9)

isolation of single
cells (n=9,600)

whole genome
amplification (n=3,300)

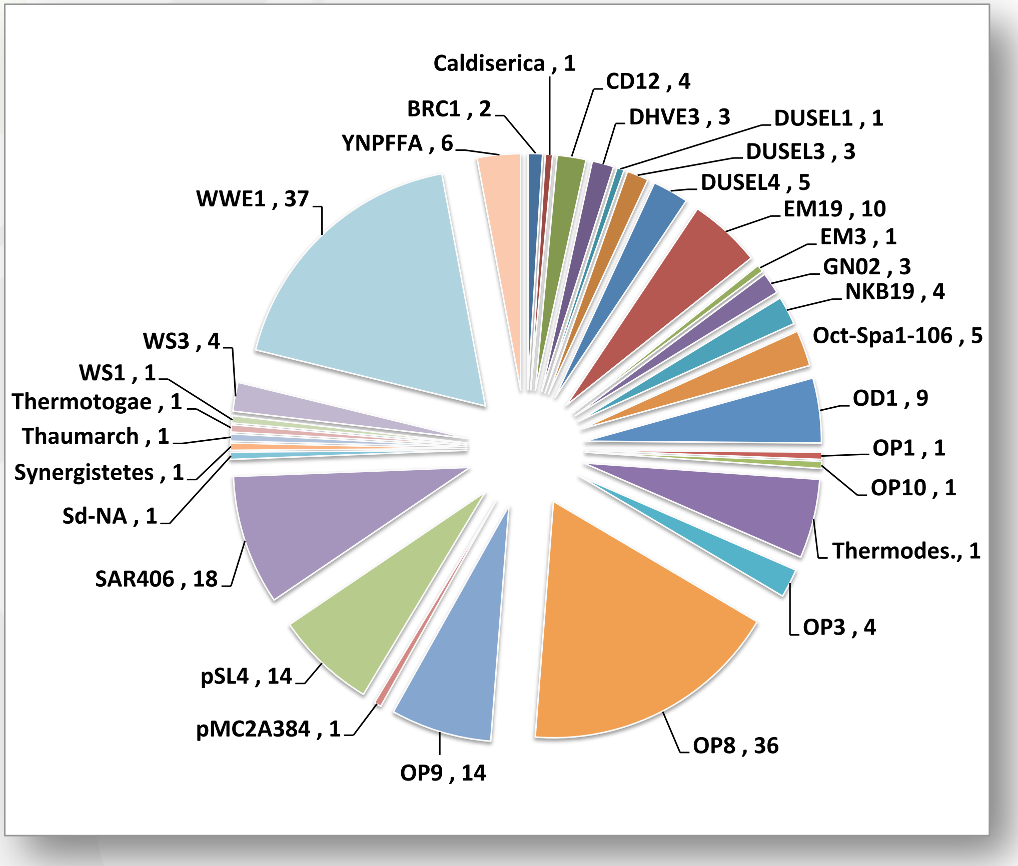
genome sequencing,
assembly and QC (n=201)

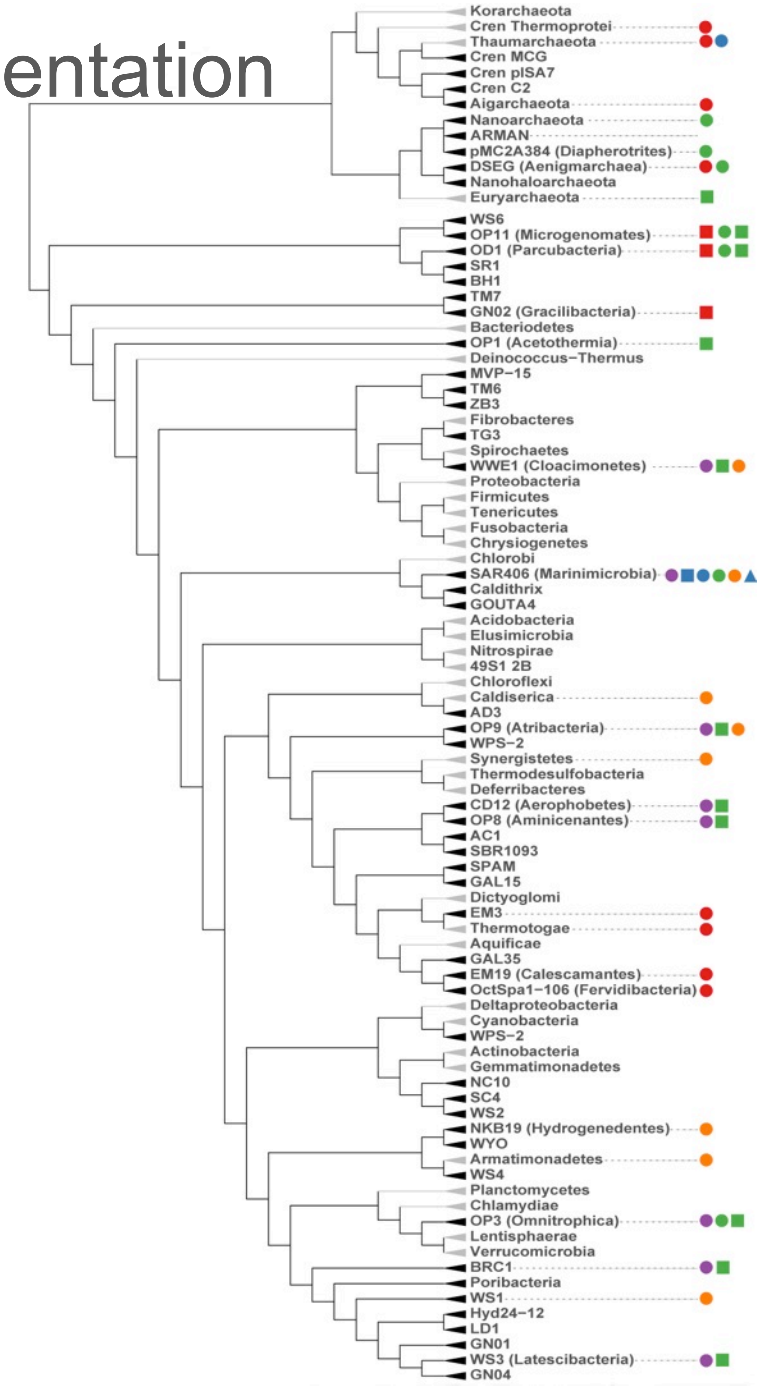
draft genomes
[201]

16S identification

Rinke et al, 2013 (Nature)

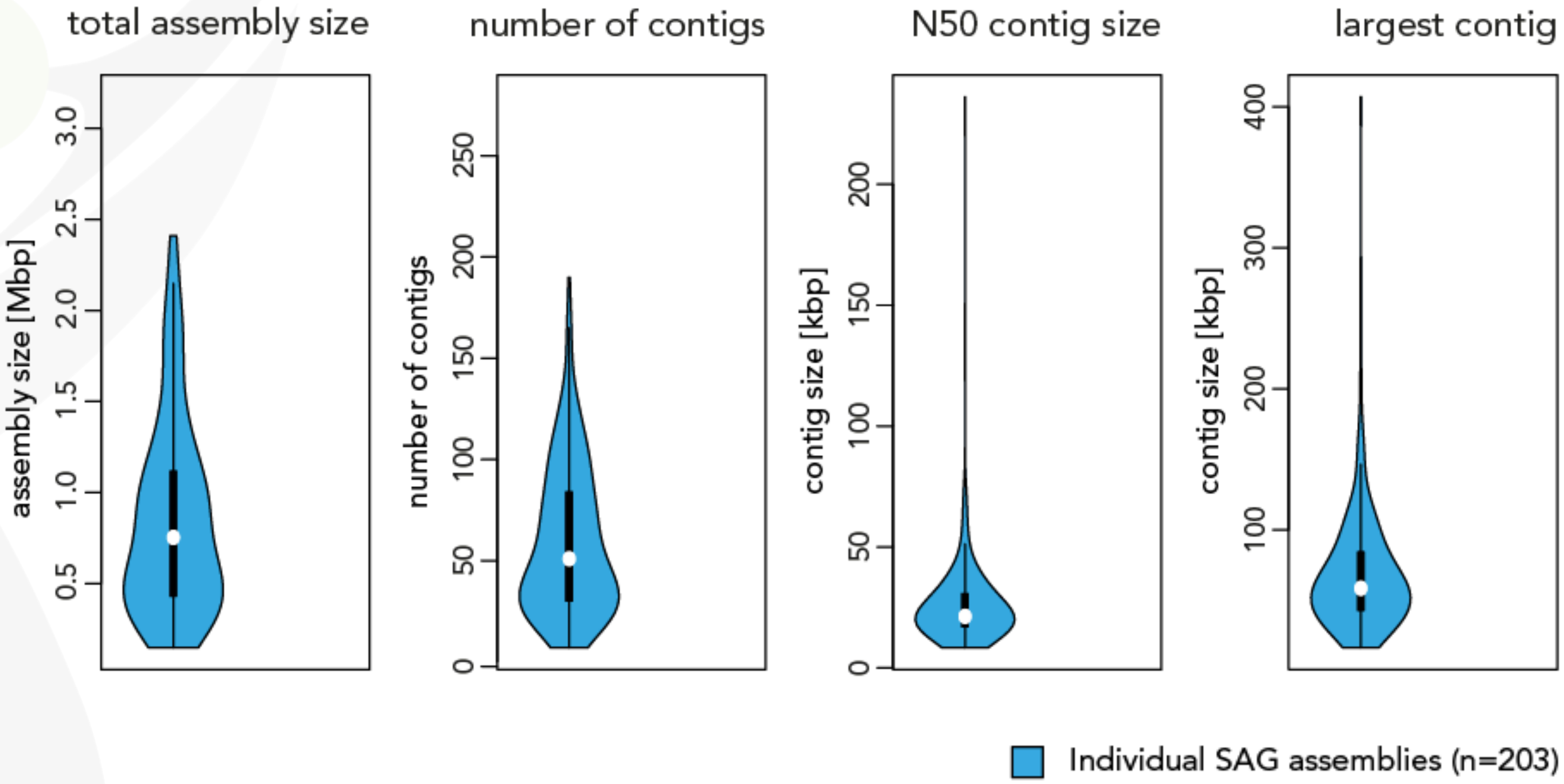
A good candidate phyla representation





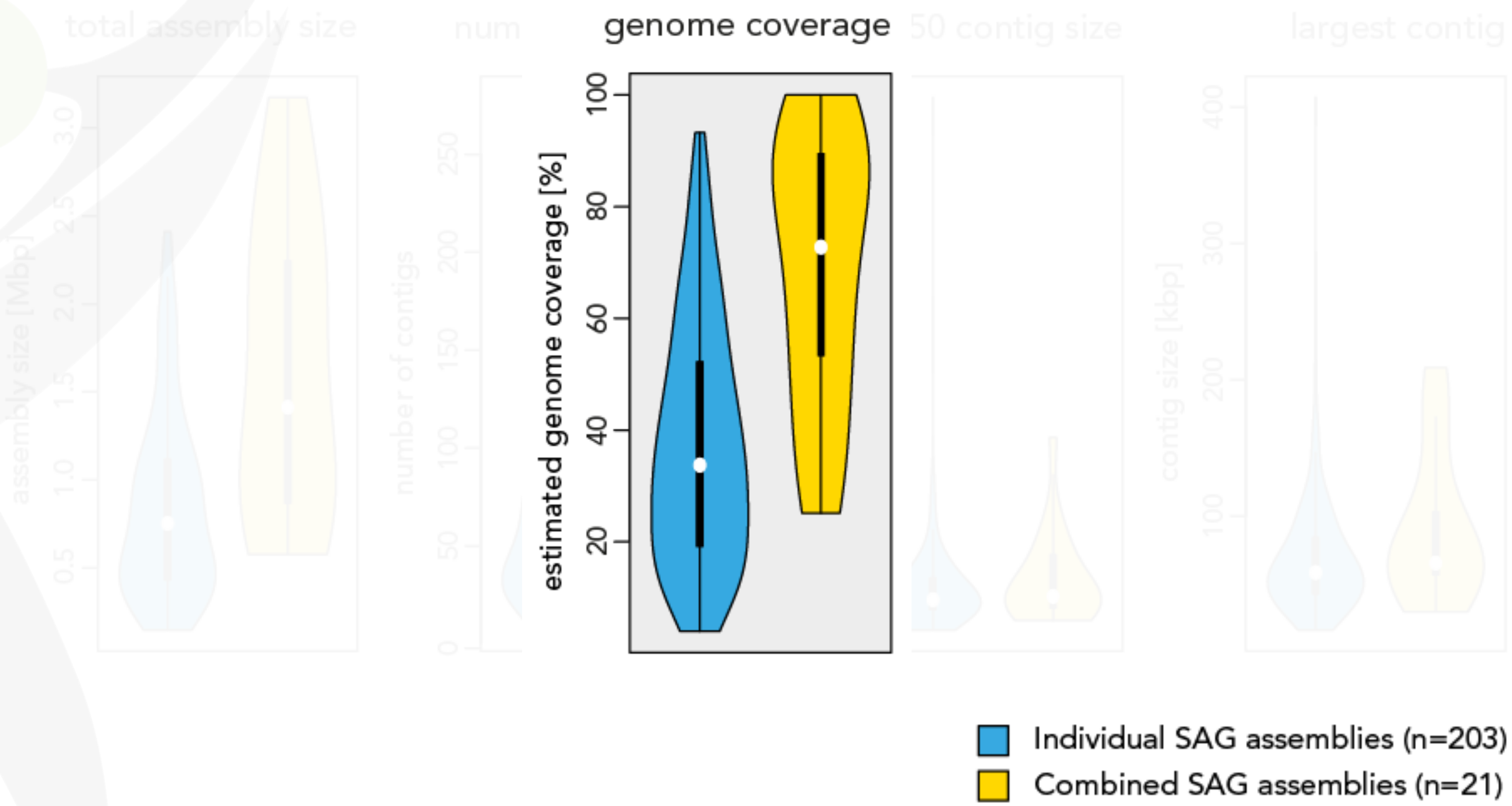
Rinke et al 2013 (Nature)

Assembly statistics of SAGs



Rinke et al 2013 (Nature)

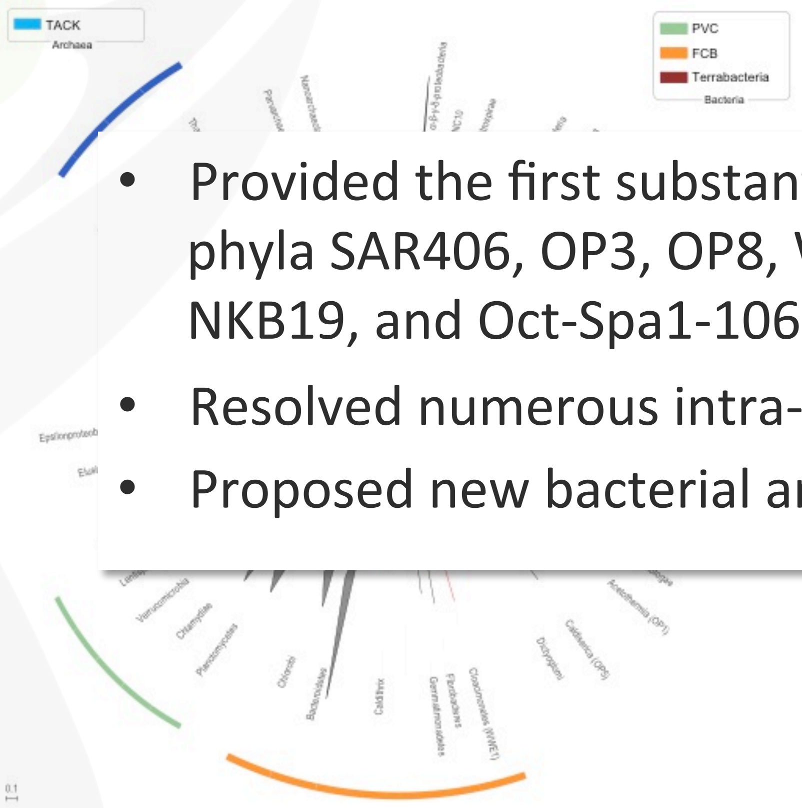
Estimated genome recoveries



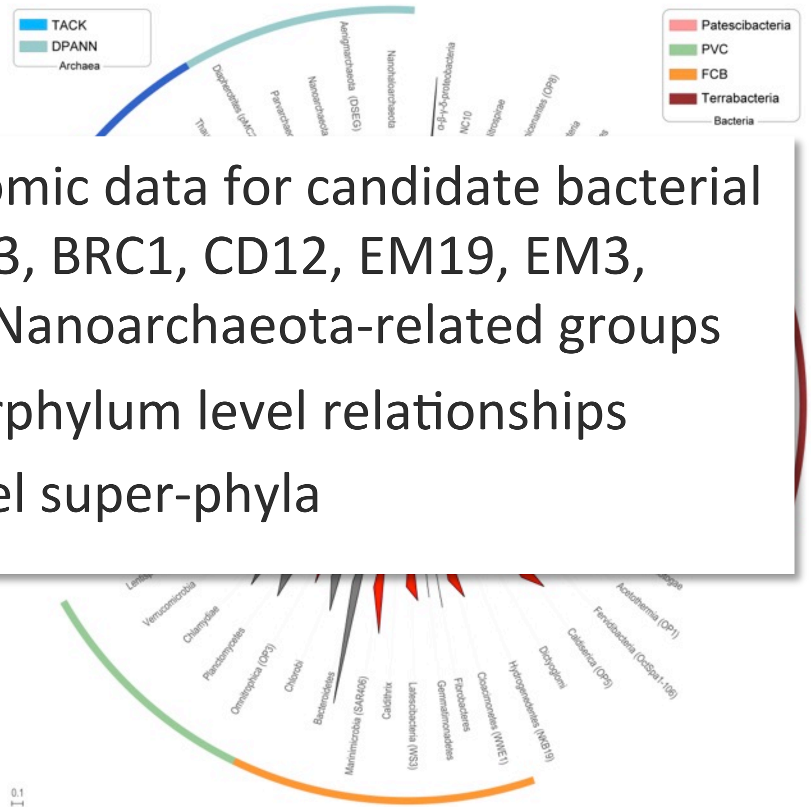
Rinke et al 2013 (Nature)

Filling some gaps in the ToL

before



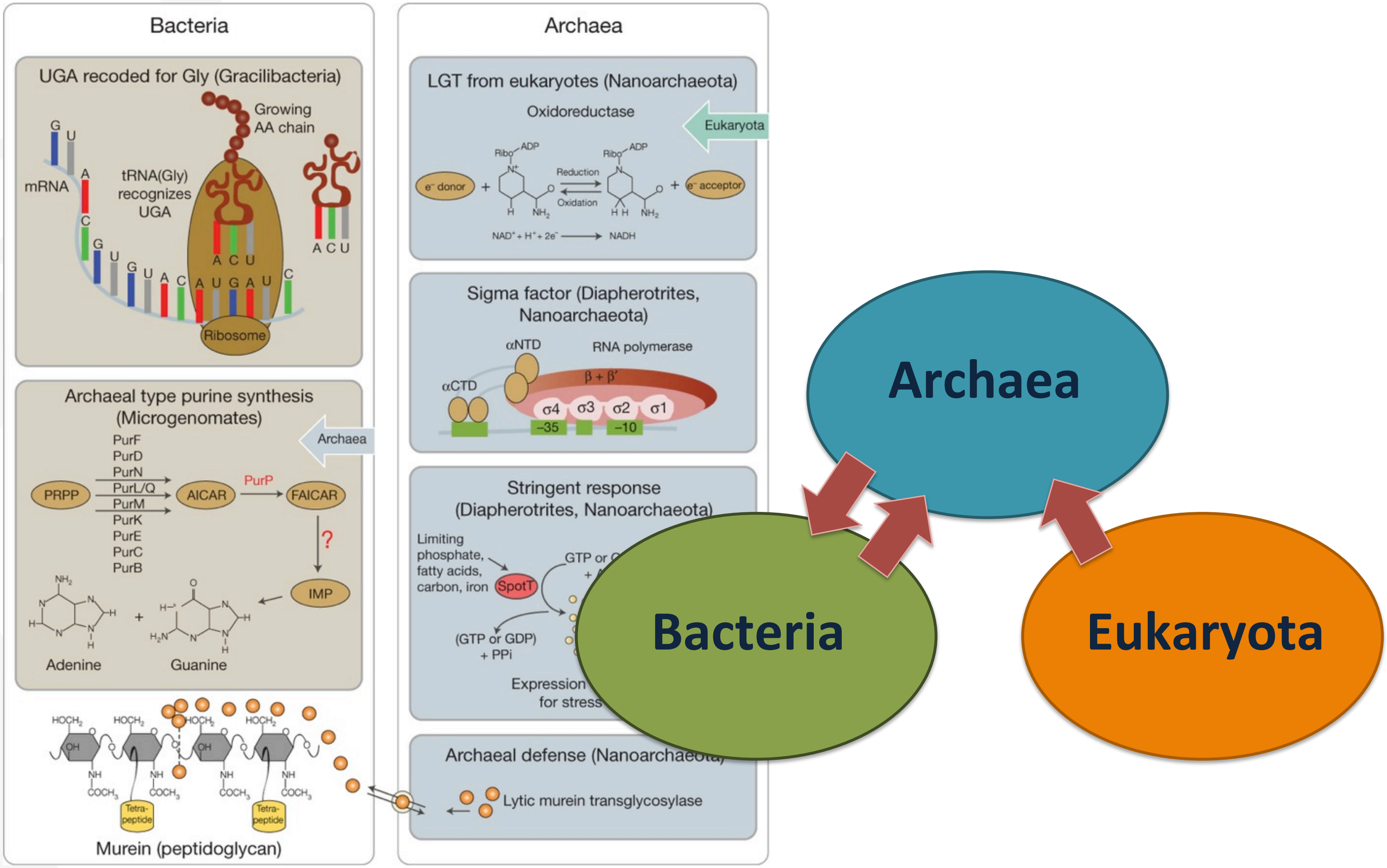
after



- Provided the first substantive genomic data for candidate bacterial phyla SAR406, OP3, OP8, WS1, WS3, BRC1, CD12, EM19, EM3, NKB19, and Oct-Spa1-106, as well Nanoarchaeota-related groups
- Resolved numerous intra-and interphylum level relationships
- Proposed new bacterial and archael super-phyla

Rinke et al 2013 (Nature)

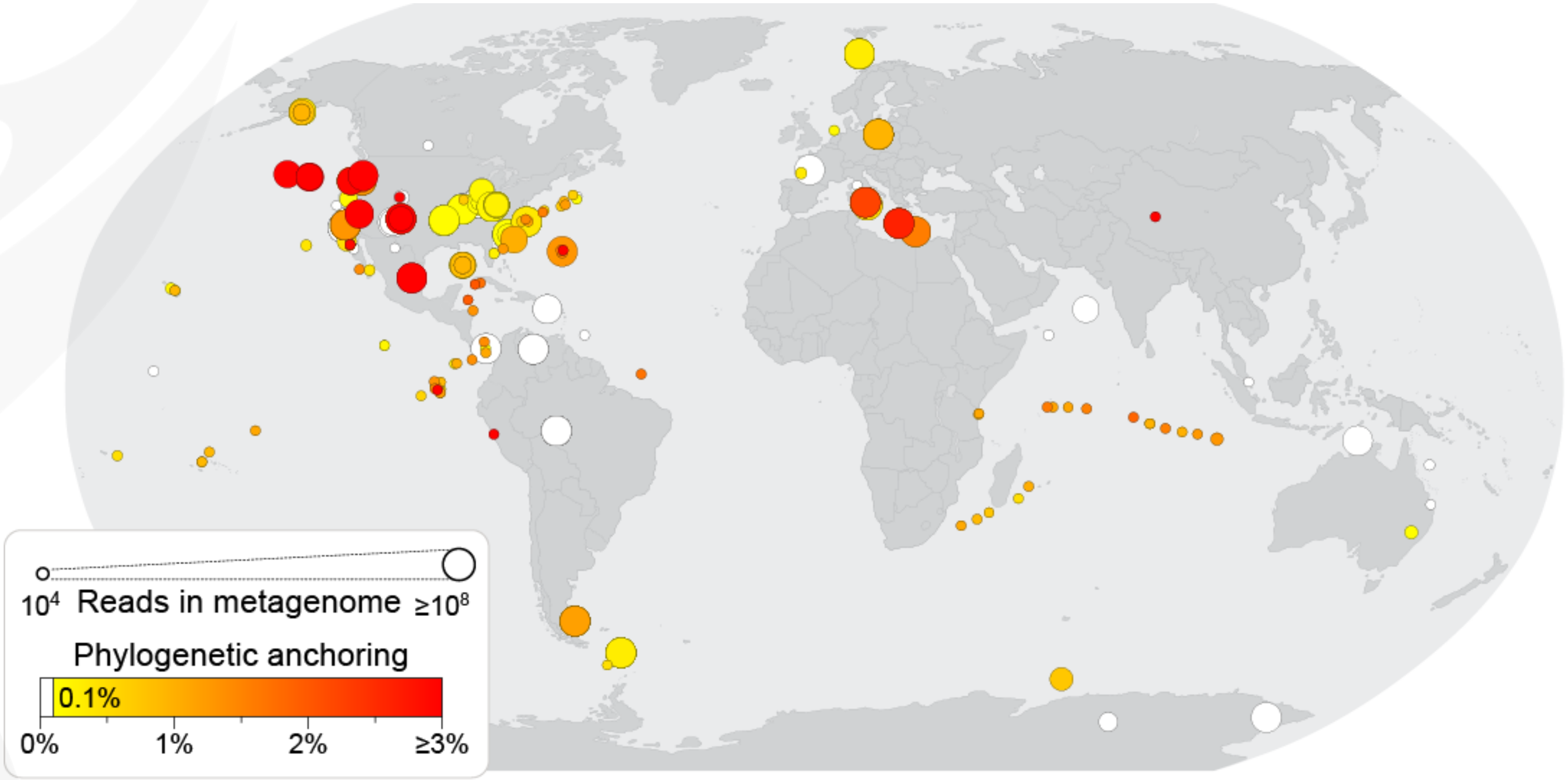
Genomic and functional novelty



Rinke et al 2013 (Nature)

Single cells as phylogenetic anchors

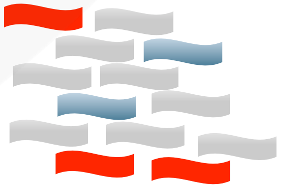
Example: Sakinaw Lake



Rinke et al 2013 (Nature)

Single cells assist in metagenome assignments

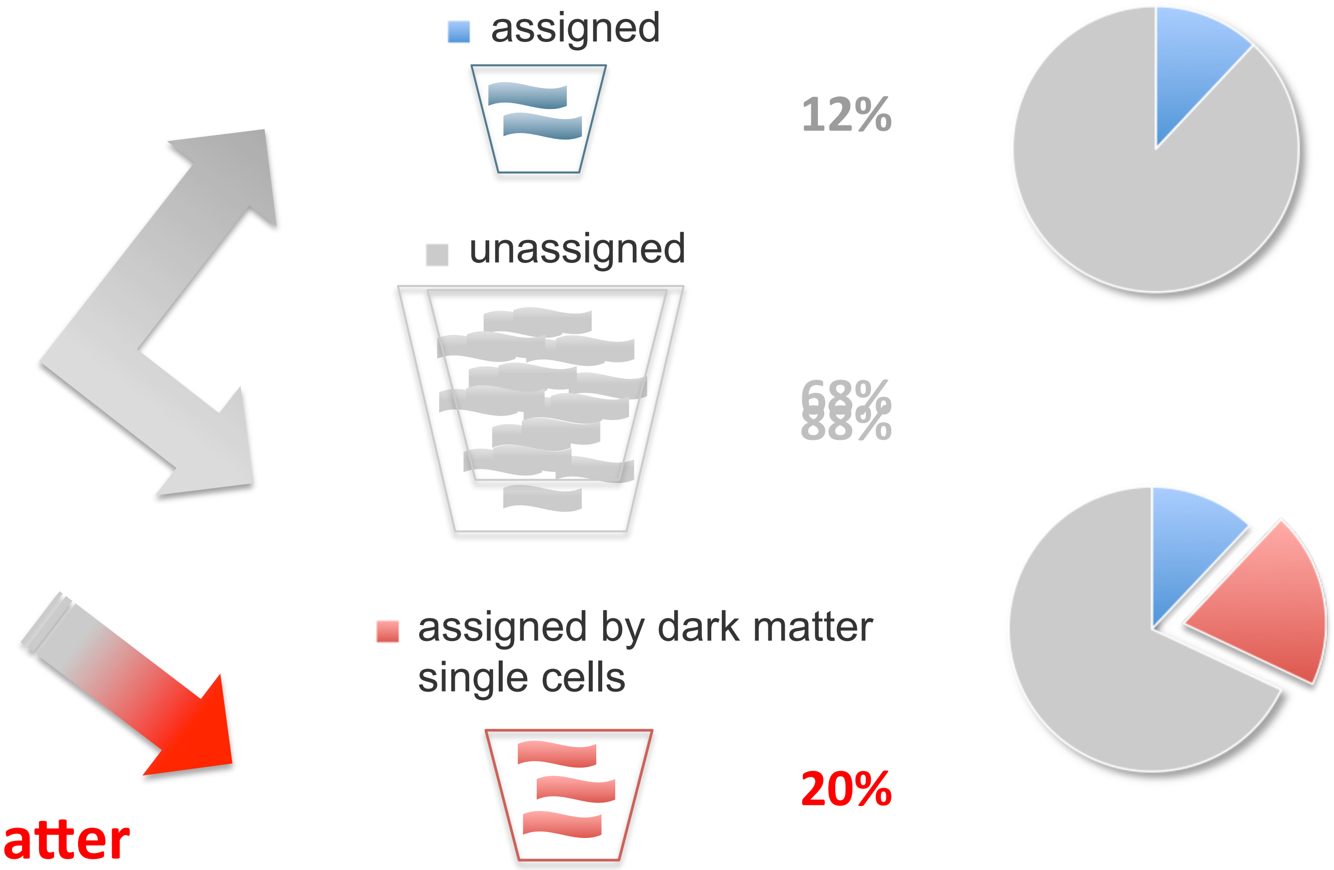
Sakinaw Lake



metagenome

With dark matter single cells

Who can do what?



🟦 assigned

12%

⬜ unassigned

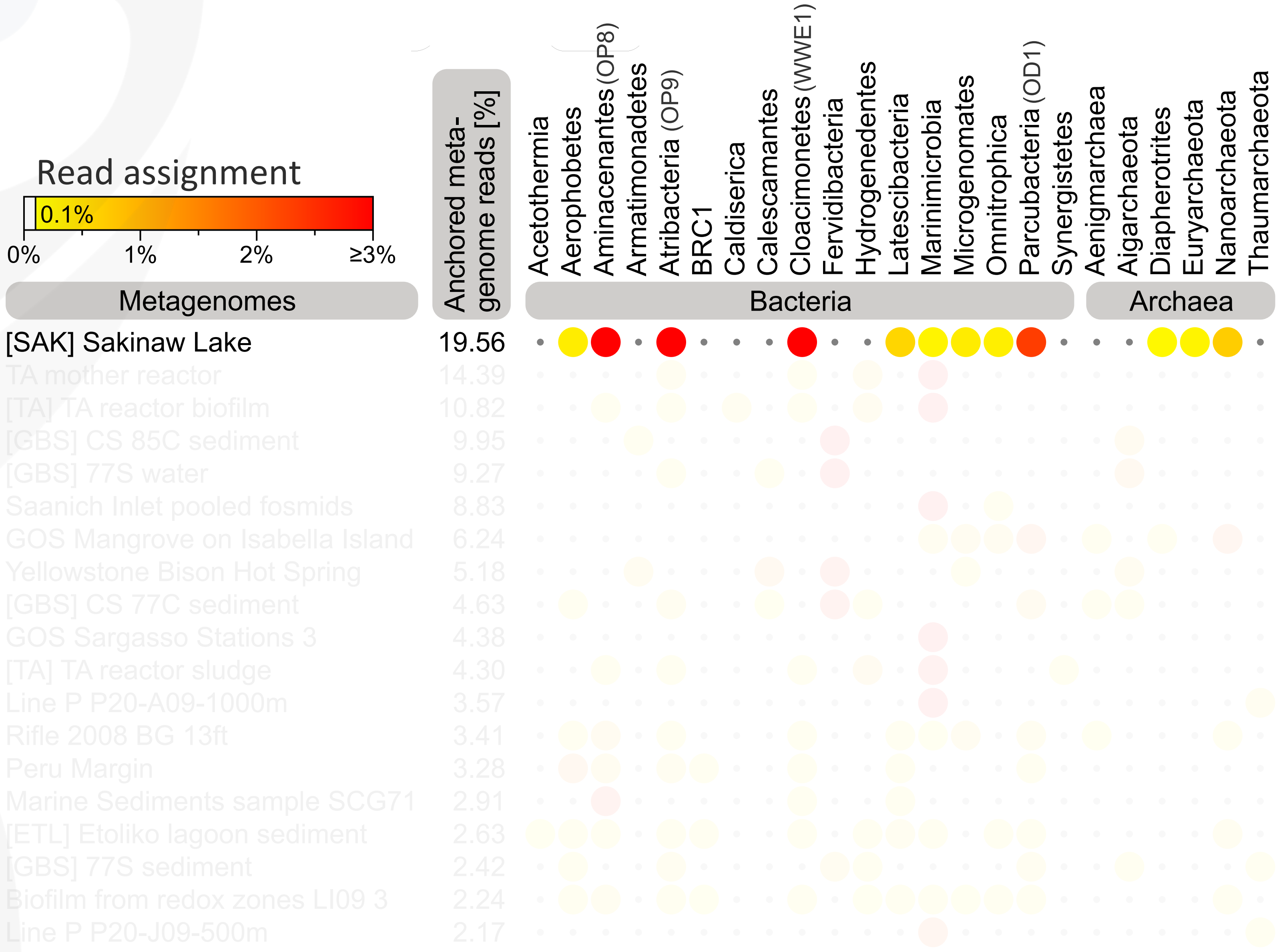
68%

88%

🟥 assigned by dark matter single cells

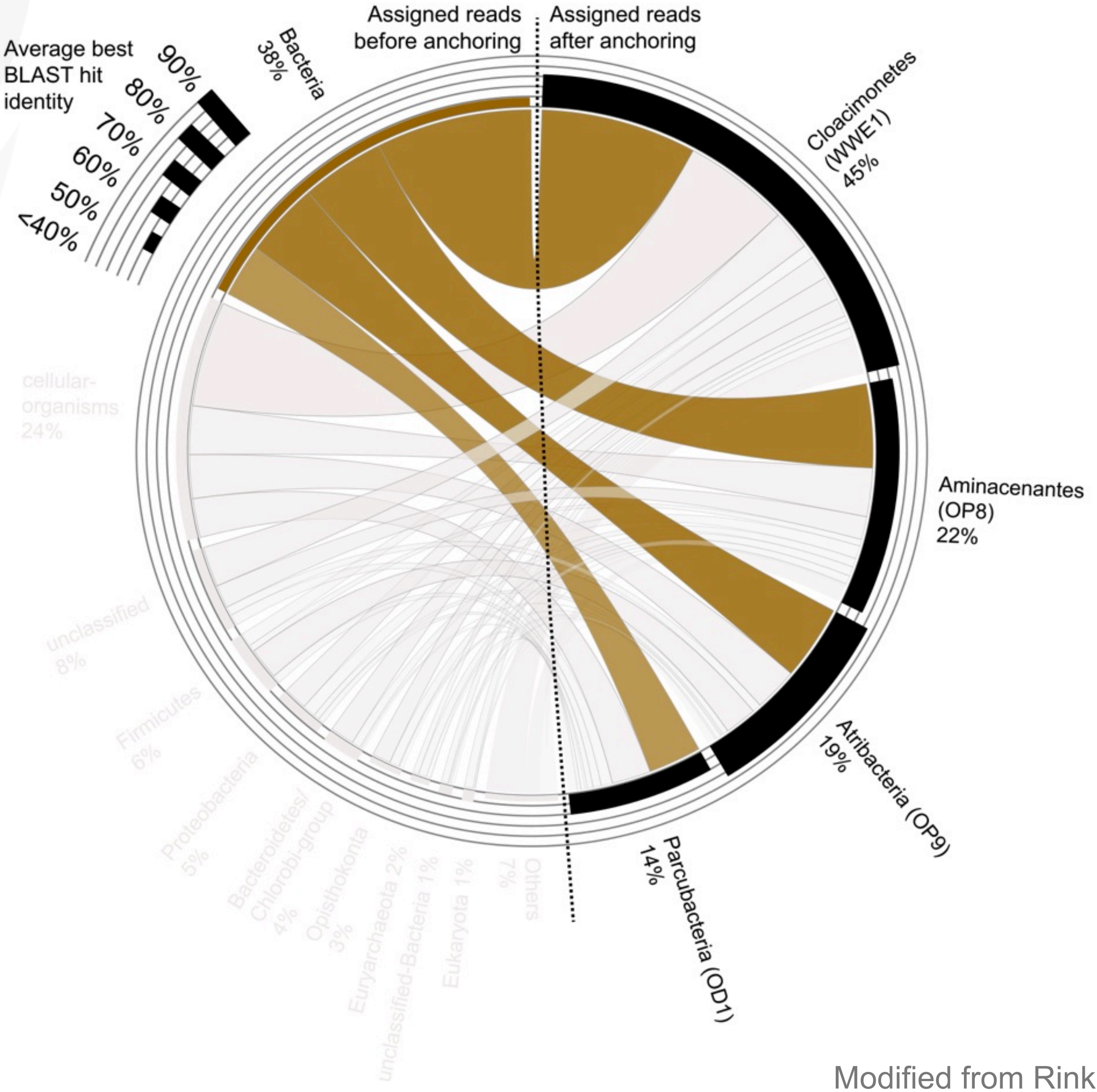
20%

Single cells as phylogenetic anchors



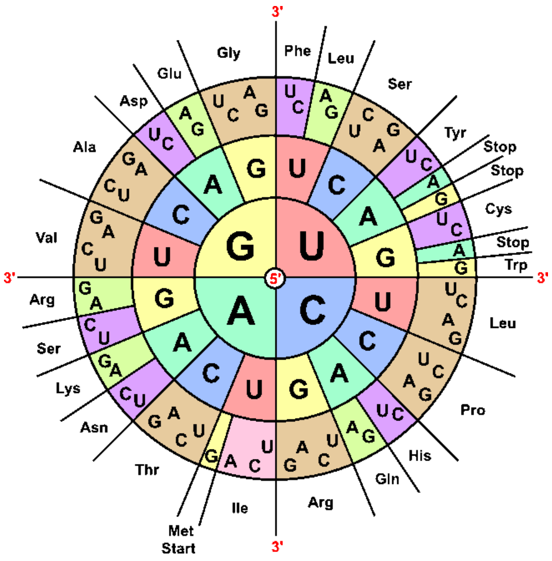
Rinke et al 2013 (Nature)

Single cells as phylogenetic anchors



Modified from Rinke et al 2013 (Nature)

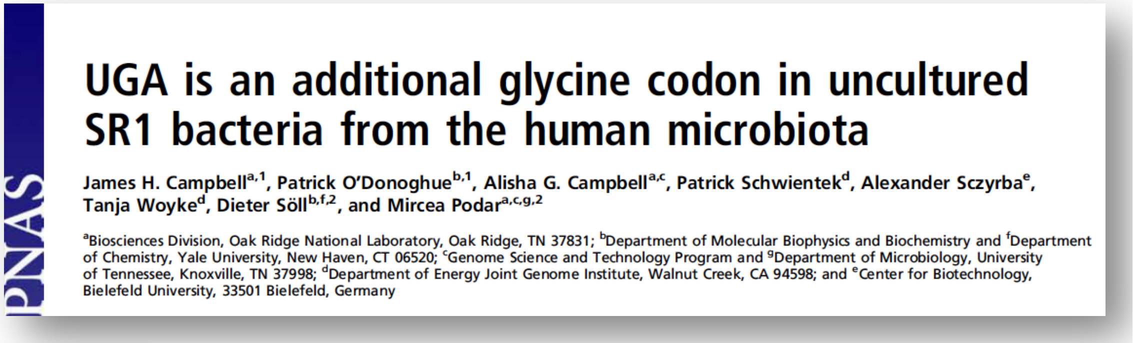
Novel stop codon reassignment



UAG ("amber")
UAA ("ochre")
~~UGA ("opal")~~ → glycine

new translation table:

“Candidate Division SR1 and Gracilibacteria Code”

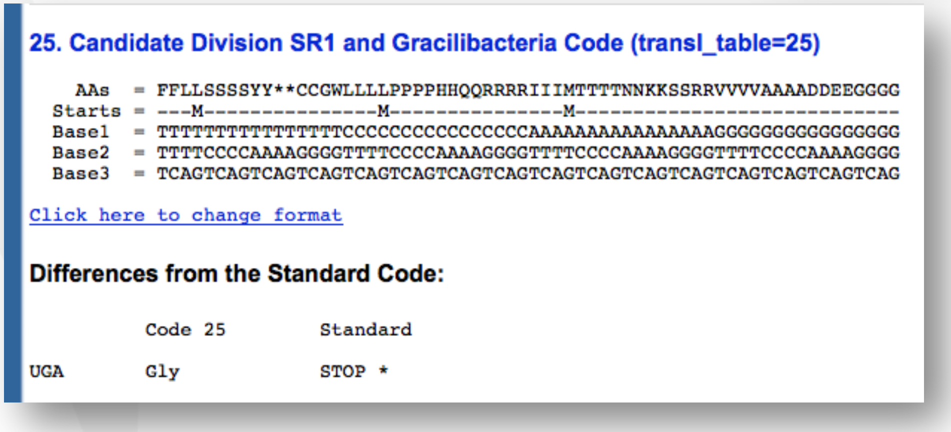


PNAS

UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota

James H. Campbell[a,1], Patrick O'Donoghue[b,1], Alisha G. Campbell[a,c], Patrick Schwientek[d], Alexander Sczyrba[e], Tanja Woyke[d], Dieter Söll[b,f,2], and Mircea Podar[a,c,g,2]

[a]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831; [b]Department of Molecular Biophysics and Biochemistry and [f]Department of Chemistry, Yale University, New Haven, CT 06520; [c]Genome Science and Technology Program and [g]Department of Microbiology, University of Tennessee, Knoxville, TN 37998; [d]Department of Energy Joint Genome Institute, Walnut Creek, CA 94598; and [e]Center for Biotechnology, Bielefeld University, 33501 Bielefeld, Germany



25. Candidate Division SR1 and Gracilibacteria Code (transl_table=25)

```
    AAs  = FFLLSSSSYY**CCGWLLLLPPPPHHQQRRRRIIIMTTTTNNKKSSRRVVVVAAAADDEEGGGG
  Starts = ---M---------------M---------------M----------------------------
  Base1  = TTTTTTTTTTTTTTTTCCCCCCCCCCCCCCCCAAAAAAAAAAAAAAAAGGGGGGGGGGGGGGGG
  Base2  = TTTTCCCCAAAAGGGGTTTTCCCCAAAAGGGGTTTTCCCCAAAAGGGGTTTTCCCCAAAAGGGG
  Base3  = TCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG
```

Click here to change format

Differences from the Standard Code:

```
          Code 25         Standard

UGA       Gly             STOP *
```

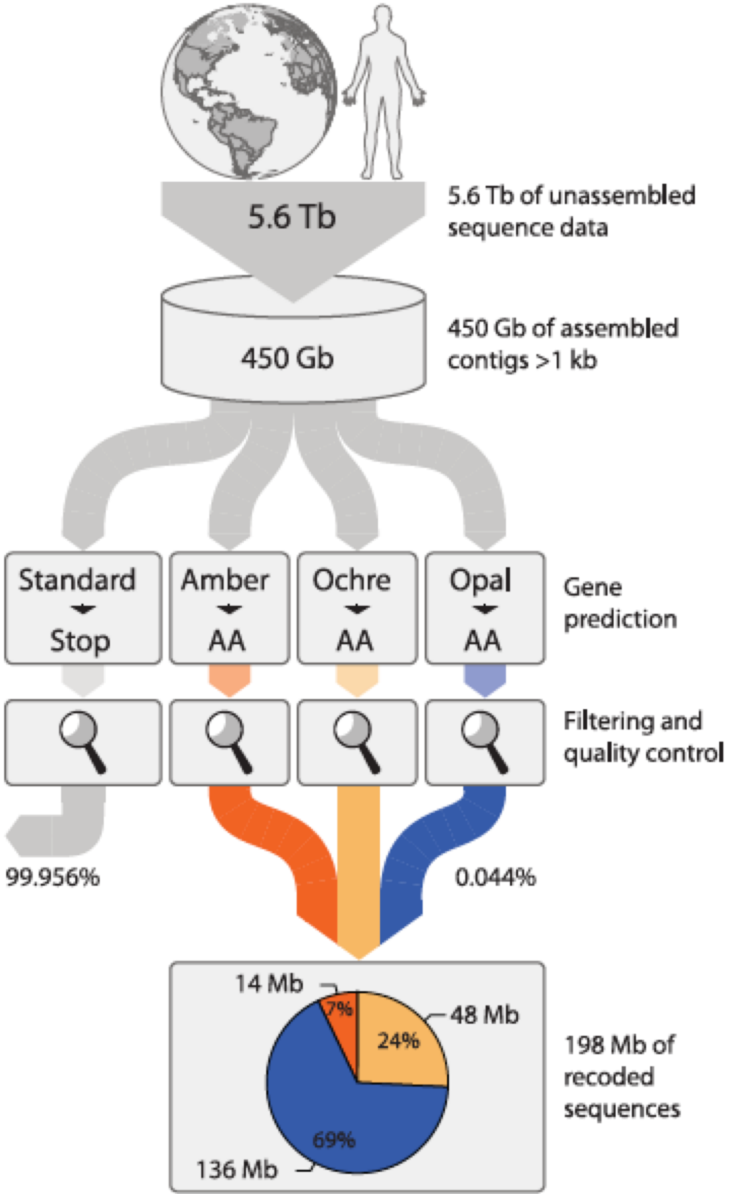
Stop codon reassignments in the wild

Eddy Rubin

Natalia Ivanova

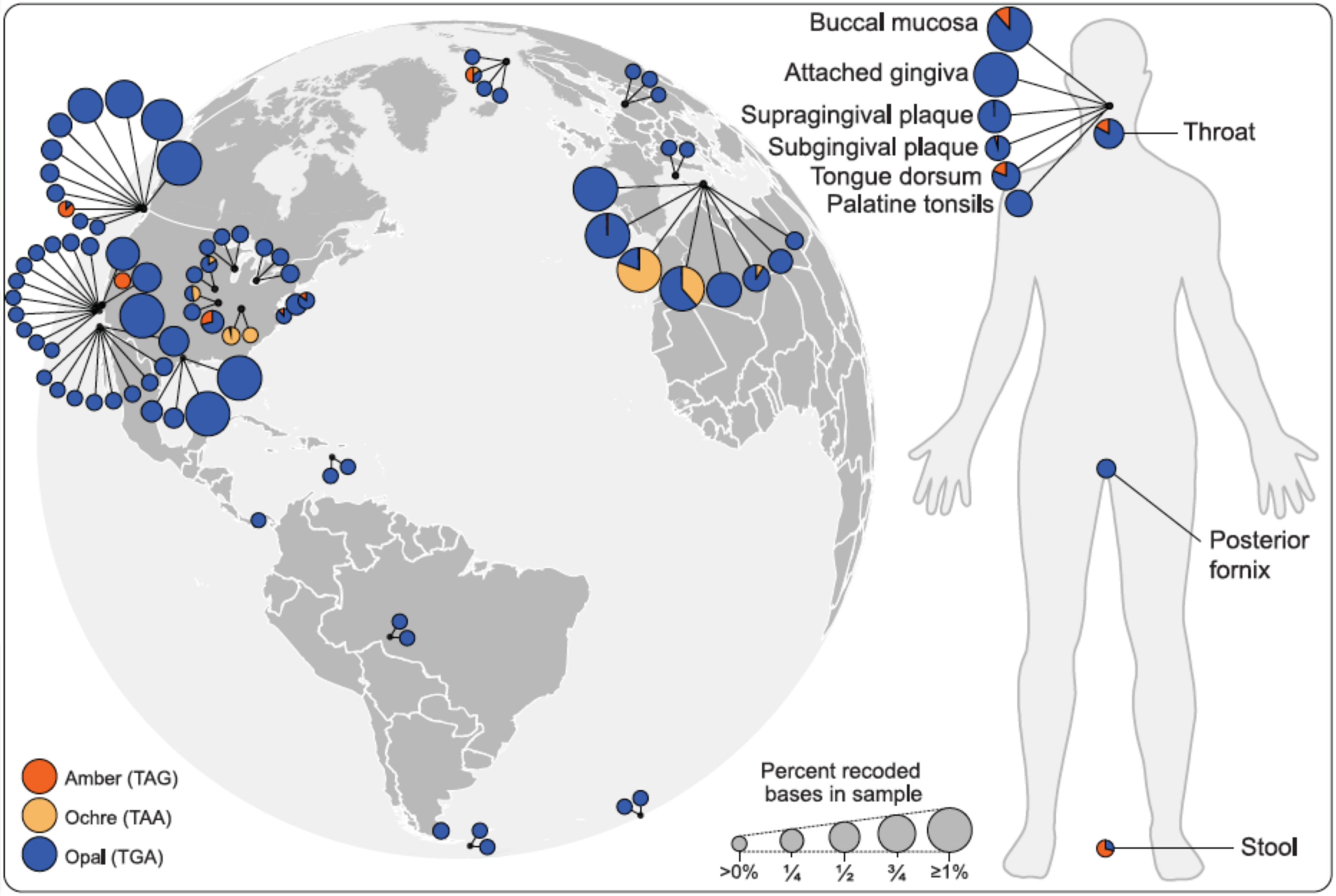
Nikos Kyrpides

Patrick Schwientek



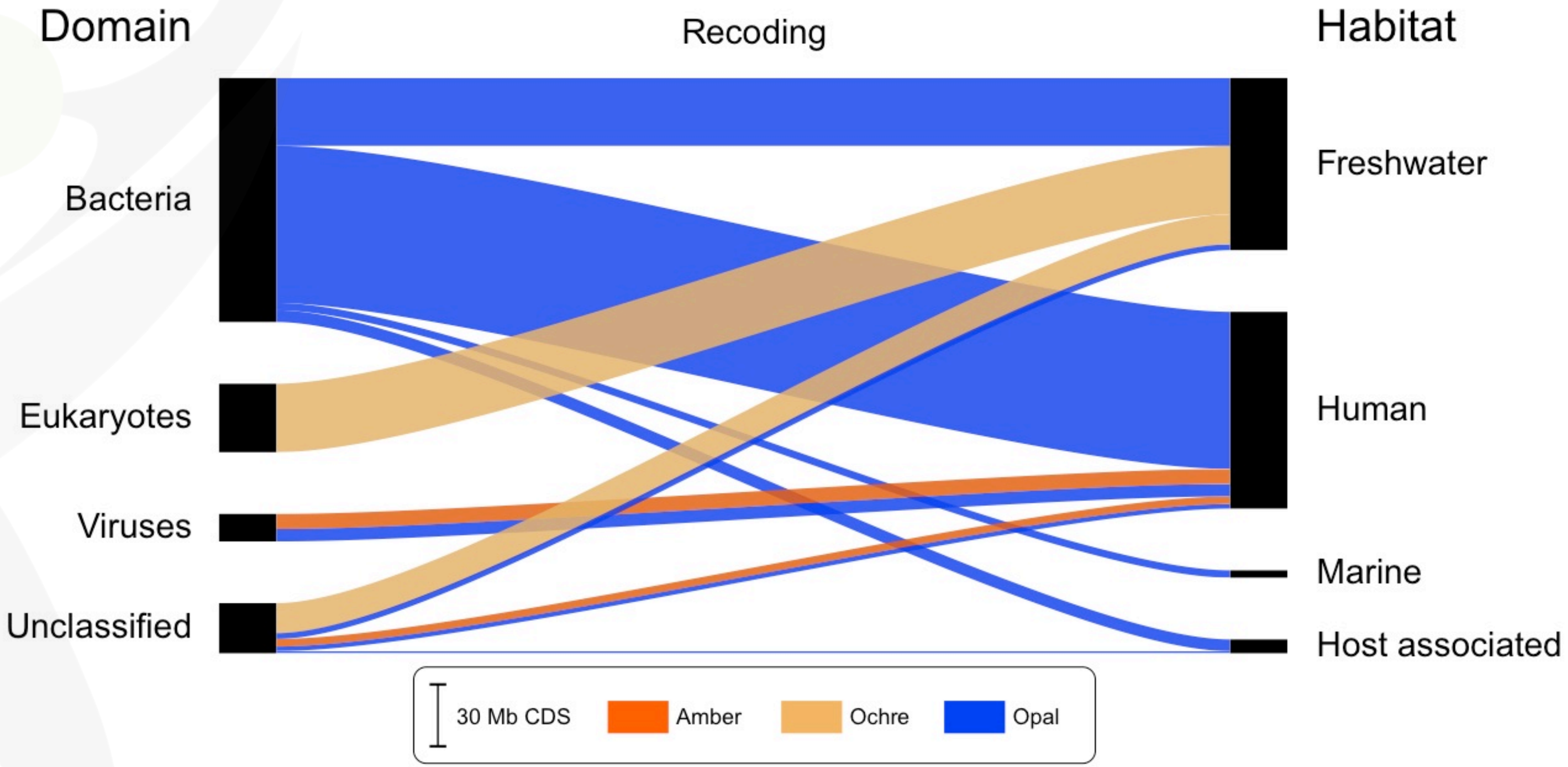
Ivanova et al 2014 (Science)

Sites of stop codon reassignments



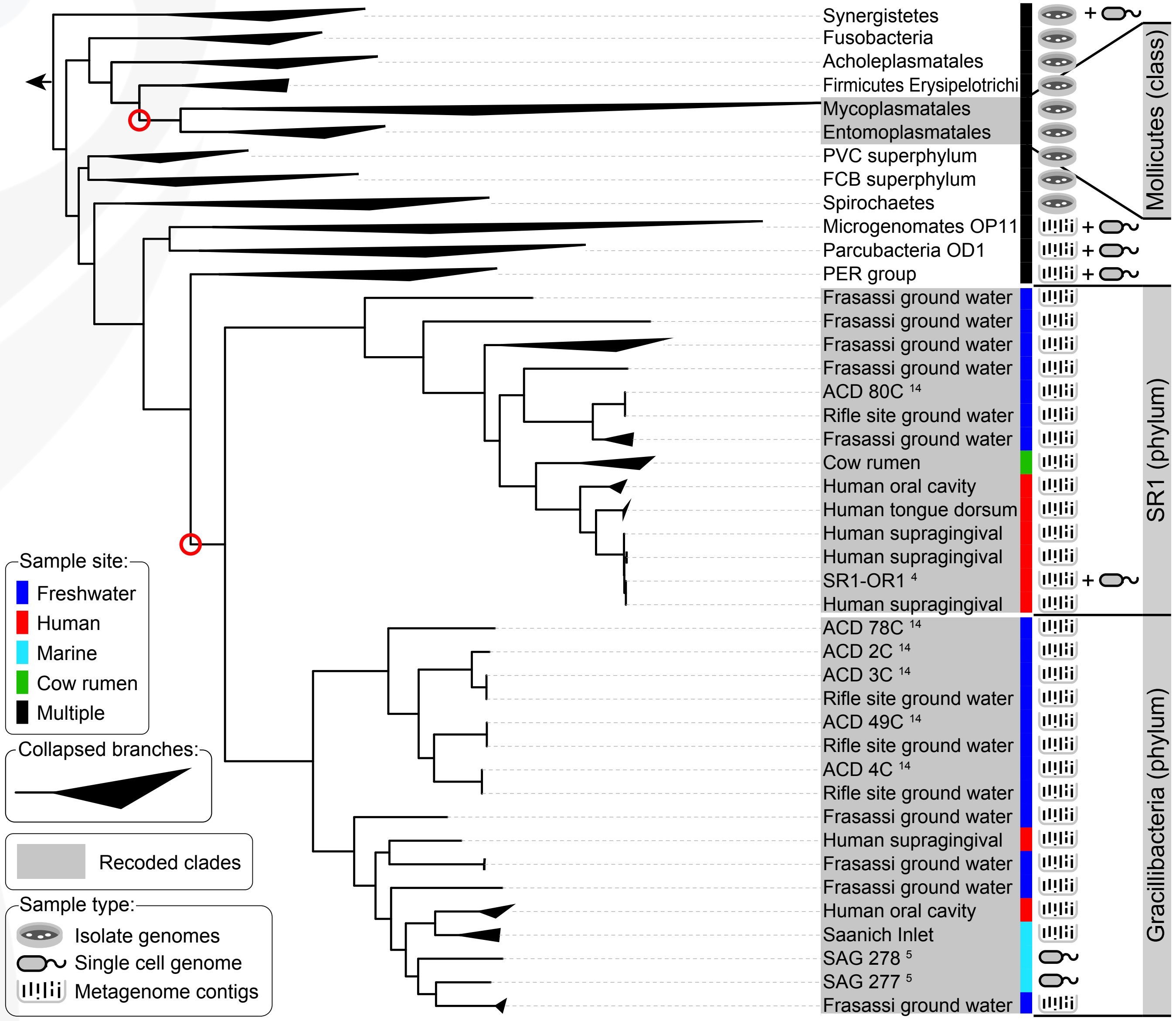
Ivanova et al 2014 (Science)

Bacterial opal recoding in freshwater and human



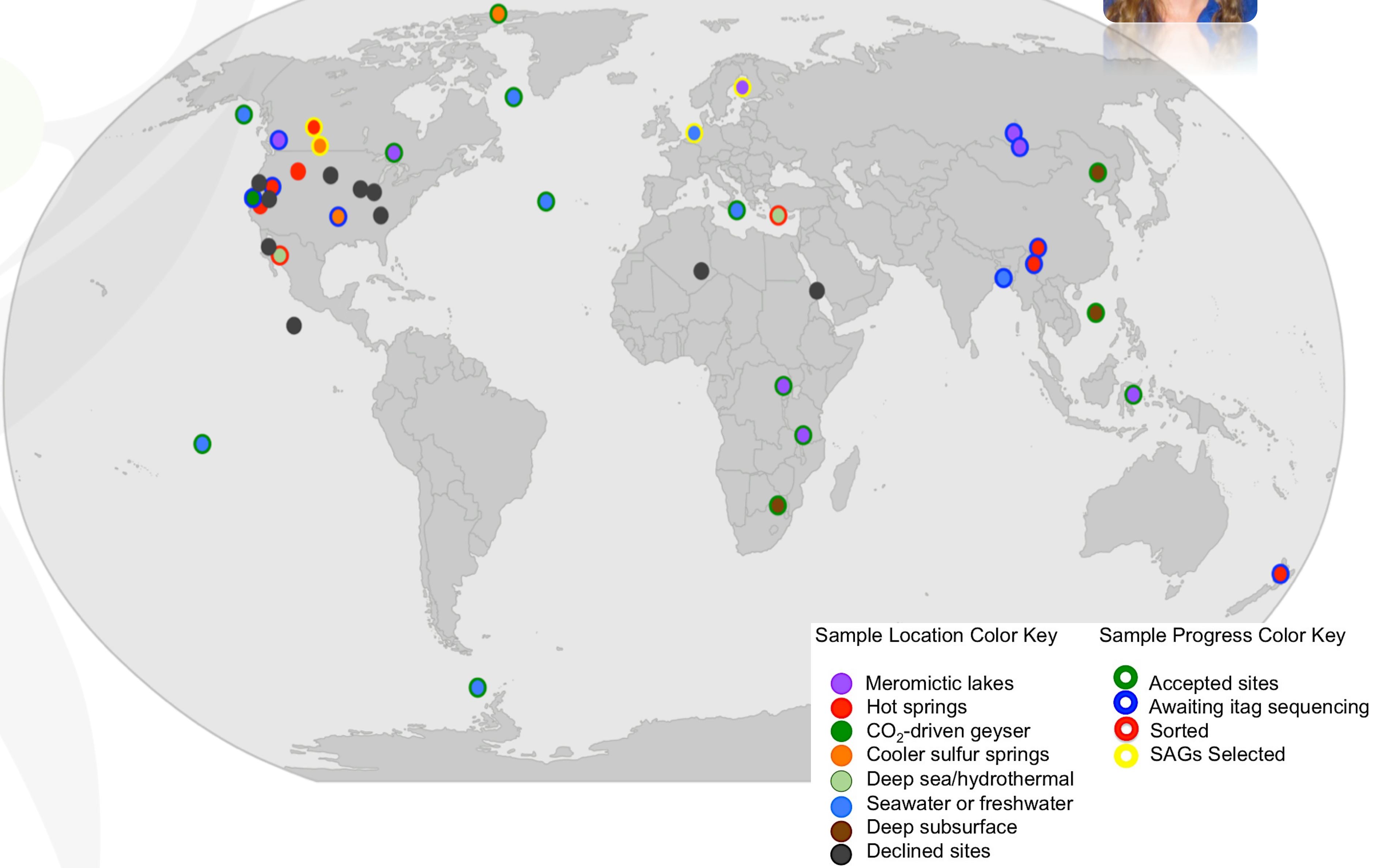
Ivanova et al 2014 (Science)

Two stop codon assignment evens in bacteria



Ivanova et al 2014 (Science)

Microbial Dark Matter phase II: A community effort



Microbial Dark Matter phase II: A community effort


Jess Jarett


Ramunas Stepanauskas
Bigelow Lab


Steven Hallam
UCB


John Spear
Colorado


Mostafa Elshaed
OK State


Sean Crowe
UBC


Brian Hedlund
U Nevada


Peter Dunfield
U of Calgary


Andreas Teske
UNC


Hailiang Dong
North China


Paul van de Wielen
KWR Water


Sari Peura
Uppsala U.


Konstaninos Kormas
U. Thessaly


Nikolai Ravin
Russian Acad. Sci.


Matt Stott
GNC Science


Chuanlun Zhang
Tongji University


Karin Rengefors
Lund University

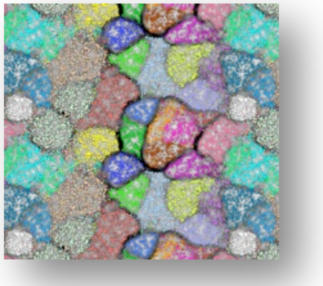

Steve Lindemann
PNNL


Nils-Kare Birkeland
U of Bergen

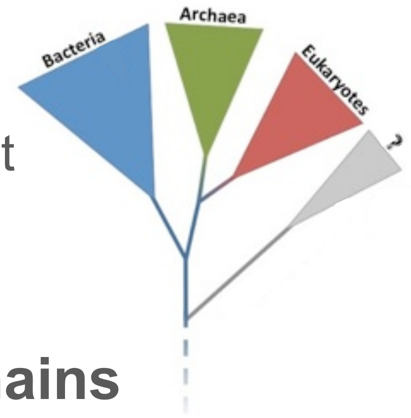
Expected outcome

- More novel discoveries

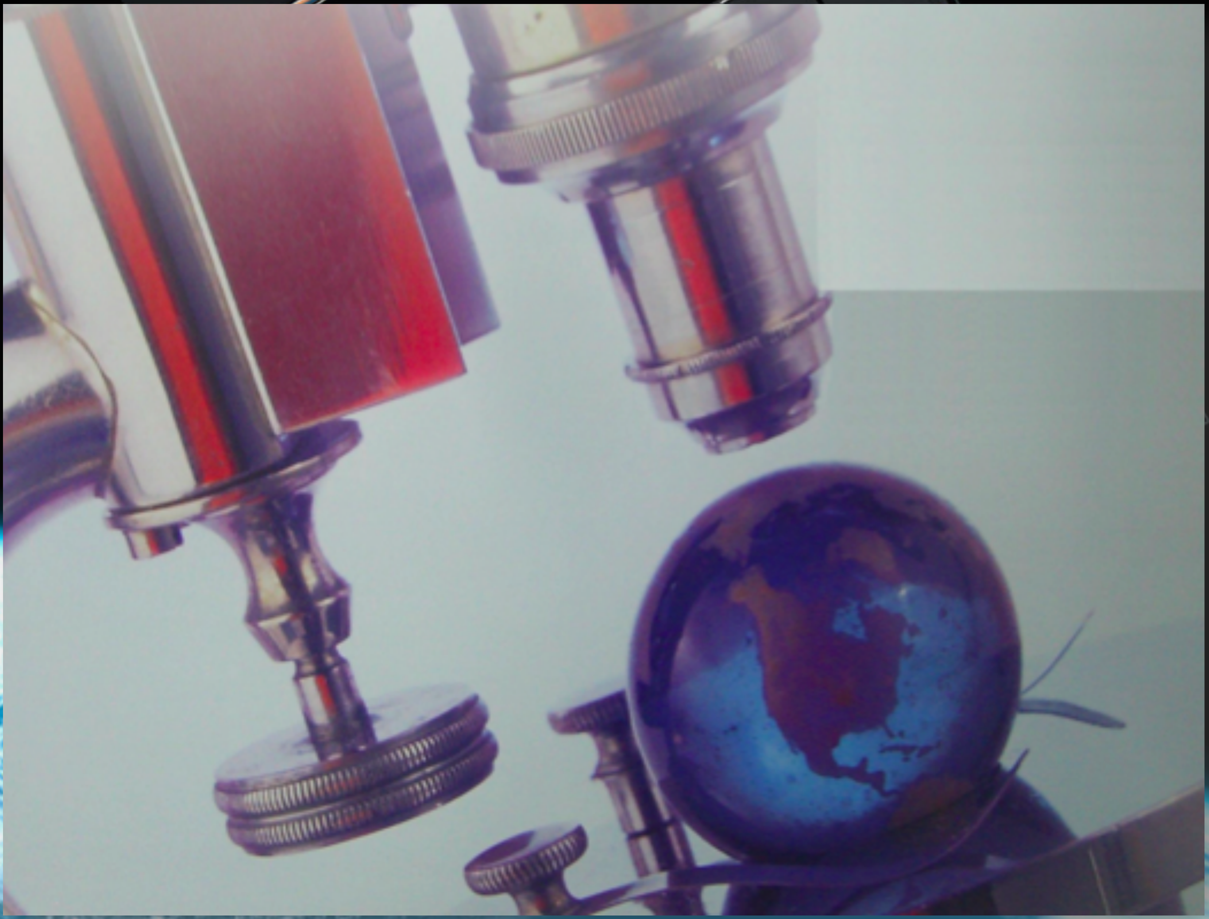
- Improved binning of metagenome data



- Understanding of DOE–relevant systems of JGI Users
- Microbial ecology & evolution
  - Functional roles of candidate phyla in the environment
  - Phylogenetic distribution of key metabolic functions
  - Co-occurrences of candidate phyla
  - **Early evolution of bacterial & archaeal domains**

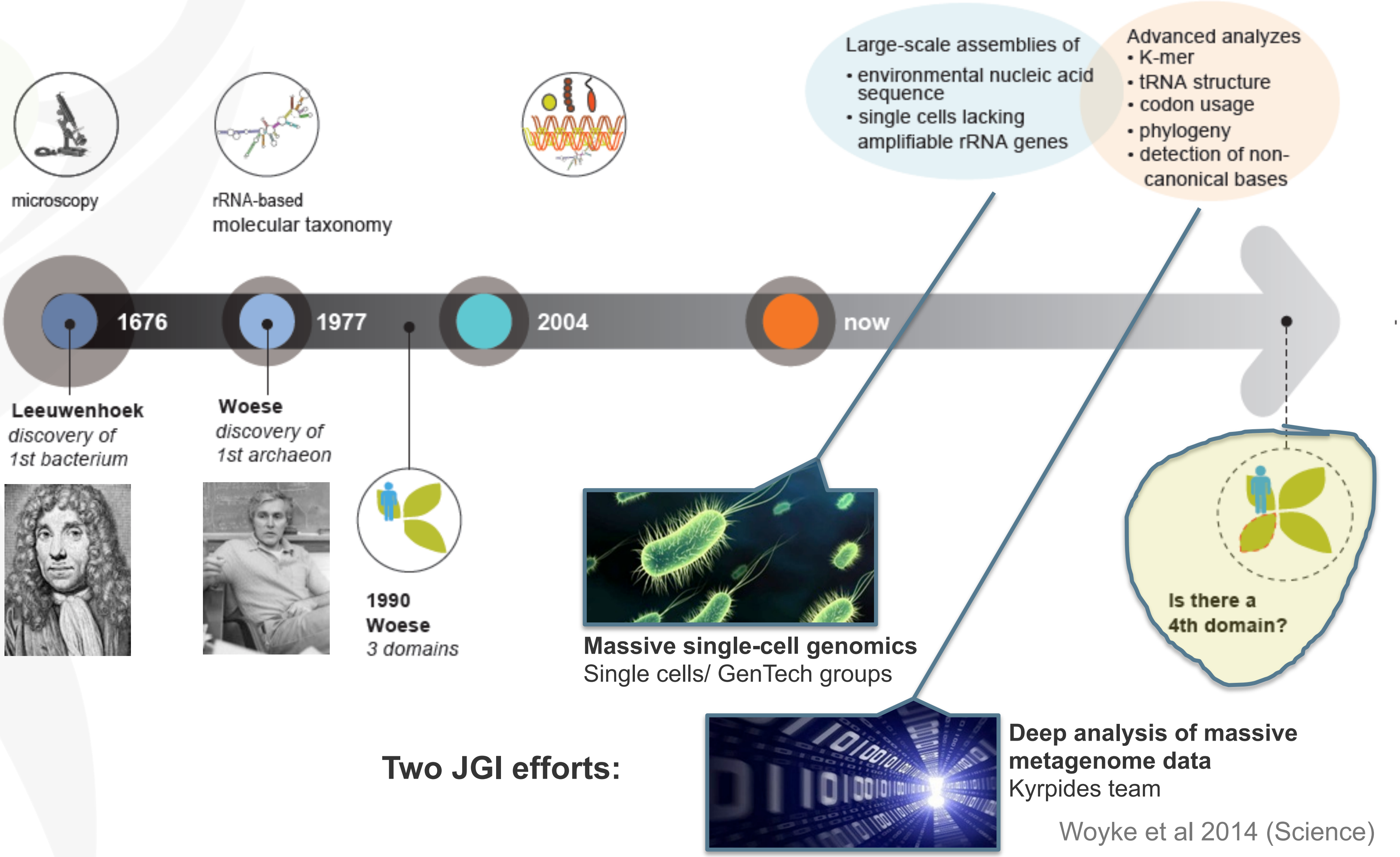


What else is out there?



The Search for new major branches

Search for new major branches



microscopy

rRNA-based molecular taxonomy

Large-scale assemblies of
- environmental nucleic acid sequence
- single cells lacking amplifiable rRNA genes

Advanced analyzes
- K-mer
- tRNA structure
- codon usage
- phylogeny
- detection of non-canonical bases

Leeuwenhoek
discovery of 1st bacterium

Woese
discovery of 1st archaeon

1990
Woese
3 domains

Massive single-cell genomics
Single cells/ GenTech groups

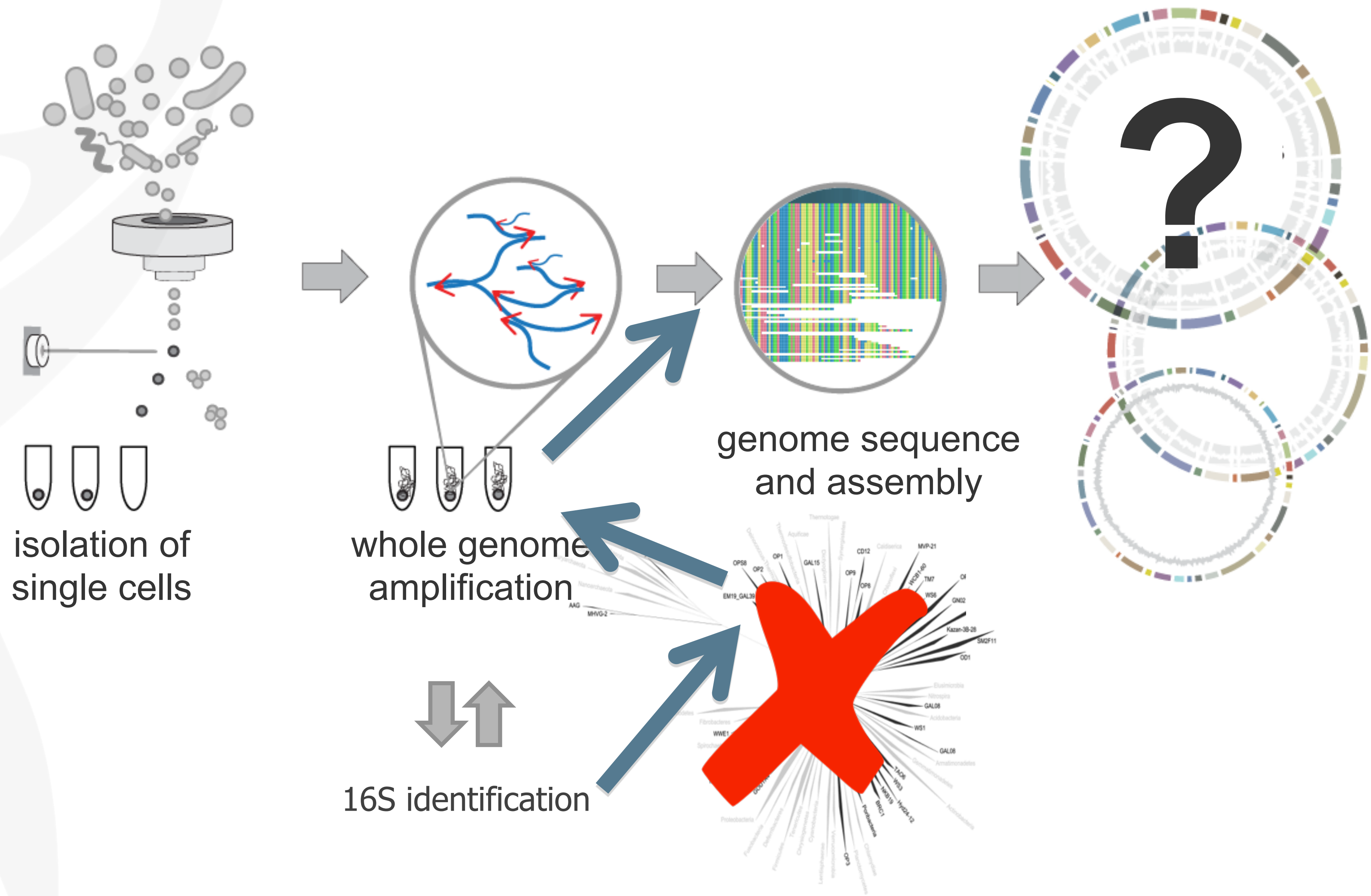
Is there a 4th domain?

Two JGI efforts:

Deep analysis of massive metagenome data
Kyrpides team

Woyke et al 2014 (Science)

One approach to look for “new major branches”



isolation of
single cells

whole genome
amplification

genome sequence
and assembly

16S identification

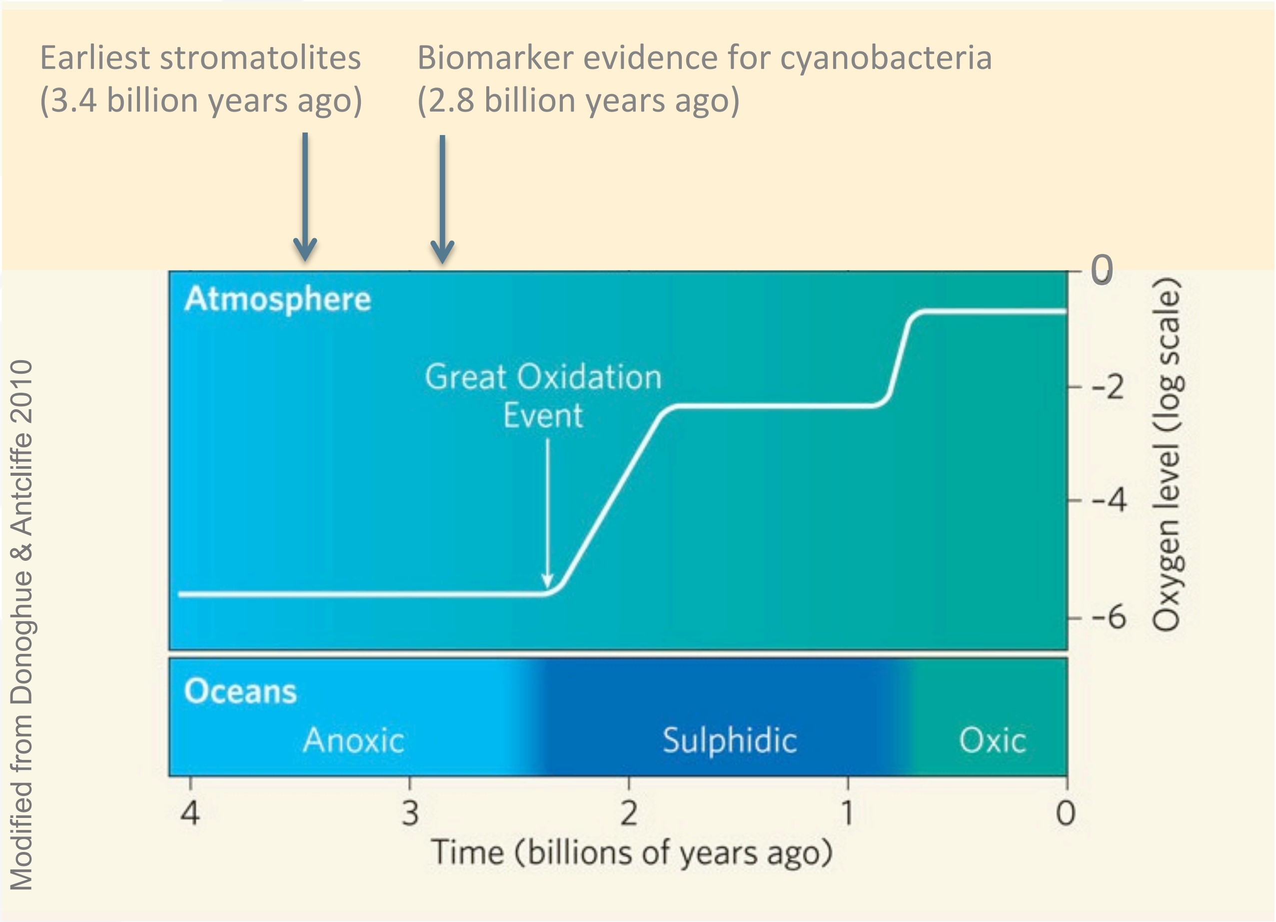
Amplified sorted single cells without 16S amplicon:

~4,000

Where else should one look?

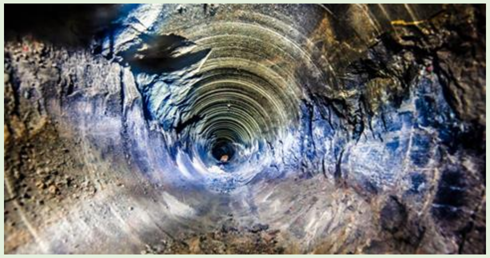
Where else should one look?

Environments mimicking that of early Earth

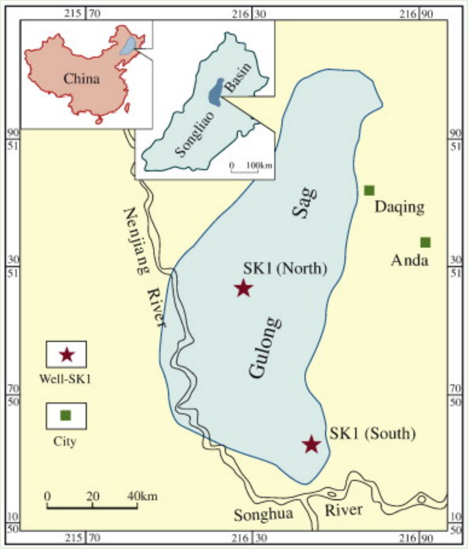


Modified from Donoghue & Antcliffe 2010

→ Deep subsurface “oxygen free” environments



Sanford Underground Research Facility (D Moser)



Songliao basin (H Dong)

- **A few words about the JGI**
  - who we are & what we do
- **Single-cell genomics**
  - the why, the how & what to expect from it

- **Single-cell science vignettes**
  - from symbionts to microbial dark matter
- **Crystal ball**
  - what the future may bring

fnx-driven
single-cell
genomics

Modified from imgkid.com

Function-driven single-cell genomics

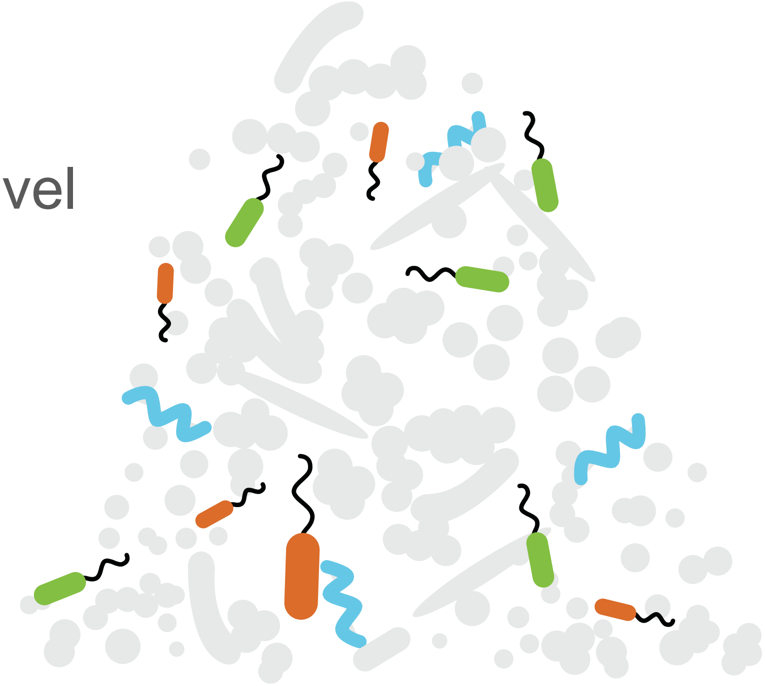
- “Next generation single-cell sequencing”:

- Identification and pre-enrichment of uncultivated environmental microbes that are involved in biogeochemical processes of interest:
  - targeted/ function-driven
  - prior to sequencing
  - without relying on any known genetic markers
  - without cultivation biases

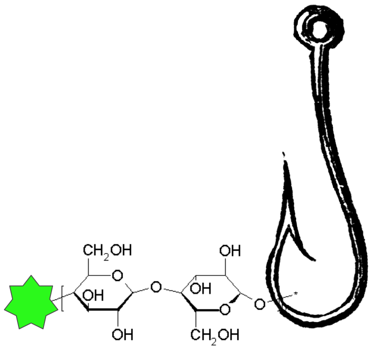
“finding the needles in the haystack”

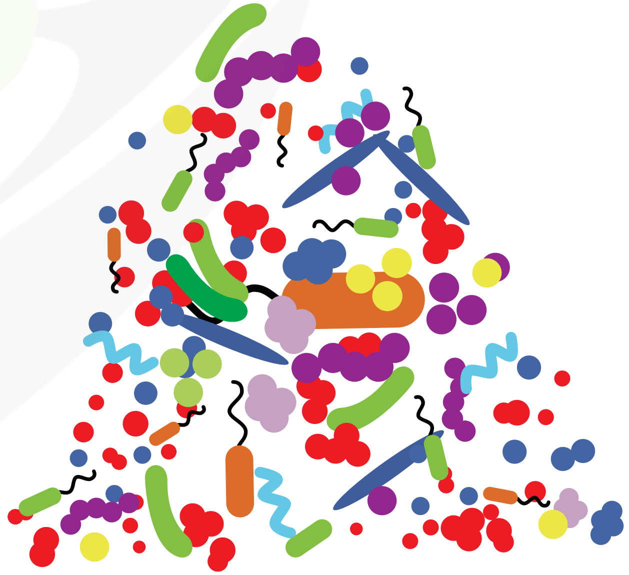
Why function-driven genomics?

- Need to move beyond sequencing yet another genome and narrow our focus on studying microbes that are involved in biogeochemical processes of interest
- Function-driven genomics is part of JGIs strategic future
- For single cells, adding a functional component is technologically highly novel (“Next-gen SCG”)



What is needed:



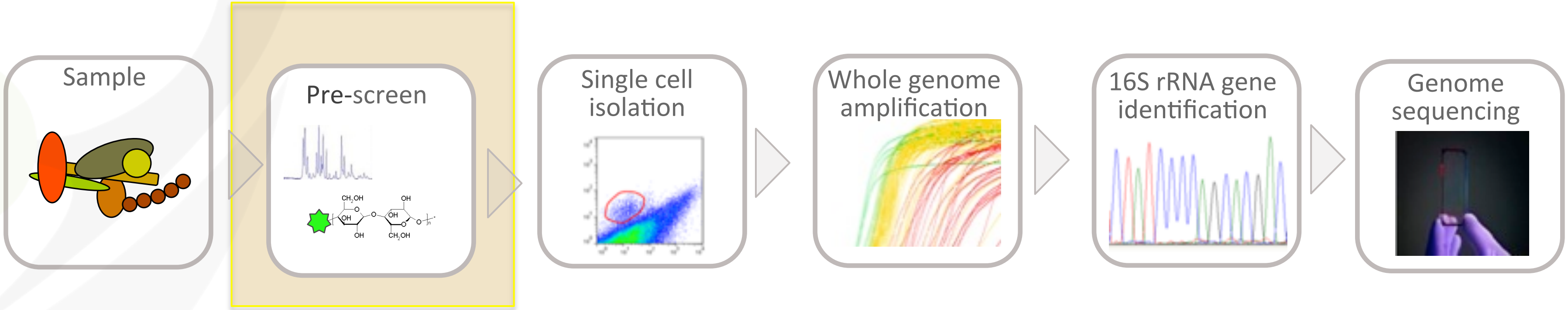


Metagenomics: high throughput, but genome context of functional genes often missing

Single-cell genomics: enables genome context, but low throughput

Neither methods is generally targeted

Function-driven single-cell genomics



Sample

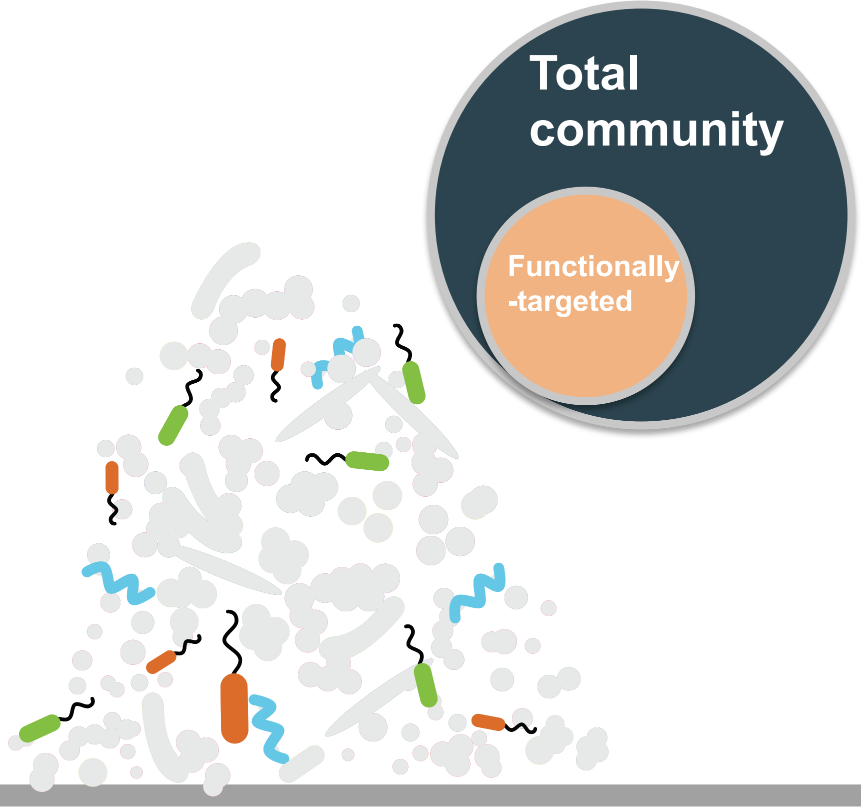
Pre-screen

Single cell isolation

Whole genome amplification

16S rRNA gene identification

Genome sequencing



Total community

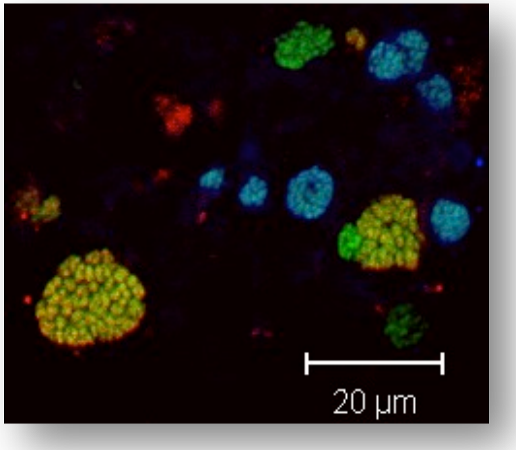
Functionally-targeted

Function-driven single-cell genomics

Two examples:

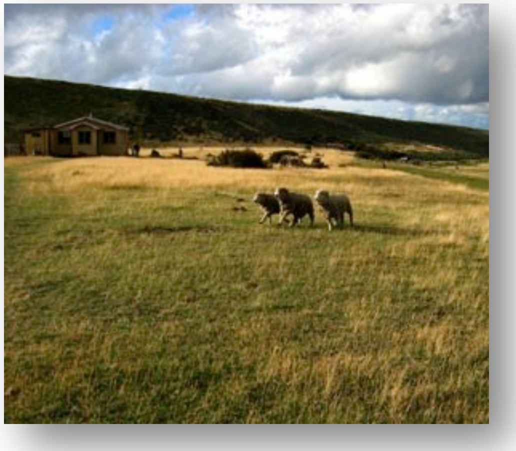
Nitrifier diversity

ETOP: Michi Wagner (Vienna)/ Roman Stocker (MIT)

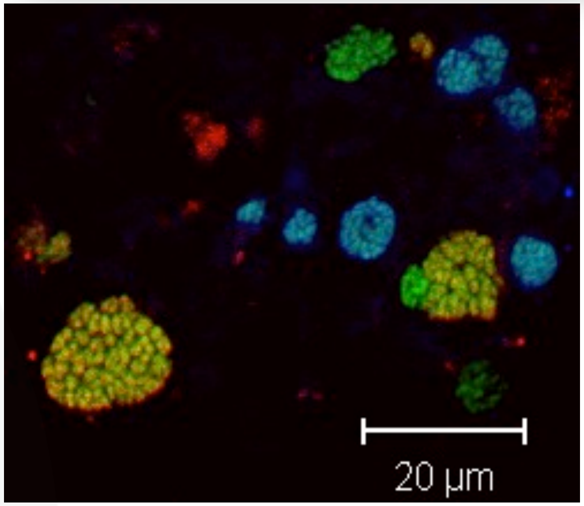


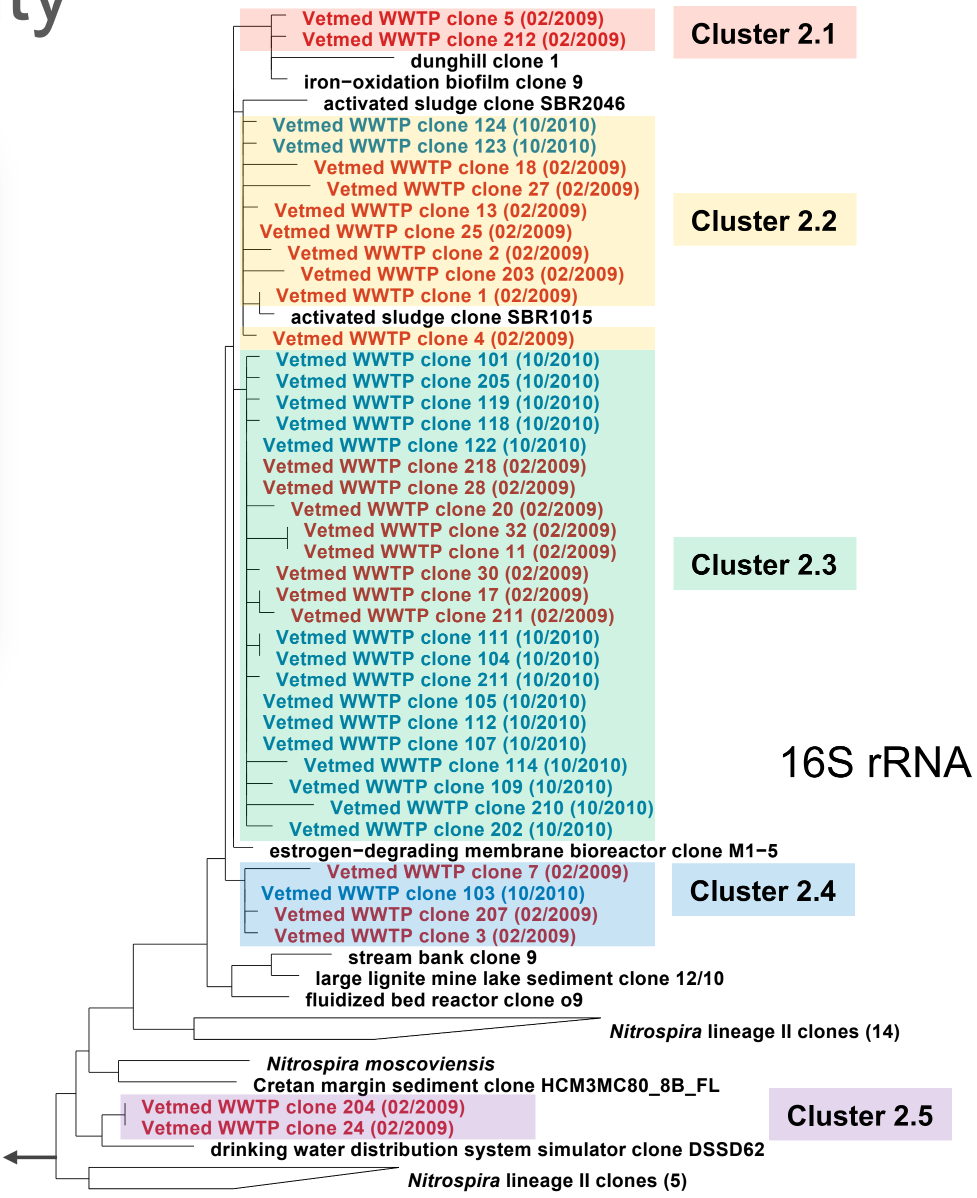
Plant carbon decomposition

LDRD



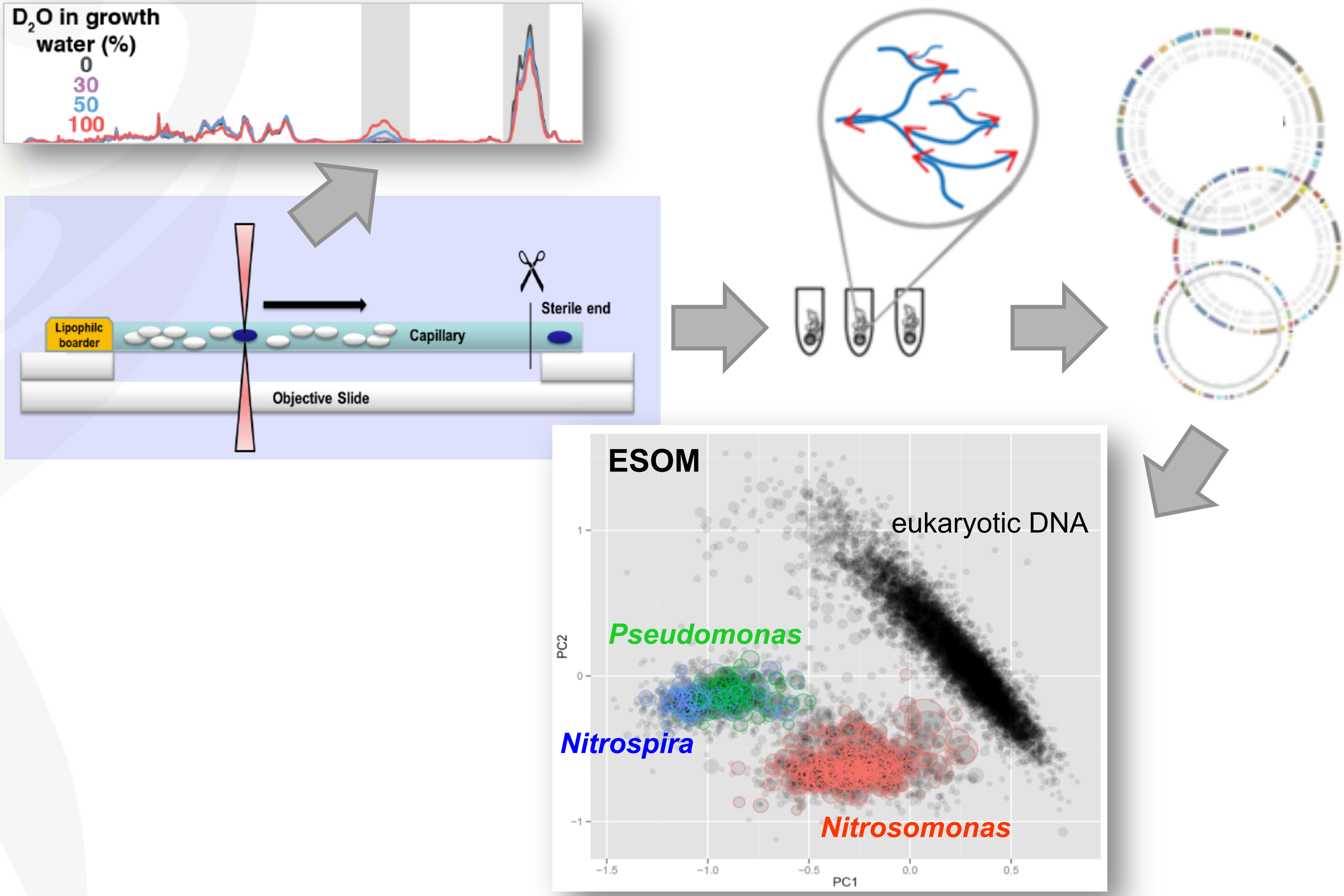
Nitrospira diversity





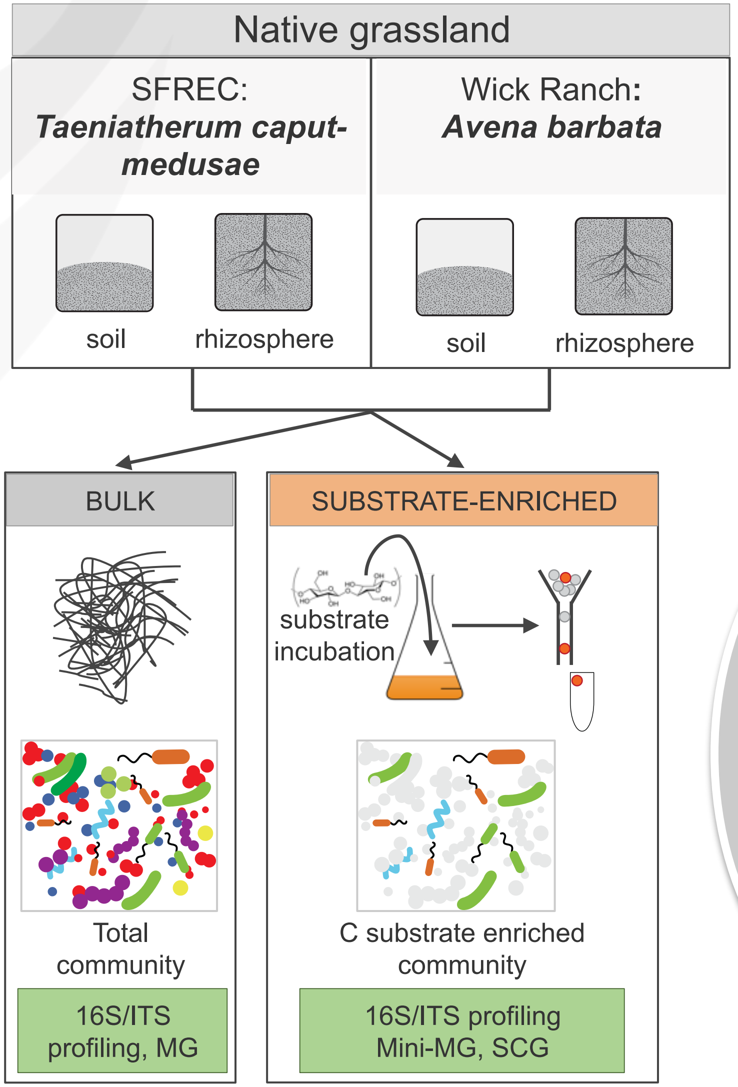
M Wagner, H Daims, TK Lee

Raman-activated colony sorting



M Wagner, H Daims, TK Lee

Plant-carbon decomposition project



Native grassland

SFREC: ***Taeniatherum caput-medusae***

soil

rhizosphere

Wick Ranch: ***Avena barbata***

soil

rhizosphere

BULK

Total community

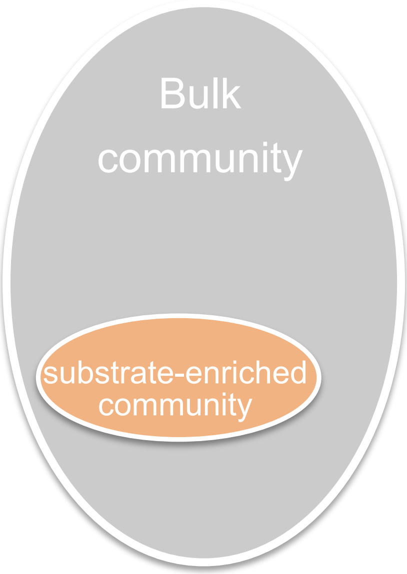
16S/ITS profiling, MG

SUBSTRATE-ENRICHED

substrate incubation

C substrate enriched community

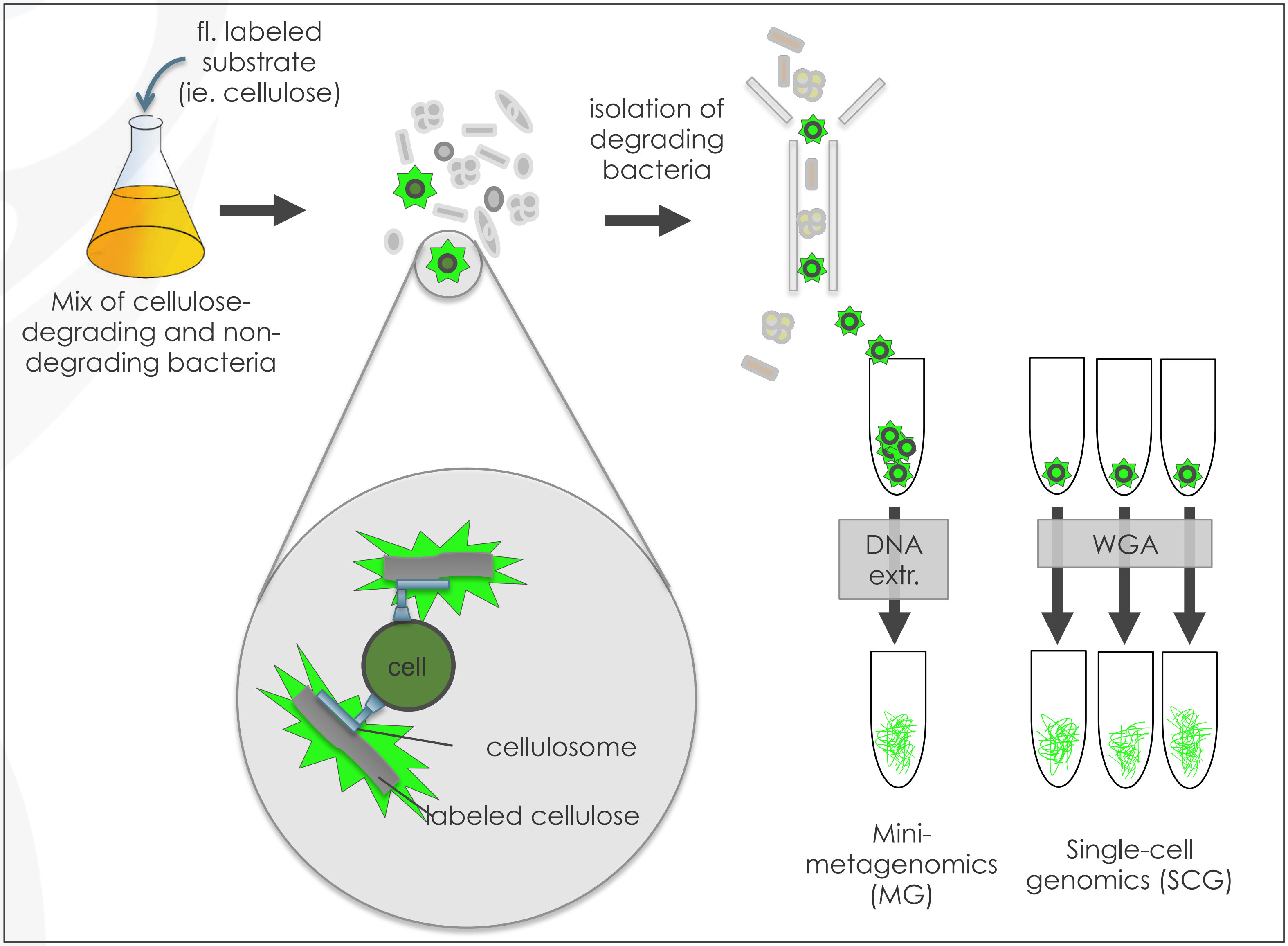
16S/ITS profiling Mini-MG, SCG



Bulk community

substrate-enriched community

Fluorescent-labeled substrate approach



Acknowledgements

Team of MDM collaborators

Phil Hugenholz (U Queensland)
Ramunas Stepanauskas (Bigelow Lab)
Jonathan Eisen (UC Davis)
Alex Sczyrba (U Bielefeld, GER)
Aaron Darling (UC Davis)
Dongying Wu (UC Davis)
Nikos Kyrpides (DOE-JGI)
Steve Hallam (U British Columbia)
Stefan Sievert (WHOI)
George Tsiamis (U Ioannina)
Wen-Tso Liu (U Illinois)
Brian Hedlund (U Nevada)

Phase II MDM collaborators

DOE-JGI

Chris Rinke
Patrick Schwientek
Natalia Ivanova
Iain Anderson
Jan-Fang Cheng
Nikos Kyrpides
Rex Malmstrom
Stephanie Malfatti
Susannah Tringe
Scott Clingenpeel
Janey Lee
Alicia Clum
Alex Copeland
Rubin, Bristow, Pennacchio Labs
Kyrpides & Markovitz Groups (IMG team)

Department of Energy

Dan Drell

Thank you!

Questions?