

# TEAM G-KNOME:

## What makes a vibrio a vibrio?

---



Hannes: The Core Master

Chiharu: The Pathway Master

Leah: The Tree Master

Stefan: The Iron Master

Maeva: MasterBLAST (aka Pie Master)







# Definition of the core genome

## Reference genome: *Vibrio splendidus*

Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
⋮				
gene 22	50	24	...	24
⋮				
gene 3971	87	85	...	85

% identity, data of full genomes from IMG database

# Definition of the core genome

## Reference genome: *Vibrio splendidus*

Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
⋮				
gene 22	50	24	...	24
⋮				
gene 3971	87	85	...	85

% identity, data of full genomes from IMG database

Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	1	1	...	1
⋮				
gene 2	1	1	...	1
⋮				
gene 22	1	0	...	0
⋮				
gene 3971	1	1	...	1

threshold identity = 25 %



# Definition of the core genome

## Reference genome: *Vibrio splendidus*

Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
⋮				
gene 22	50	24	...	24
⋮				
gene 3971	87	85	...	85

% identity, data of full genomes from IMG database

Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236	
gene 1	1	1	...	1	core
⋮					
gene 2	1	1	...	1	core
⋮					
gene 22	1	0	...	0	
⋮					
gene 3971	1	1	...	1	core

threshold identity = 25 %

# Definition of the core genome

## Reference genome: *Vibrio splendidus*

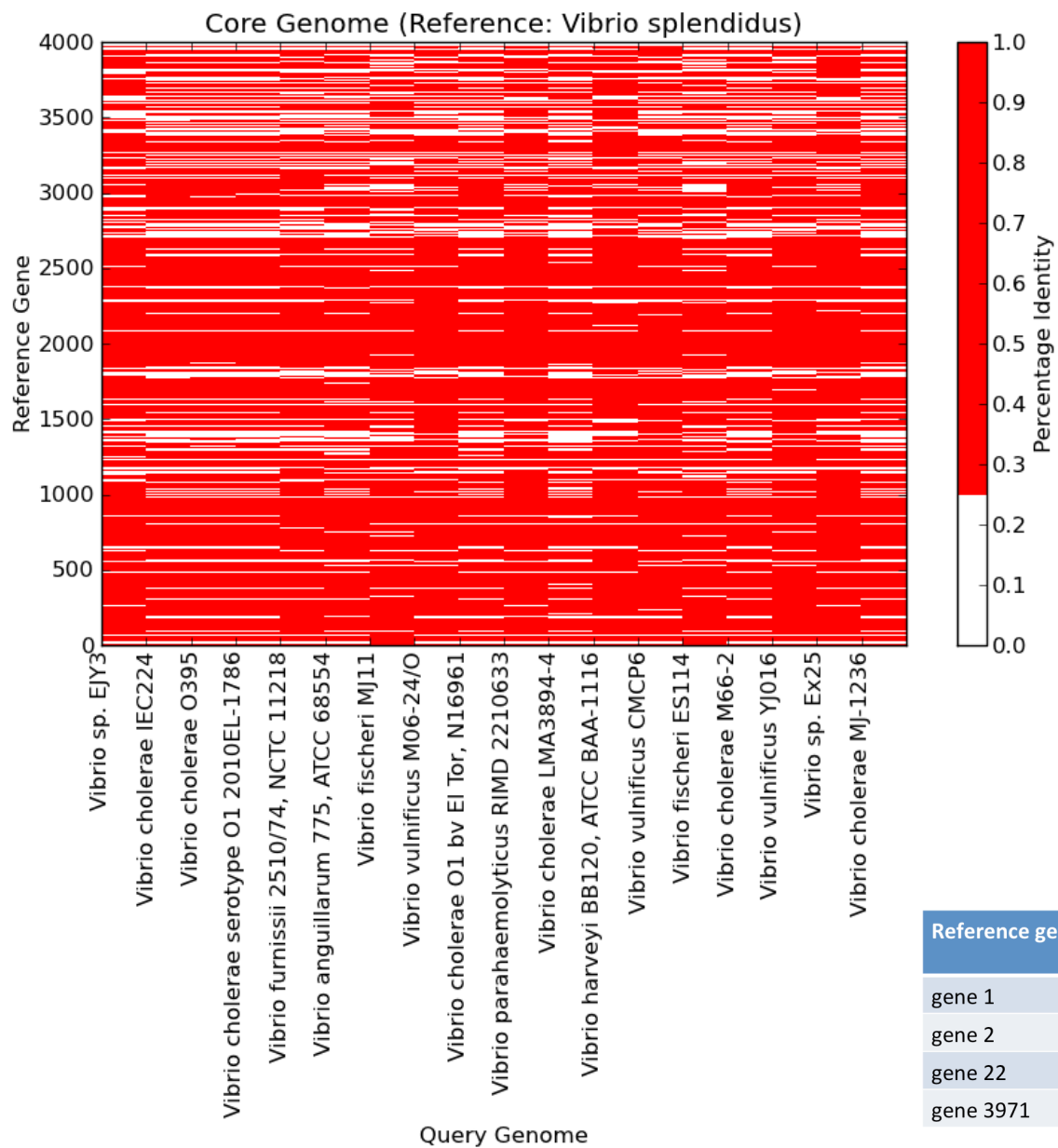
Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
⋮				
gene 22	50	24	...	24
⋮				
gene 3971	87	85	...	85

% identity, data of full genomes from IMG database

Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236	
gene 1	1	1	...	1	core
gene 2	1	1	...	1	core
⋮					
gene 22	1	0	...	0	
⋮					
gene 3971	1	1	...	1	core

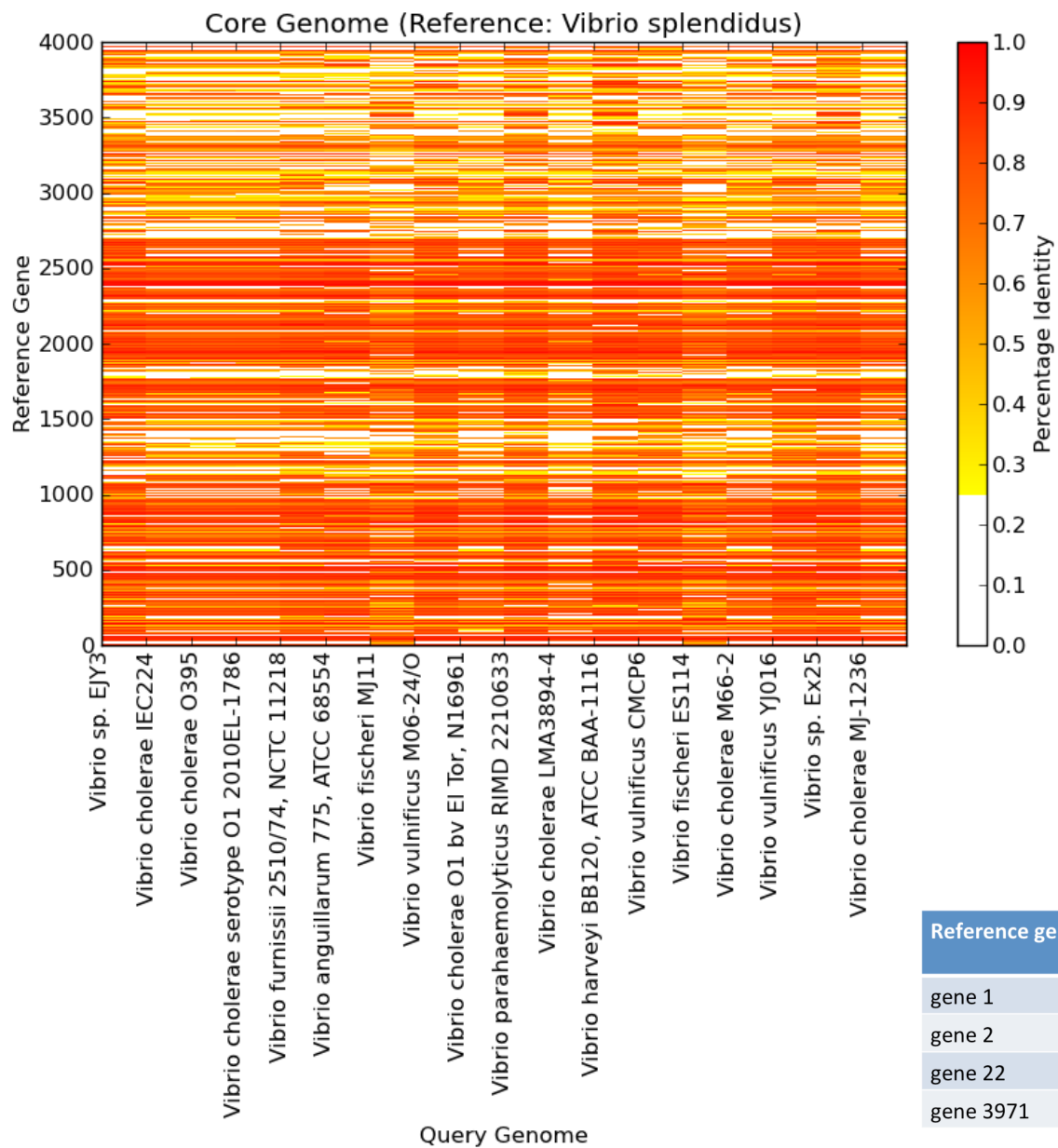
core genome

threshold identity = 25 %

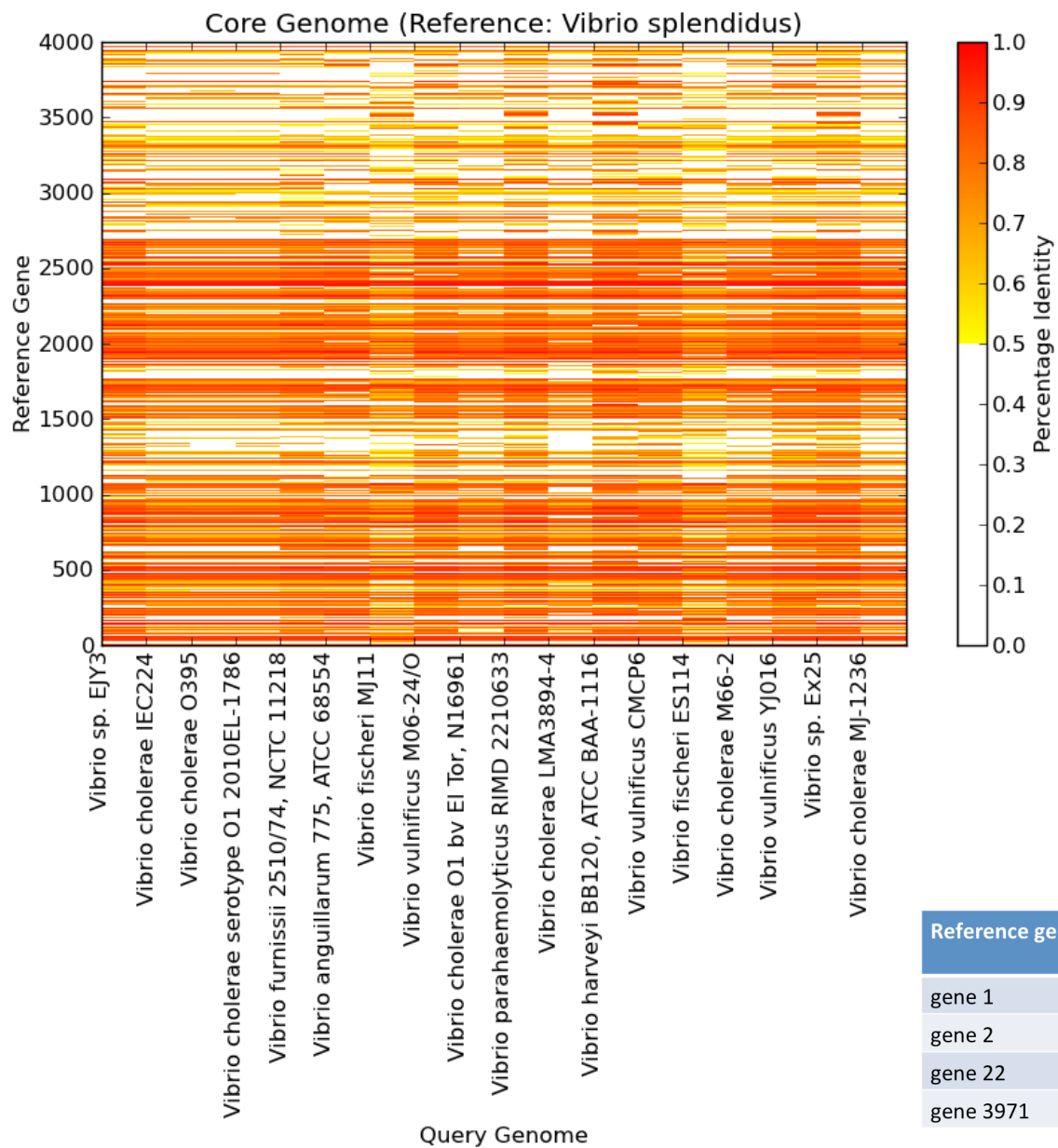


Reference gene	<i>Vibrio</i> sp. EYJ3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	1	1	...	1
gene 2	1	1	...	1
gene 22	1	0	...	0
gene 3971	1	1	...	1

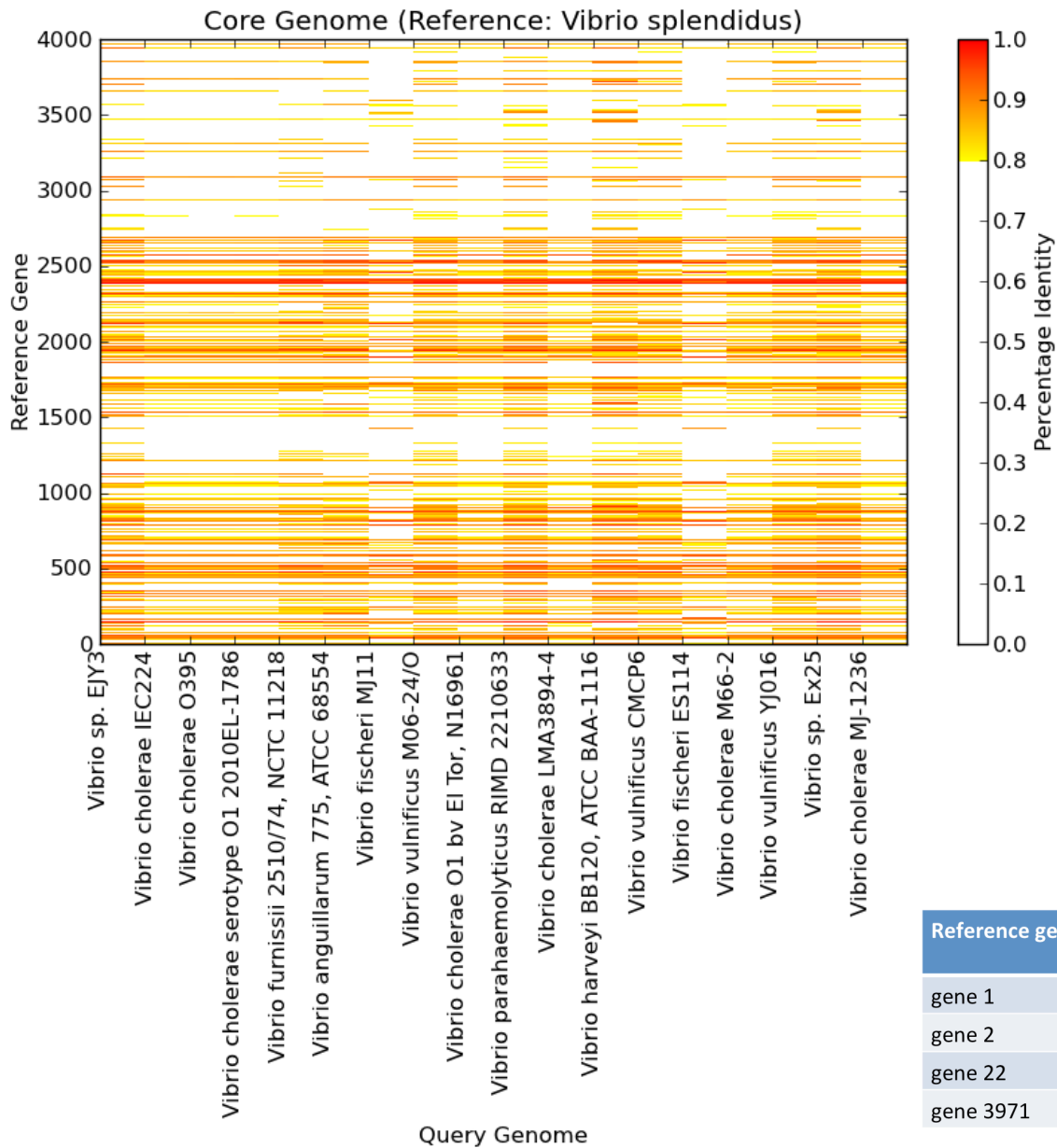




Reference gene	<i>Vibrio</i> sp. EYJ3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85



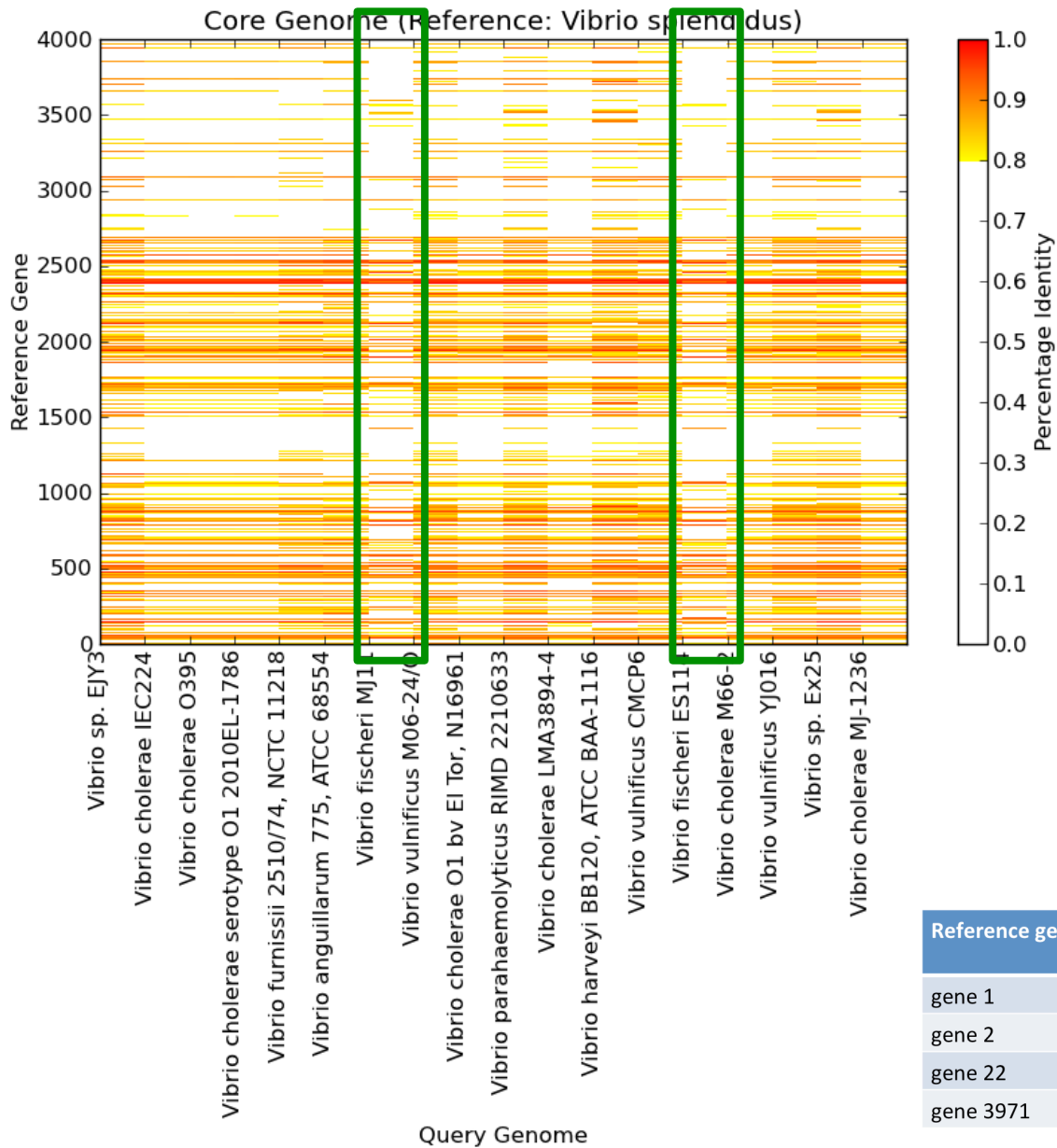
Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85



- Chiharu: genetic markers for different environments
- Leah: tree diagrams
- Stefan: iron uptake and proteins containing iron
- Maeva: flexible genome

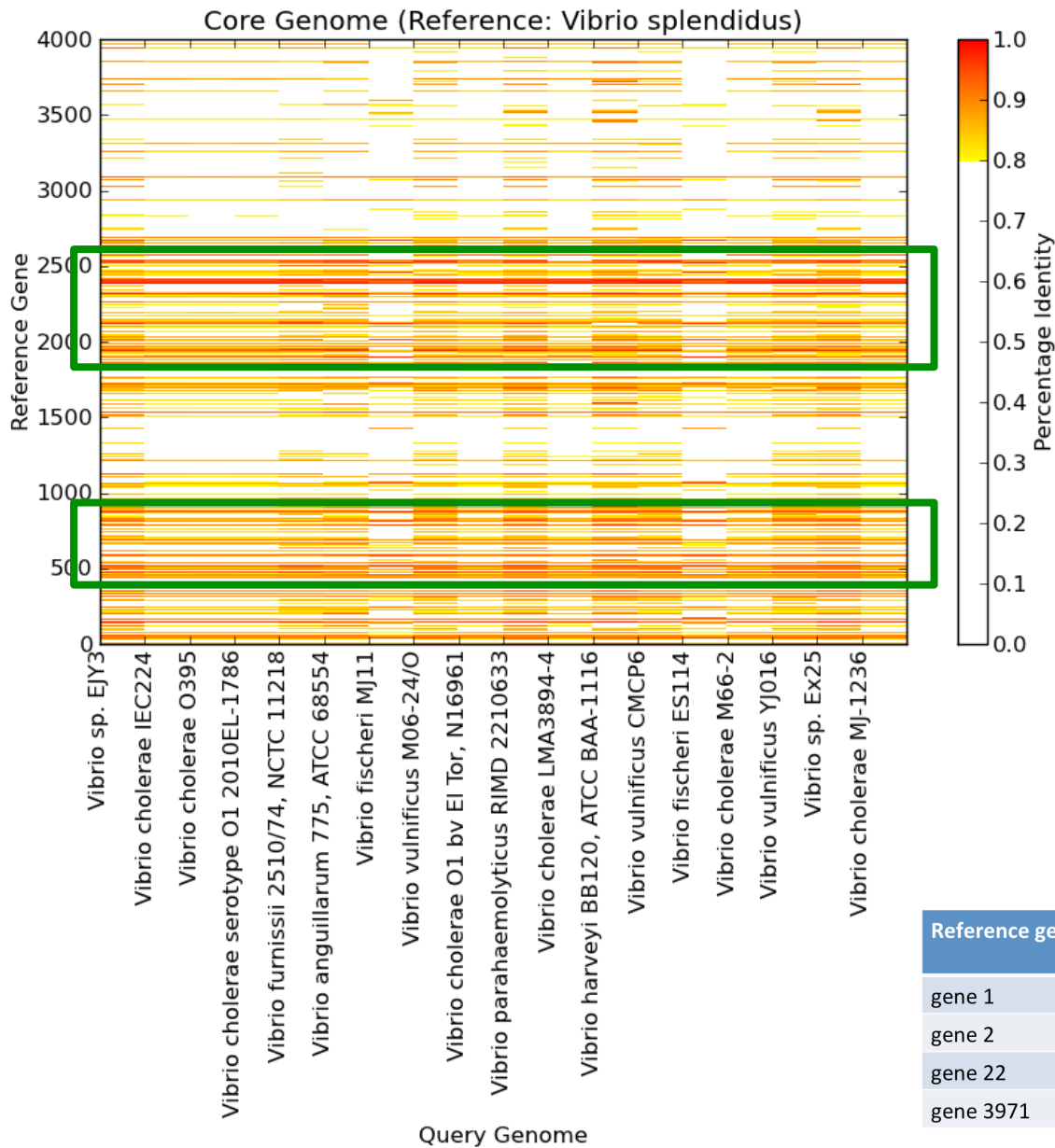
Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85





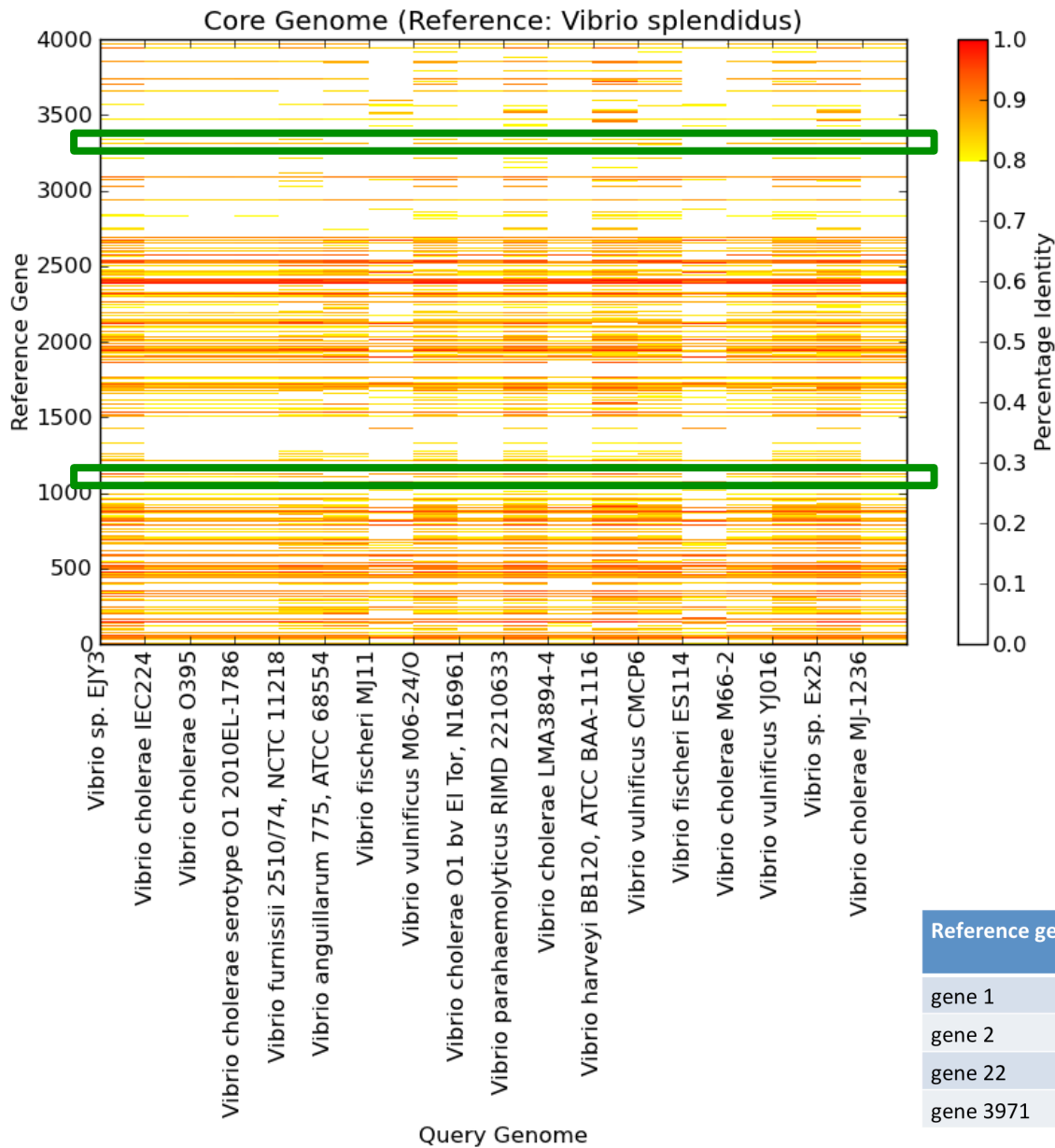
- Chiharu: genetic markers for different environments
- Leah: tree diagrams
- Stefan: iron uptake and proteins containing iron
- Maeva: flexible genome

Reference gene	<i>Vibrio</i> sp. EYJ3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85



- Chiharu: genetic markers for environments
- Leah: tree diagrams
- Stefan: iron uptake and proteins containing iron
- Maeva: flexible genome

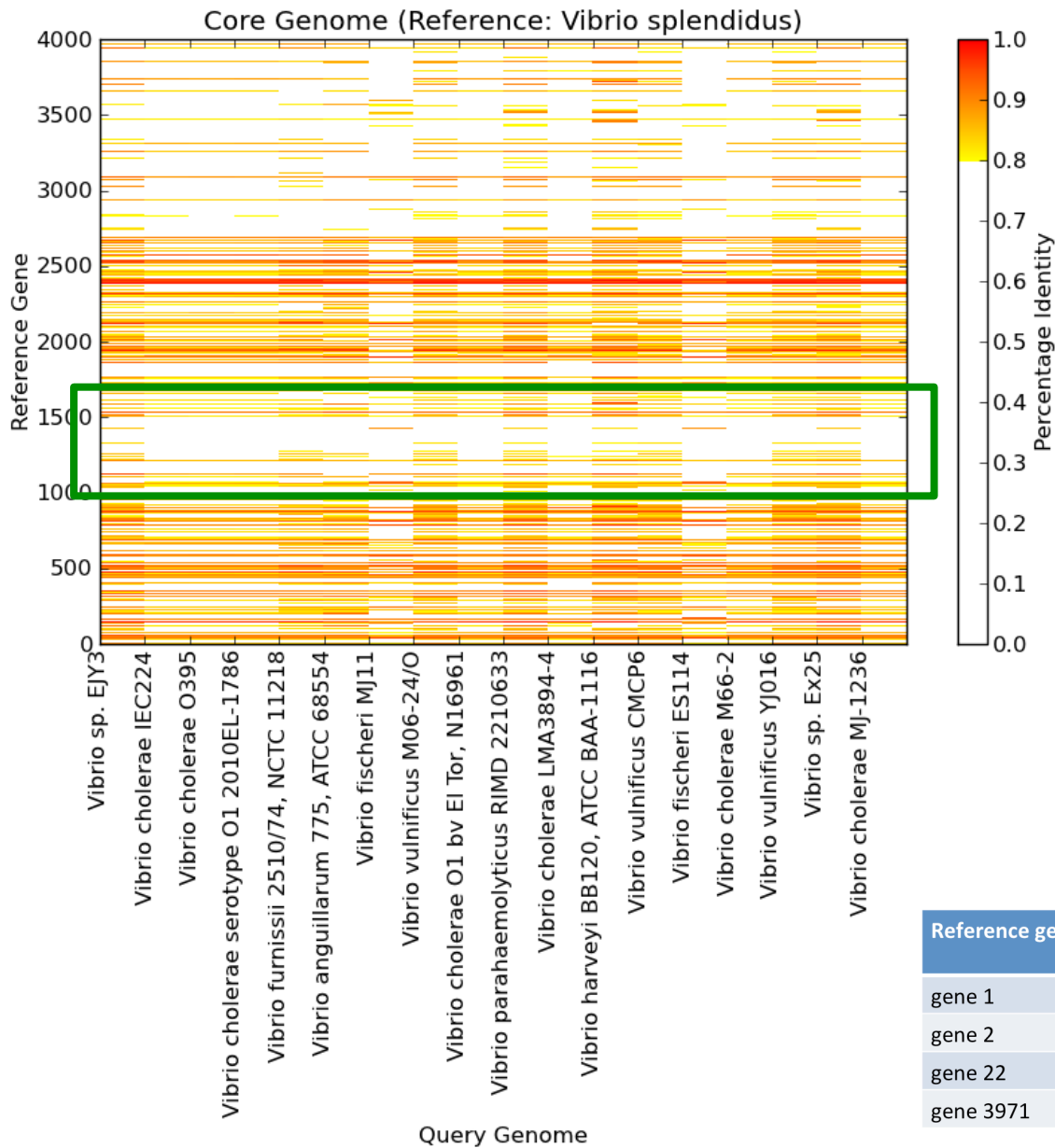
Reference gene	<i>Vibrio</i> sp. EYJ3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85



- Chiharu: genetic markers for environments
- Leah: tree diagrams
- **Stefan: iron uptake and proteins containing iron**
- Maeva: flexible genome

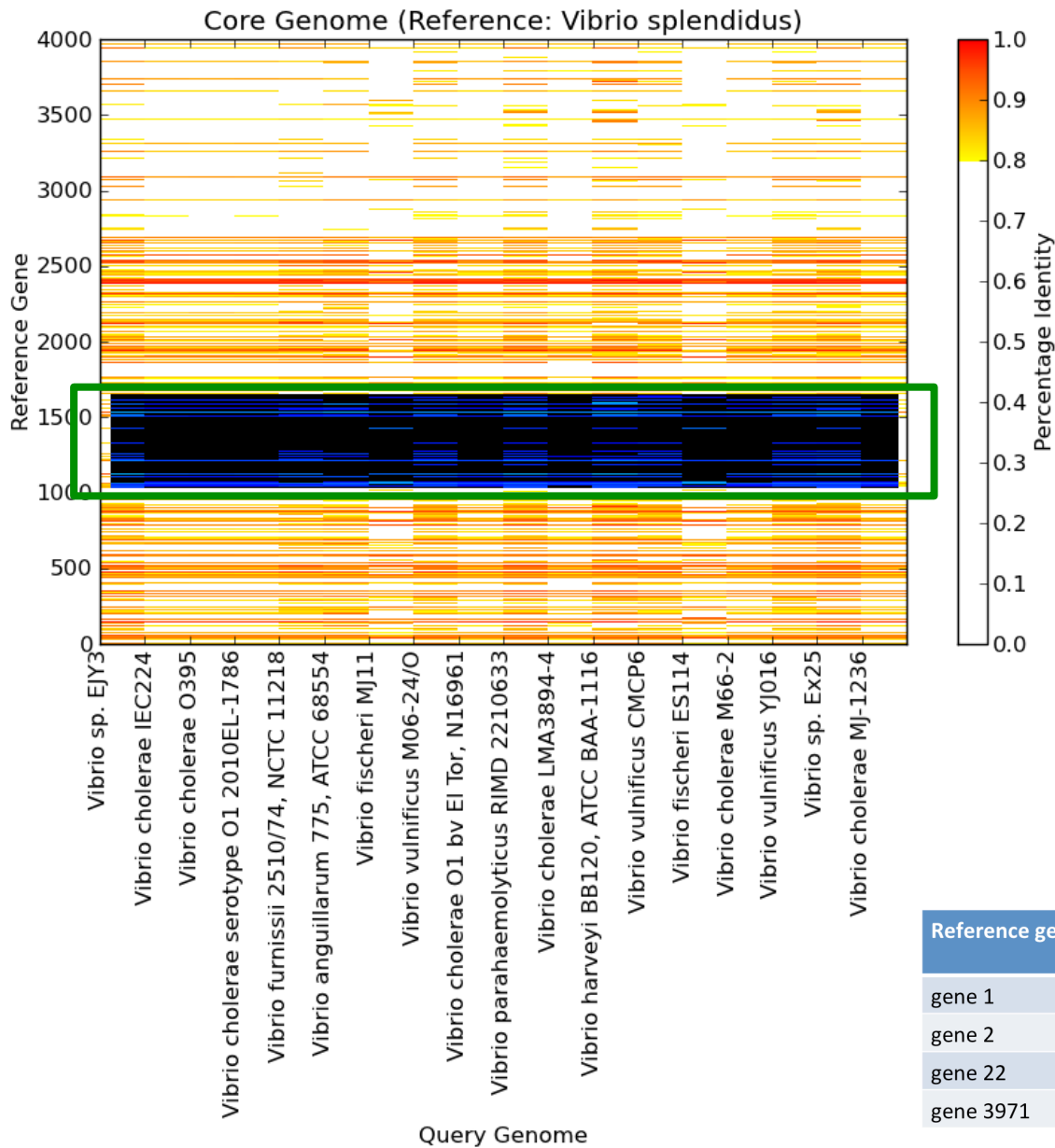
Reference gene	<i>Vibrio</i> sp. EYJ3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85





- Chiharu: genetic markers for environments
- Leah: tree diagrams
- Stefan: iron uptake and proteins containing iron
- **Maeva: flexible genome**

Reference gene	<i>Vibrio</i> sp. EJY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85

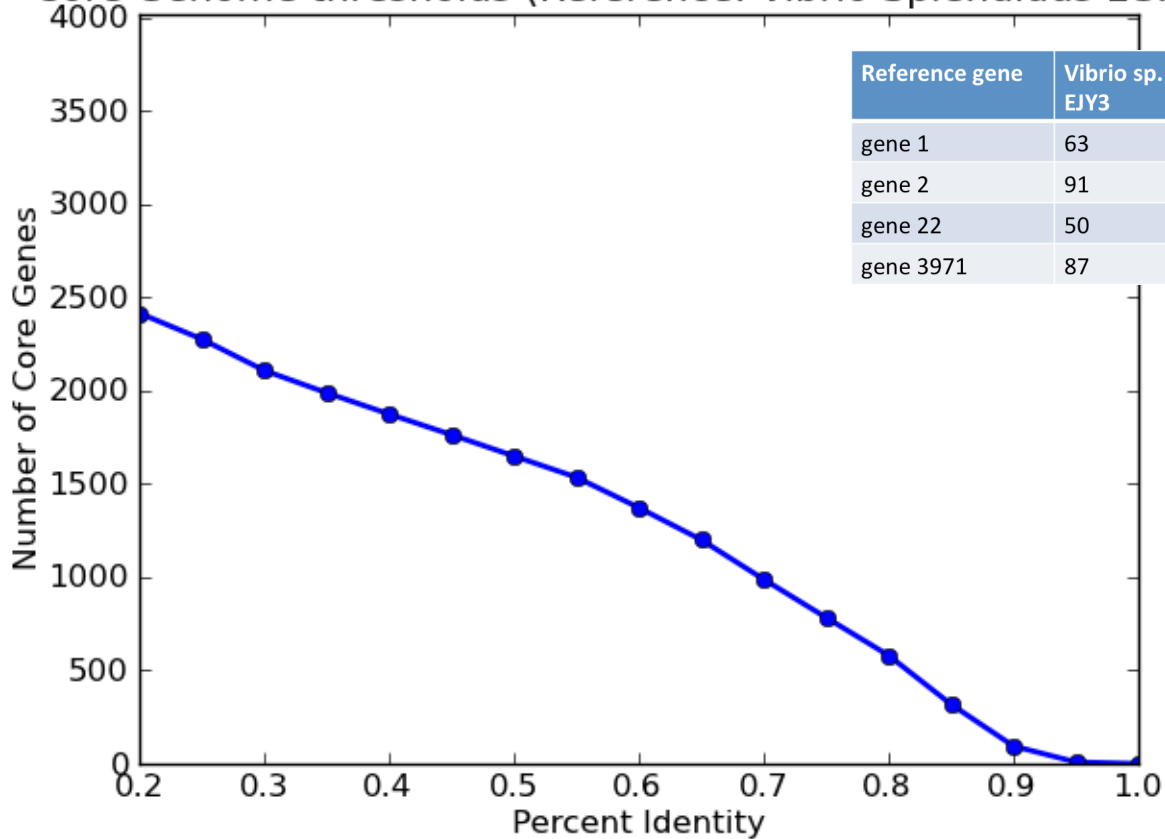


- Chiharu: genetic markers for environments
- Leah: tree diagrams
- Stefan: iron uptake and proteins containing iron
- **Maeva: flexible genome**

Reference gene	<i>Vibrio</i> sp. EYJ3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85

# Core genome: How big are you?

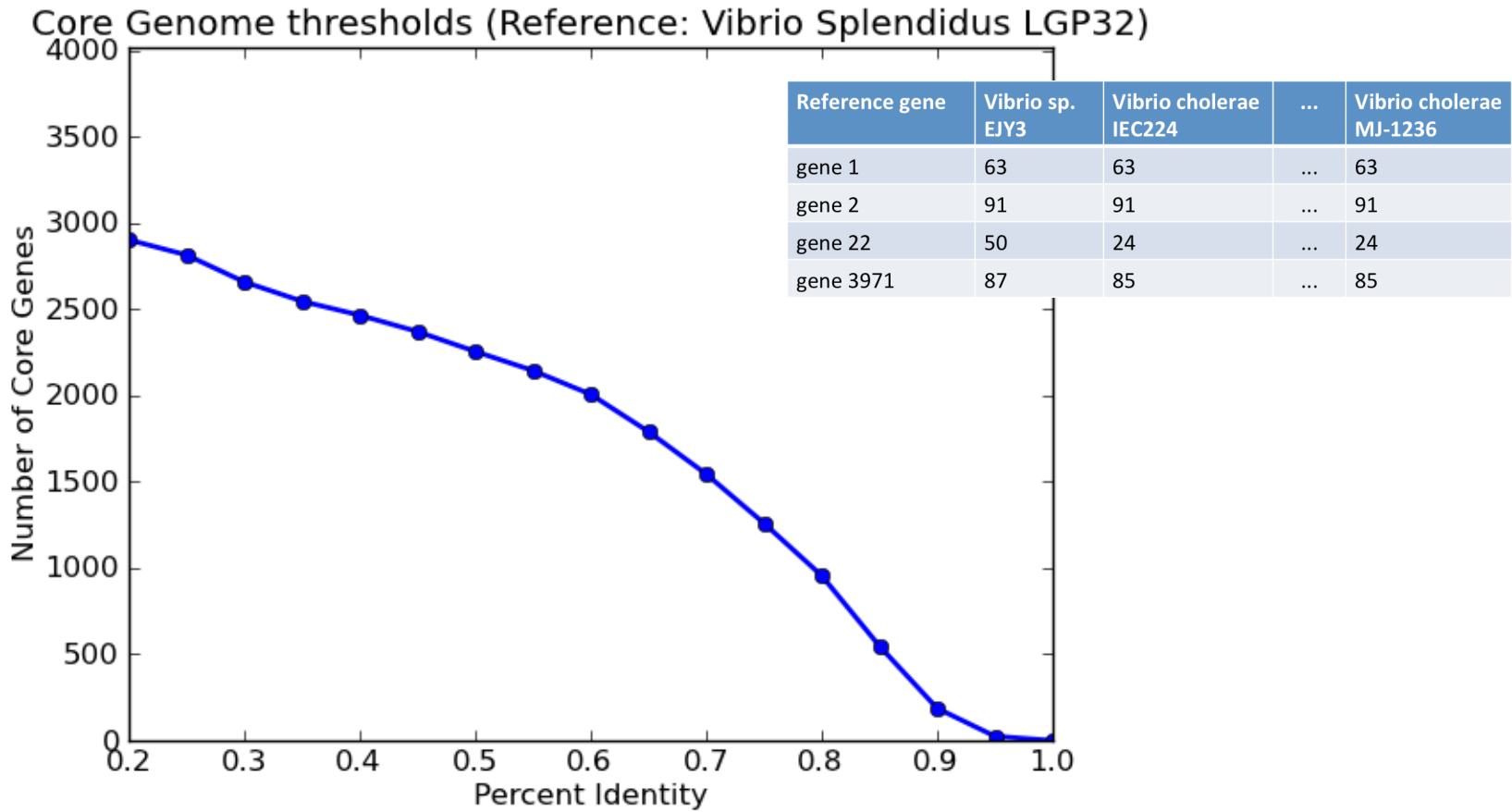
Core Genome thresholds (Reference: *Vibrio Splendidus* LGP32)



Reference gene	<i>Vibrio</i> sp. EY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85

% Core Match  100.00

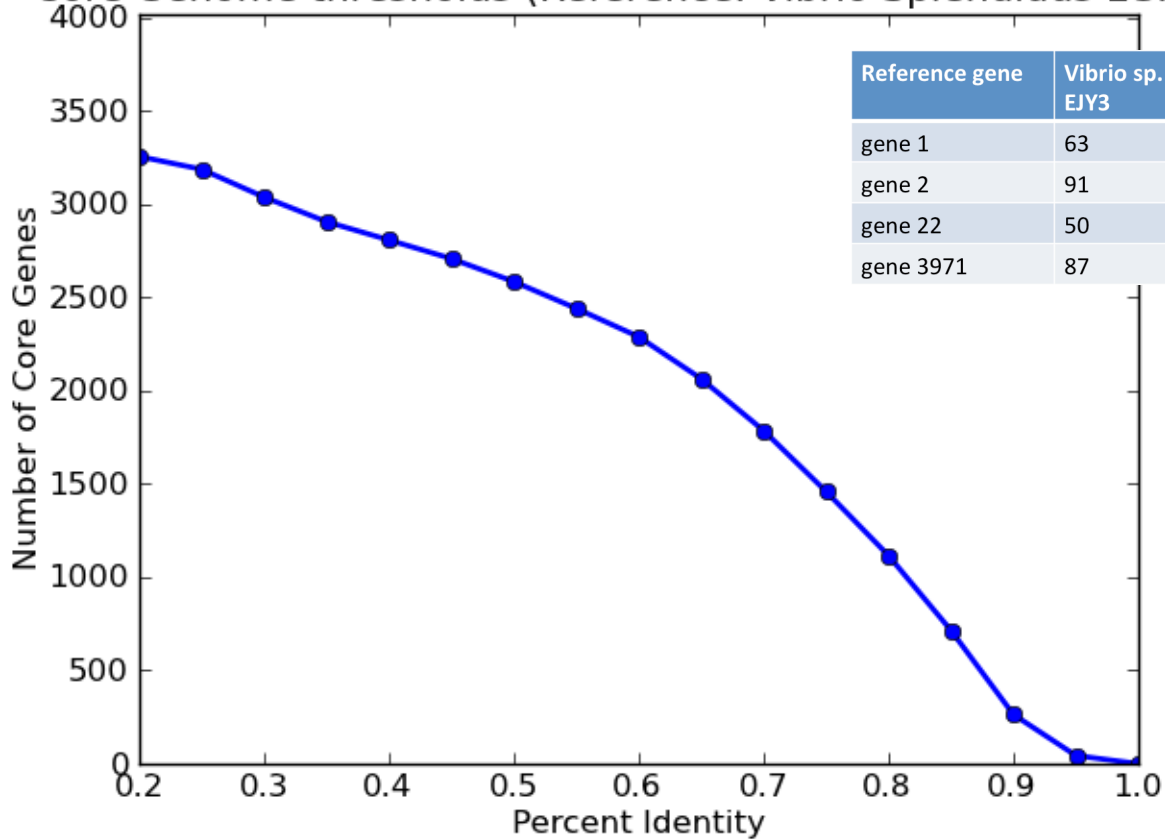
# Core genome: How big are you?



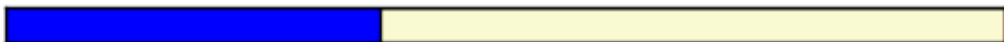
% Core Match  85.00

# Core genome: How big are you?

Core Genome thresholds (Reference: *Vibrio Splendidus* LGP32)



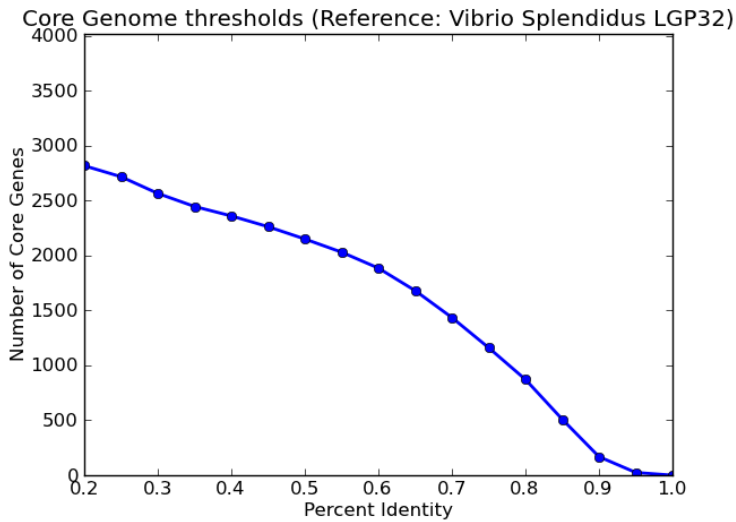
Reference gene	<i>Vibrio</i> sp. EY3	<i>Vibrio cholerae</i> IEC224	...	<i>Vibrio cholerae</i> MJ-1236
gene 1	63	63	...	63
gene 2	91	91	...	91
gene 22	50	24	...	24
gene 3971	87	85	...	85

% Core Match  50.00



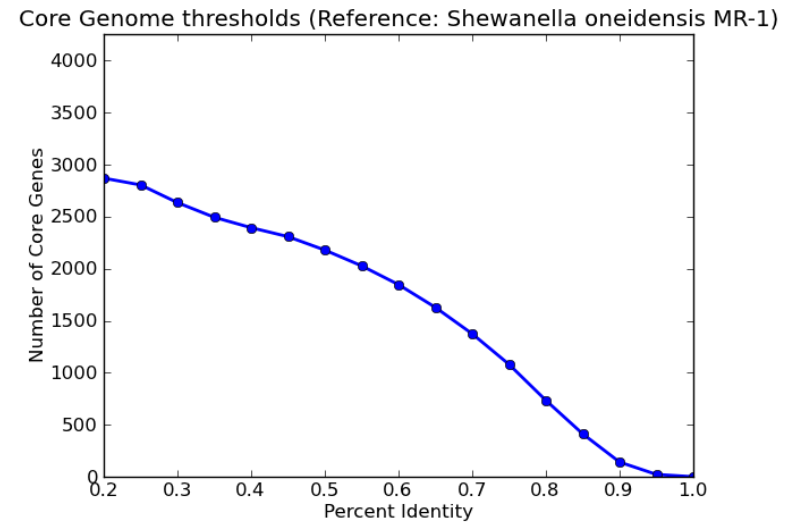
# What makes a Vibrio a Vibrio?

## Vibrio splendidus



% Core Match  90.00

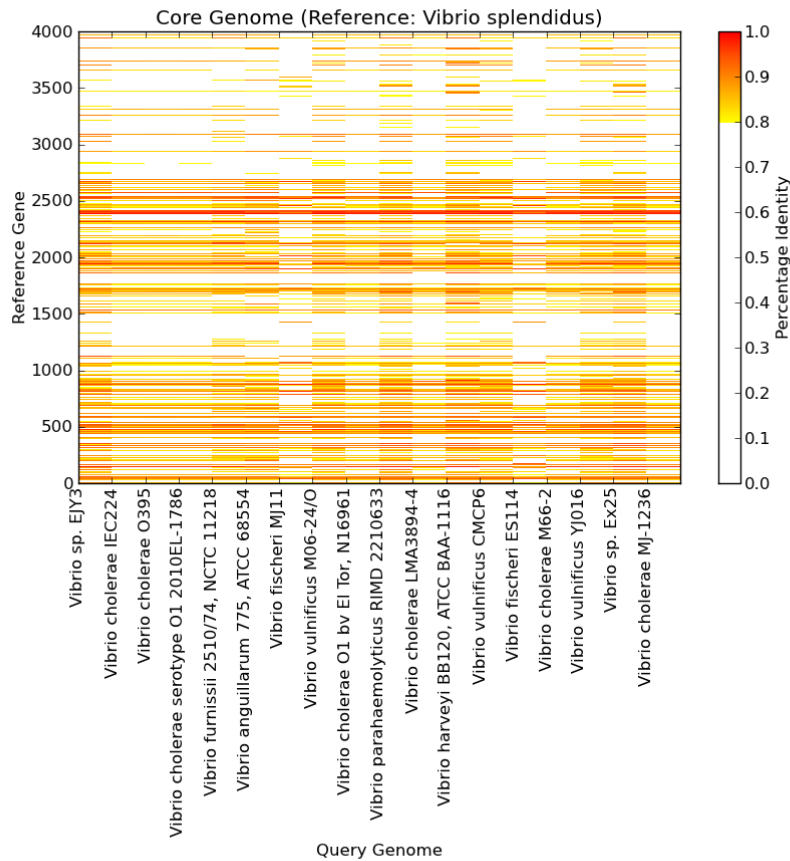
## Shewanella oneidensis



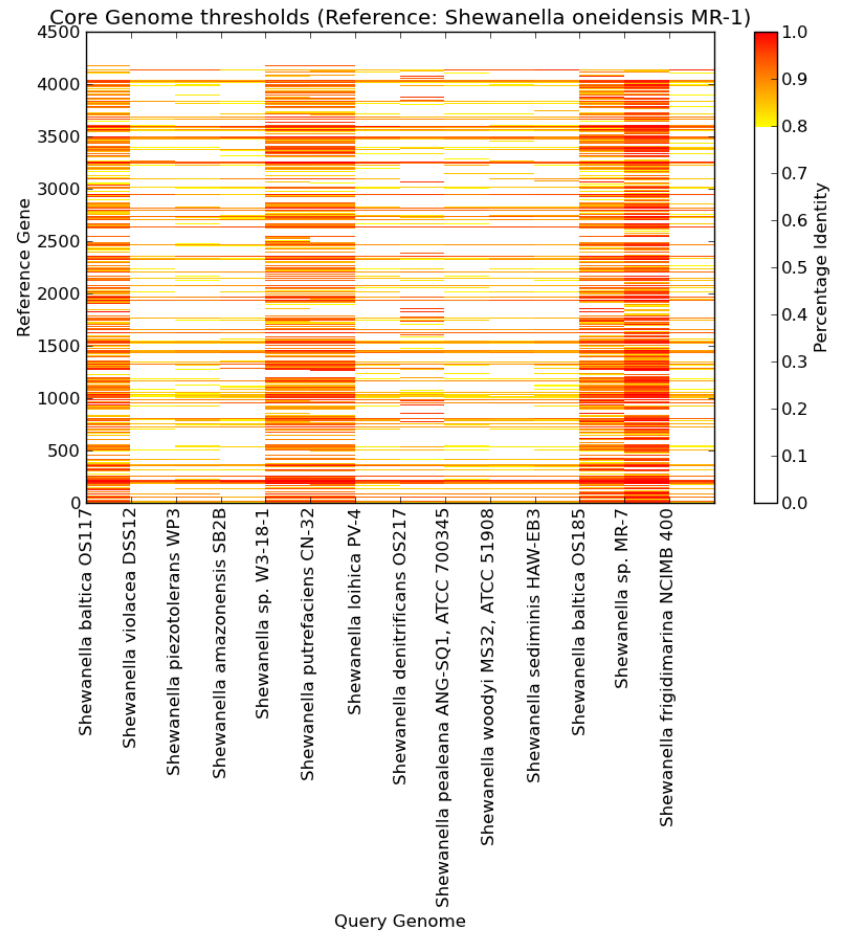
% Core Match  90.00

# What makes a Vibrio a Vibrio?

## Vibrio splendidus



## Shewanella oneidensis



My question

What is the difference of Vibrio between their locations ??

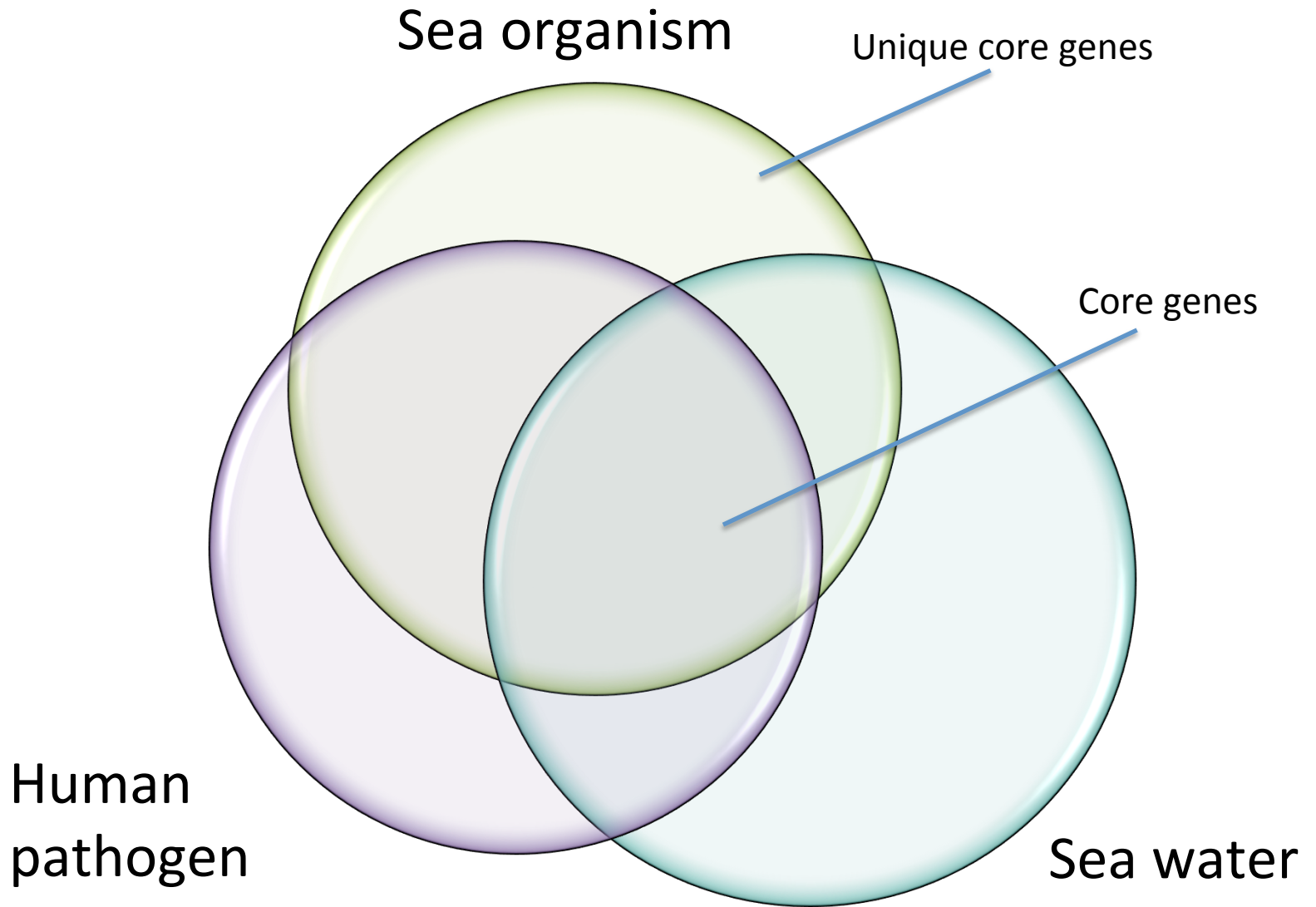
Proposal Name	category
Vibrio cholerae IEC224	H
Vibrio cholerae LMA3894-4	H
Vibrio cholerae M66-2	H
Vibrio cholerae MJ-1236	H
Vibrio cholerae O1 bv El Tor, N16961	H
Vibrio cholerae O395	H
Vibrio cholerae O395	H
Vibrio cholerae serotype O1 2010EL-1786	H
Vibrio furnissii 2510/74, NCTC 11218	H
Vibrio parahaemolyticus RIMD 2210633	H
Vibrio vulnificus CMCP6	H
Vibrio vulnificus M06-24/O	H
Vibrio vulnificus YJ016	H
Vibrio anguillarum 775	SO
Vibrio fischeri ES114	SO
Vibrio fischeri MJ11	SO
Vibrio splendidus LGP32	SO
Vibrio harveyi BB120, ATCC BAA-1116	SW
Vibrio sp. EJY3	SW
Vibrio sp. Ex25	SW

Human pathogen vs  
sea organism's symbiote vs  
free living in sea water

Hypothesis

- Does they have category  
unique genes?

→ Does the unique genes gives  
benefits to live their own  
location ?



# Methods



**img** INTEGRATED MICROBIAL GENOME

On July 23, 2013 time TBD Our web servers will undergo maintenance. All IMG systems will be down for an hour. Sorry for the inconvenience. [View Updates](#)

**NEW 4** **IMG 4 Data Management**

The mission of the Integrated Microbial Genomes (IMG) system is to support the annotation, analysis, and distribution of genome and metagenome datasets, as long as they agree with the [IMG data release policy](#) and follow the [IMG submission site](#).

**Note to Users - Posted June 28, 2013**

We are committed to support scientists worldwide with their genome and metagenome data analysis needs. However, we do not support for non-JGI users, that is, users who do not use JGI for sequencing.

We will continue to accept genome and metagenome datasets for annotation and integration into IMG and rely increasingly on our data release policy and metadata requirements will be strictly enforced and our ability to answer questions by email will be limited.

**Data distribution** for genomes and metagenomes is provided solely through individual genome and metagenome data port services; no other type of data distribution (data download) is provided.

**IMG Data Warehouse**

**IMG data warehouse** integrates genome and metagenome datasets sequenced at JGI or provided by IMG users with a comprehensive set of functional annotations, and data quality, in terms of the coherence of annotations. About 300 samples have been sequenced at DOE's Joint Genome Institute.

Through recent updates, IMG aims at providing high levels of **genome diversity**, in terms of the number of genomes in terms of the number of functional annotations, and **data quality**, in terms of the coherence of annotations. About 300 samples have been sequenced at DOE's Joint Genome Institute.

As of March 2013, IMG has about 3,100 registered users from 66 countries across North America (57%), Europe (20%), South America (9%), Oceania (4.5%), and Africa (3%).

[Register](#)

Microbial Genomes Systems Metagenome Analysis Systems

**Gene Comparison**

Reference Locus Tag	Query Gene Name	Query Locus Tag	Query Genome	Percent Identity	Bit Score	E-value
VIBHAR_00259	ATPase	VEJY3_14845	Vibrio sp. EJY3	73.3	48.9	7.40E-07
VIBHAR_00259	hypothetical protein	VS_2997	Vibrio splendidu	73.1	42.4	6.40E-05
VIBHAR_00259	predicted ATPase	vfu_A00487	Vibrio furnissii 2	69.2	42	8.20E-05
VIBHAR_00259	ATP-binding region	VAA_00489	Vibrio anguillaru	82.6	41.2	0.00012
VIBHAR_00259	conserved domain pr	VFMJ11_2558	Vibrio fischeri M	73.1	43.5	2.60E-05
VIBHAR_00259	hypothetical protein	VVM_00280	Vibrio vulnificus	39.7	43.5	2.90E-05
VIBHAR_00259	hypothetical protein	VP2937	Vibrio parahaem	56.6	57.8	1.50E-09
VIBHAR_00259	hypothetical protein	VF2434	Vibrio fischeri E	73.1	43.9	1.90E-05
VIBHAR_00259	predicted ATPase	VV0167	Vibrio vulnificus	39.7	43.1	4.10E-05

# Methods



Make the gene list of unique core

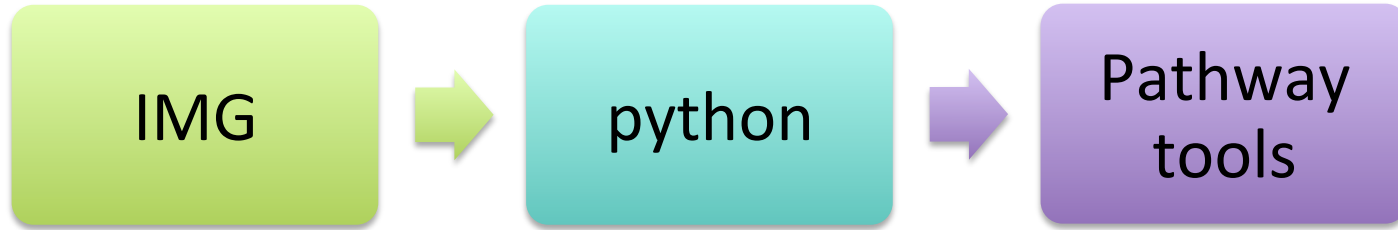
REFERENCE C	Vibrio harveyi	Vibrio fischeri	Vibrio vulnificus	Vibrio anguillar	Vibrio cholerae	Vibrio cholerae	Vibrio fischeri	Vibrio sp. EJY3	Vibrio furnissii	Vibrio cholerae
VV0001	1	1	1	1	1	1	1	1	1	1
VV0002	1	1	1	1	0	1	0	1	0	1
VV0003	1	1	0	0	0	1	0	0	1	1



Reference Locus Tag	Query Gene Name	Query Locus Tag	Query Genome	Percent Identity	Bit Score	E-value
VIBHAR_00259	ATPase	VEJY3_14845	Vibrio sp. EJY3	73.3	48.9	7.40E-07
VIBHAR_00259	hypothetical protein	VS_2997	Vibrio splendidi	73.1	42.4	6.40E-05
VIBHAR_00259	predicted ATPase	vfu_A00487	Vibrio furnissii 2	69.2	42	8.20E-05
VIBHAR_00259	ATP-binding region	VAA_00489	Vibrio anguillar	82.6	41.2	0.00012
VIBHAR_00259	conserved domain pr	VFMJ11_2558	Vibrio fischeri M	73.1	43.5	2.60E-05
VIBHAR_00259	hypothetical protein	VVM_00280	Vibrio vulnificus	39.7	43.5	2.90E-05
VIBHAR_00259	hypothetical protein	VP2937	Vibrio parahaem	56.6	57.8	1.50E-09
VIBHAR_00259	hypothetical protein	VF2434	Vibrio fischeri E	73.1	43.9	1.90E-05
VIBHAR_00259	predicted ATPase	VV0167	Vibrio vulnificus	39.7	43.1	4.10E-05



# Methods



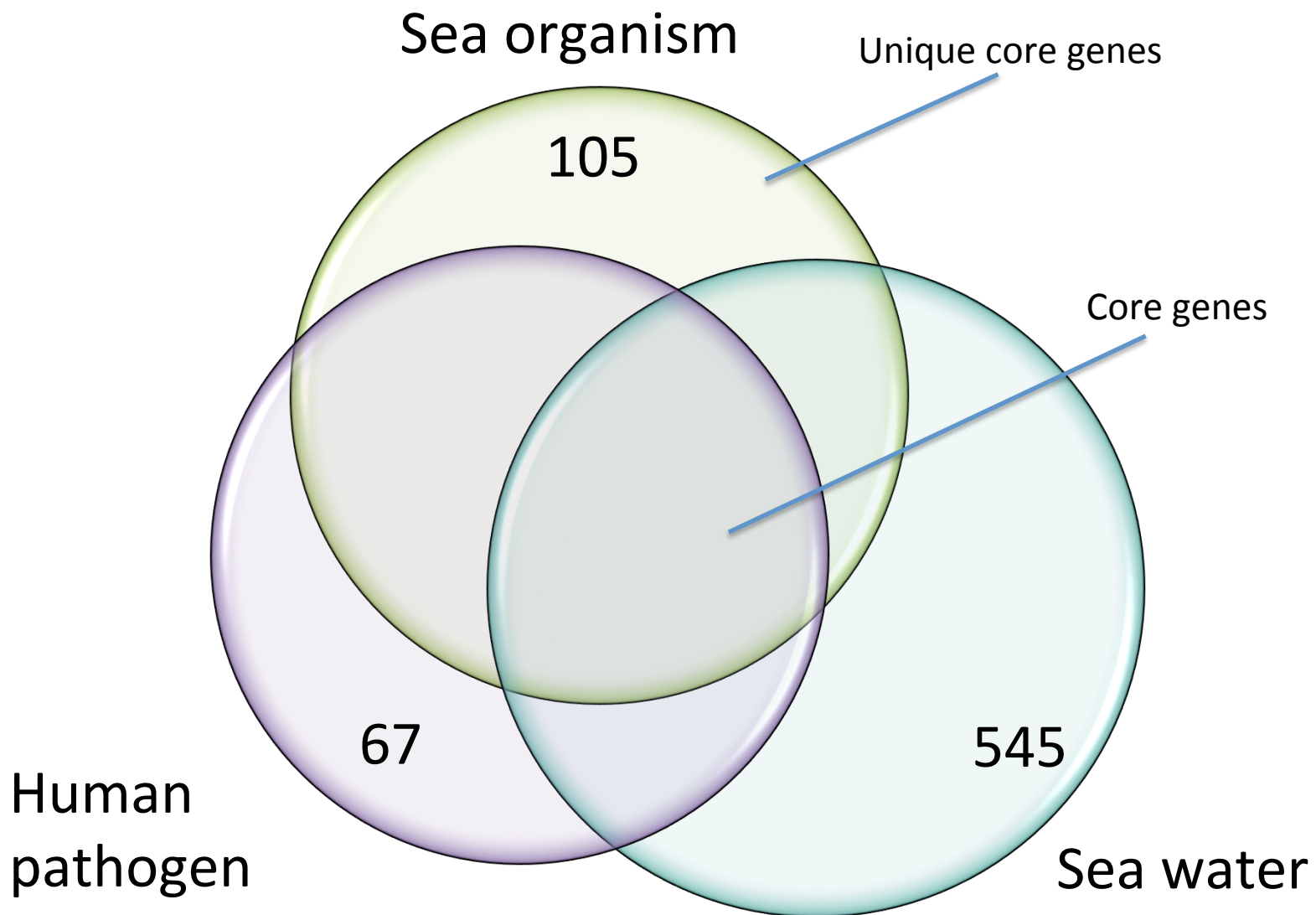
Make the gene list of unique core

REFERENCE C	<i>Vibrio harveyi</i>	<i>Vibrio fischeri</i>	<i>Vibrio vulnificus</i>	<i>Vibrio anguillar</i>	<i>Vibrio cholerae</i>	<i>Vibrio cholerae</i>	<i>Vibrio fischeri</i>	<i>Vibrio sp. EJY2</i>	<i>Vibrio furnissii</i>	<i>Vibrio cholerae</i>
VV0001	1	1	1	1	1	1	1	1	1	1
VV0002	1	1	1	1	0	1	0	1	0	1
VV0003	1	1	0	0	0	1	0	0	1	1

*Vibrio splendidus* LGP32 Cellular Overview

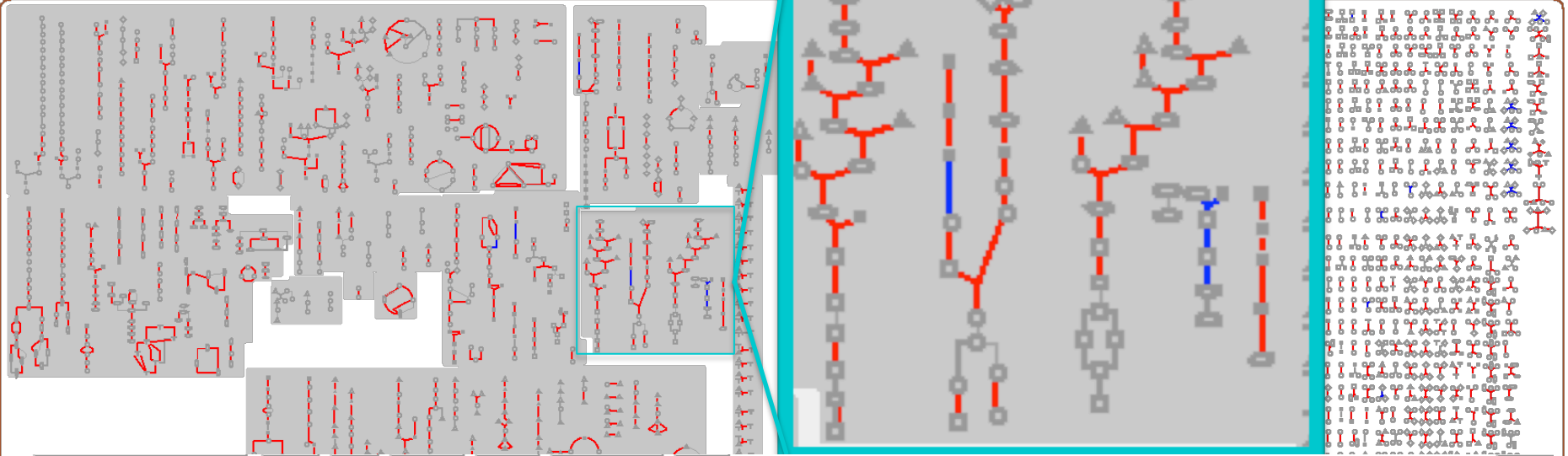
Mapping the genes on reference genome





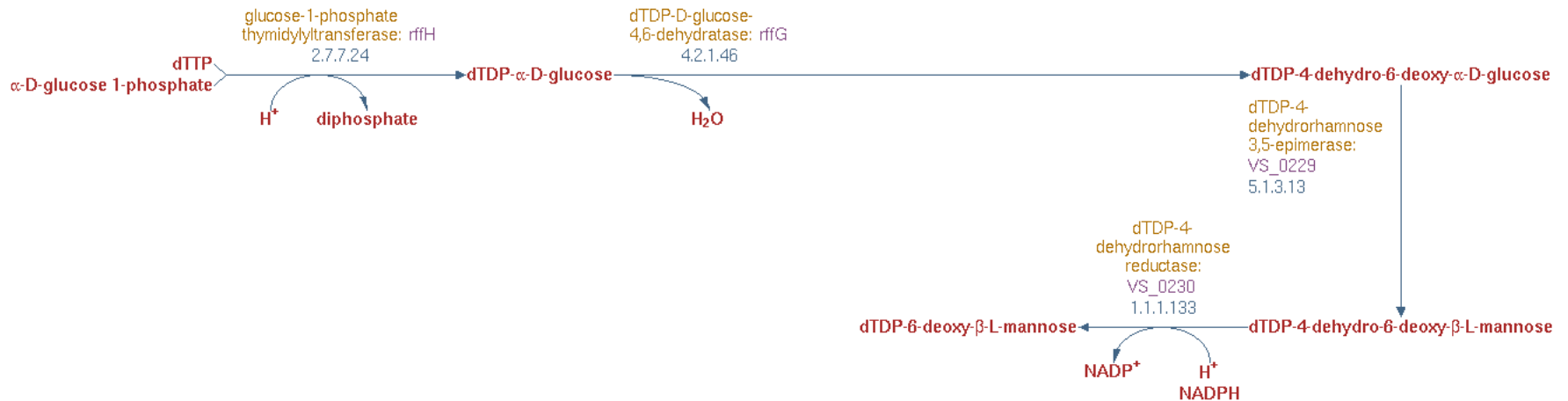
# Sea Organism group unique gene

## *Vibrio splendidus* LGP32 Cellular Overview



## *Vibrio splendidus* LGP32 Pathway: dTDP-L-rhamnose biosynthesis I

[More Detail](#) [Less Detail](#) [Species Comparison](#)



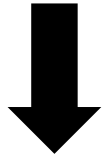
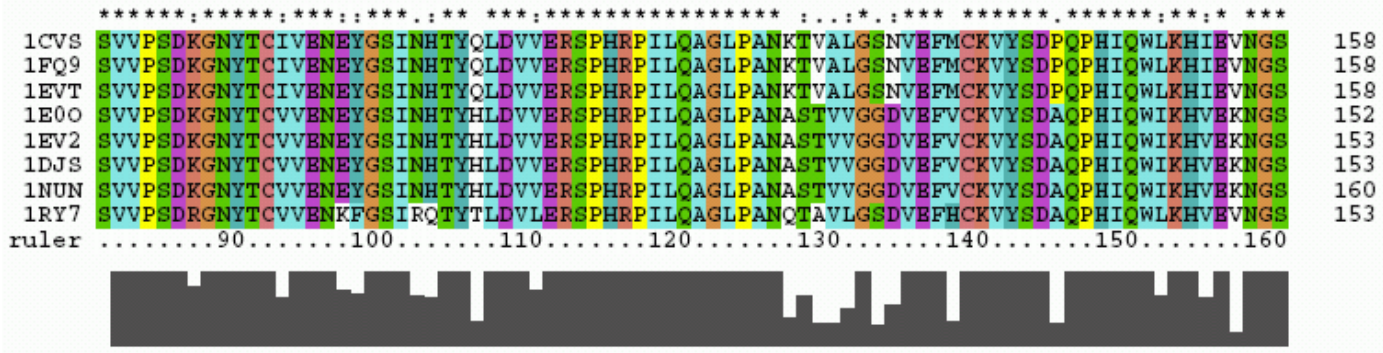
# Draft Genome Data

- 7 “draft” vibrio genomes
  - Annotated contigs
- Isolated from anemones
- Samples chosen based on MLST
- 2009-2012



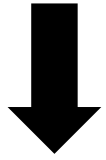
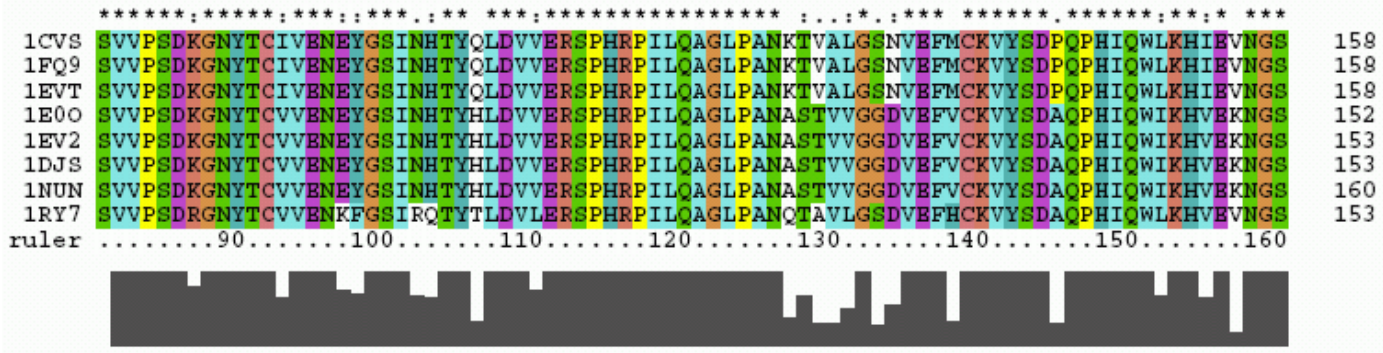
<http://www.stanford.edu/~siegelr/california/waddellbeach.html>

# Sequence alignment





# Sequence alignment



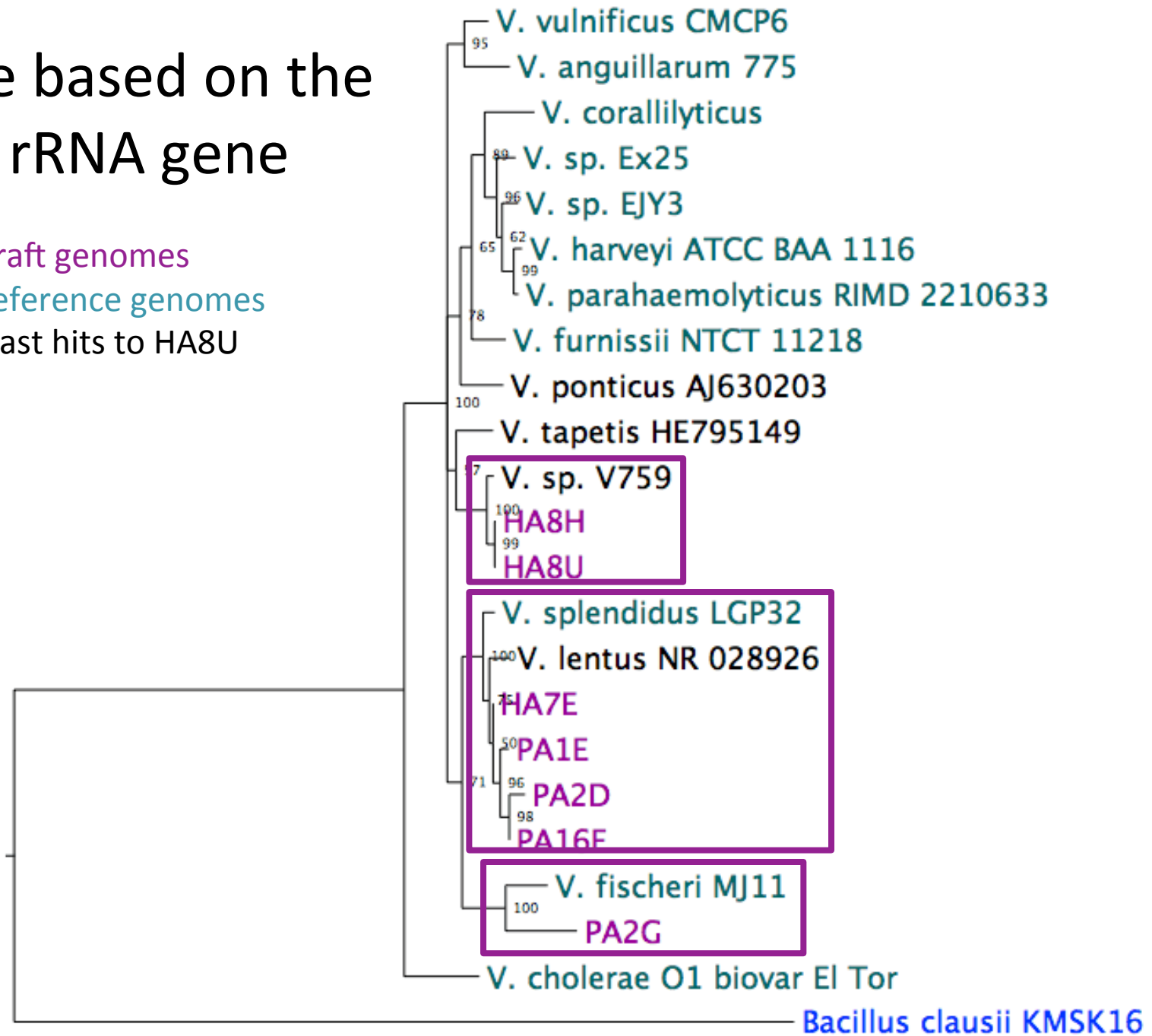


# Tree based on the 16S rRNA gene

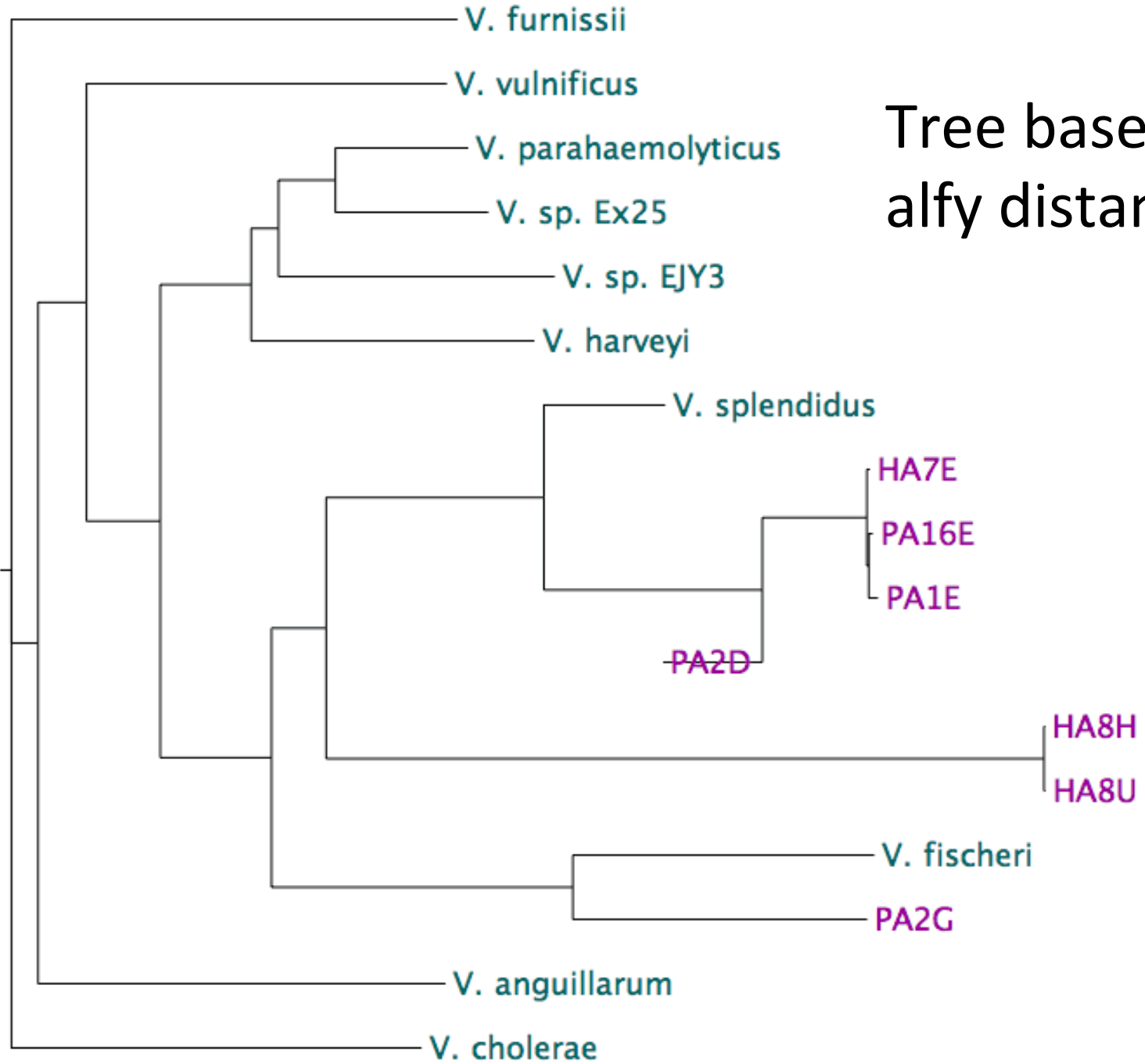
Draft genomes

Reference genomes

Blast hits to HA8U

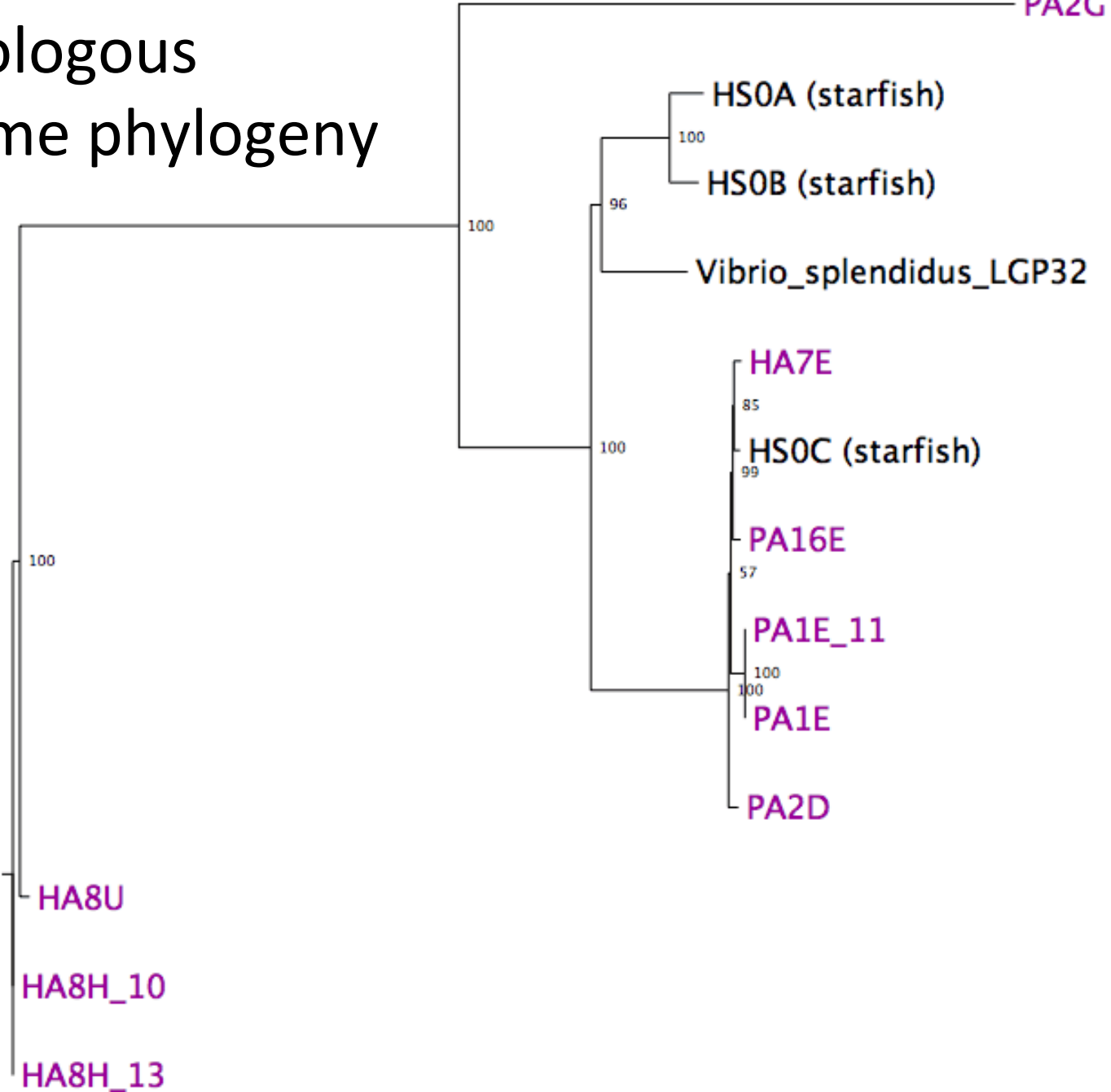


Tree based on kr/alfy distances



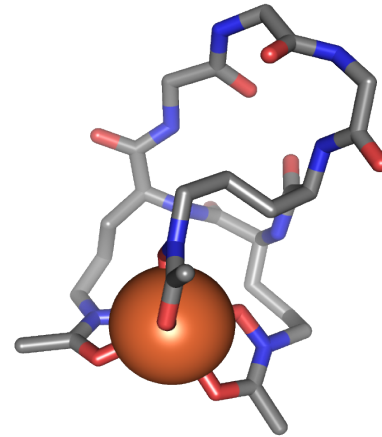
0.02

# Homologous genome phylogeny

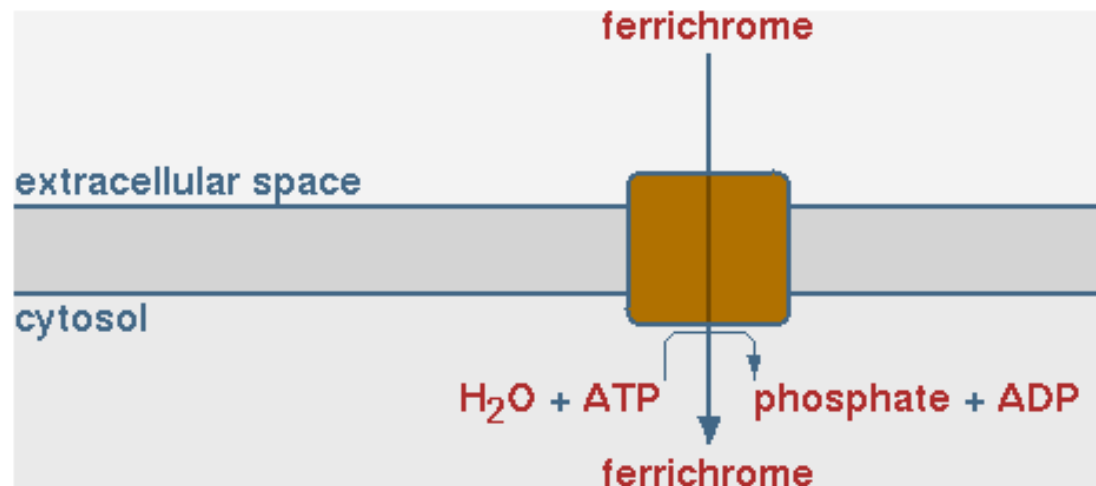


# Ferrichrome ABC Transporter: ATP-binding protein

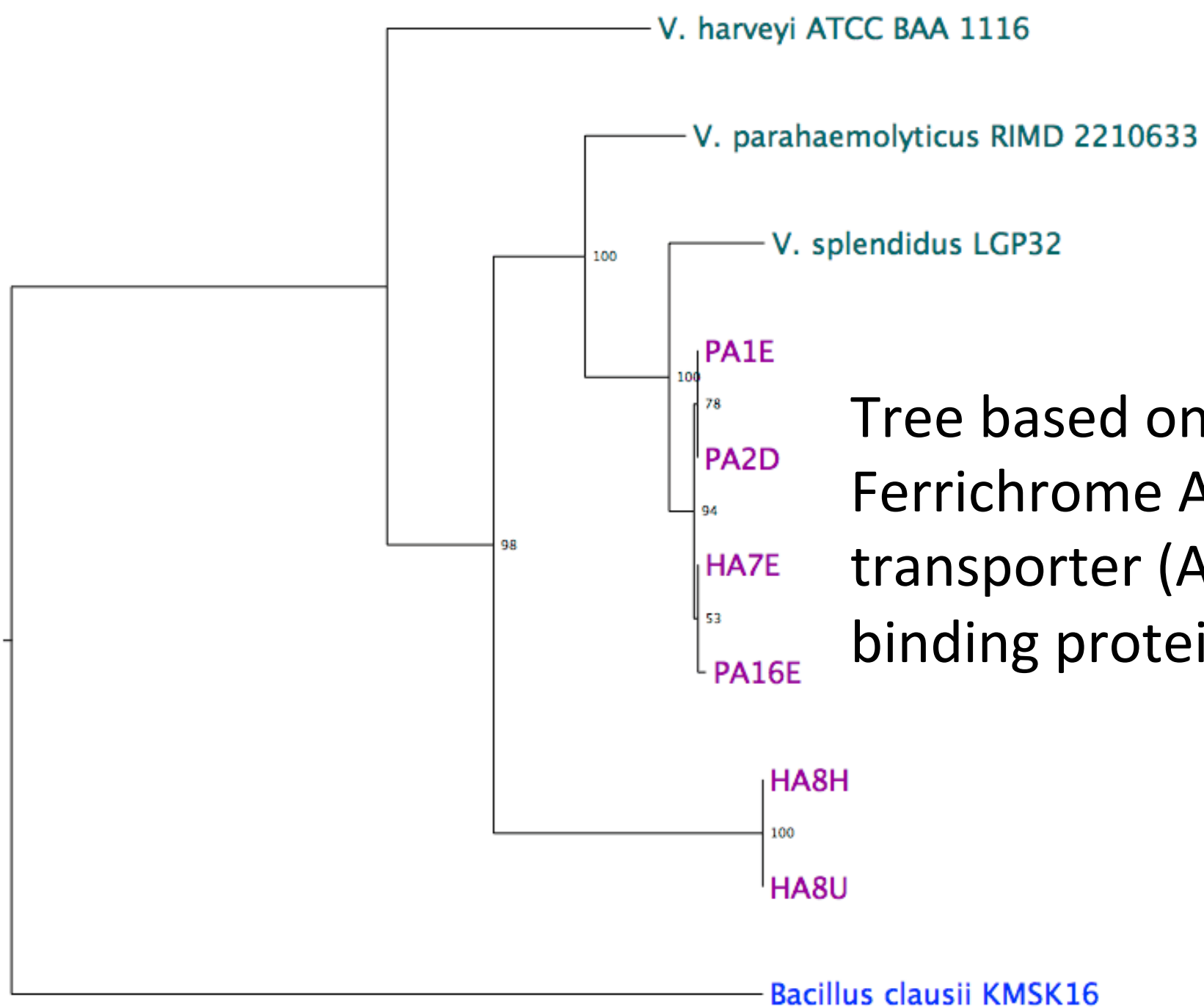
- Cyclic hexa-peptide that complexes with iron
  - Siderophore
- Transmembrane protein
- Utilizes energy from hydrolysis of ATP



<http://en.wikipedia.org/wiki/File:Ferrichrome.png>



<http://biocyc.org/BHAL272558/NEW-IMAGE?type=GENE&object=GJC5-3395>



Tree based on the Ferrichrome ABC transporter (ATP binding protein)

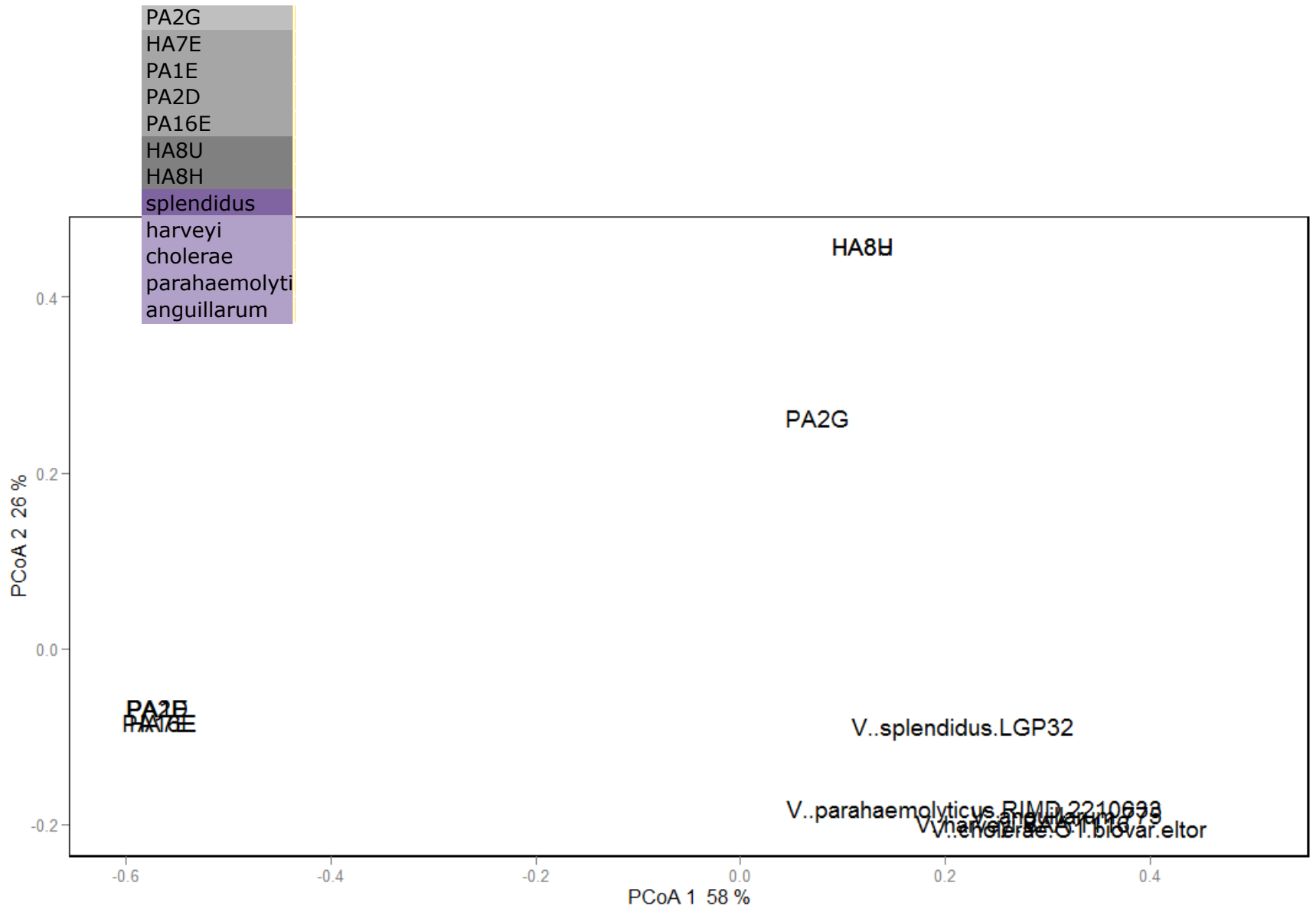
0.09

# Iron-acquisition and iron-containing genes

- Annotated genes of our 7 Vibrio-strains
- Annotated genes of 5 database strains
- SEED list of genes for iron acquisition and metabolism (reference)
- Script to look for iron-acquisition genes
- Script to look for iron-containing genes



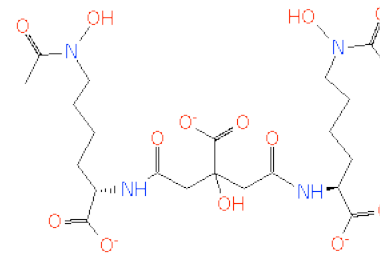




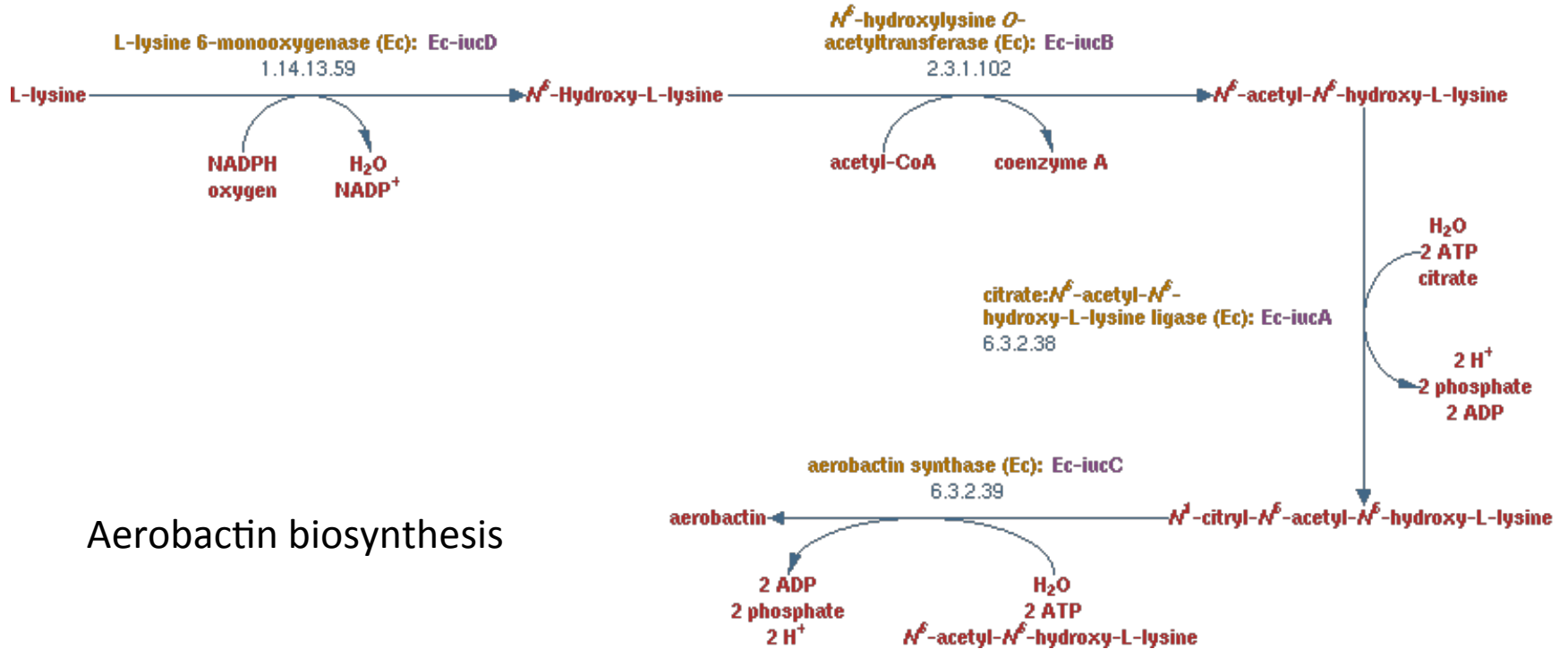


<b>Genome Name</b>	<b>HA7E</b>	<b>PA1E</b>	<b>PA2D</b>	<b>PA16E</b>	<b>HA8U</b>	<b>HA8H</b>
Iron compound ABC uptake transporter permease protein	0	0	0	0	1	1
Hypothetical protein	1	1	1	1	0	0
4'-phosphopantetheinyl transferase (EC 2.7.8.-) [enterobactin]	1	1	1	1	0	0
Transcriptional regulator, VCA0231	1	1	1	1	0	0
L-ornithine 5-monooxygenase (EC 1.13.12.-), PvdA of pyoverdin	1	1	1	1	0	0
Non-ribosomal peptide synthetase modules	1	0	0	1	0	0
Hypothetical MbtH-like protein, PA2412	1	0	0	1	0	0
Vibrioferrin ligase/carboxylase protein	1	1	1	1	0	0
Vibrioferrin receptor	1	1	1	1	0	0
<b>TonB-dependent</b>	0	0	0	0	1	1
TonB-dependent heme receptor	1	1	1	1	0	0
TonB-dependent hemin , ferrichrome	0	0	0	0	1	1
<b>Siderophore biosynthesis protein</b>	1	1	1	1	0	0
Citrate:6-N-acetyl-6-N-hydroxy-L-lysine ligase, alpha subunit (EC 6.3.2.27), aerobactin biosynthesis protein IucA @ Siderophore	0	0	0	0	1	1
N6-hydroxylysine O-acetyltransferase (EC 2.3.1.102), aerobactin biosynthesis protein IucB @ Siderophore synthetase small com	0	0	0	0	1	1
Citrate:6-N-acetyl-6-N-hydroxy-L-lysine ligase, alpha subunit (EC 6.3.2.27), aerobactin biosynthesis protein IucA @ Siderophore	0	0	0	0	1	1
L-lysine 6-monooxygenase [NADPH] (EC 1.14.13.59), aerobactin biosynthesis protein IucD @ Siderophore biosynthesis protein,	0	0	0	0	1	1
Aerobactin siderophore receptor	0	0	0	0	1	1
Ferric aerobactin ABC transporter, ATPase	0	0	0	0	1	1
Ferric aerobactin ABC transporter, periplasmic substrate binding	0	0	0	0	1	1
Ferric aerobactin ABC transporter, permease	0	0	0	0	1	1
Hypothetical protein in aerobactin uptake	1	1	1	1	0	0
Ferrichrome transport system permease protein FhuB (TC	1	1	1	1	0	0
Ferrichrome-binding periplasmic protein precursor (TC	1	1	1	1	0	0
Ferrichrome transport ATP-binding protein FhuC (TC	1	1	1	1	0	0
Ferric siderophore receptor	1	1	1	1	0	0
Ferric hydroxamate outer membrane receptor	1	1	1	1	0	0
Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), permease component	1	1	1	1	0	0
Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), periplasmic substrate binding protein	1	1	1	1	0	0
Ferric hydroxamate ABC transporter (TC 3.A.1.14.3), ATP-binding protein	1	1	1	1	0	0

# Genes related to aerobactin (iron chelating compound, siderophore)



	HA7E	PA1E	PA2D	PA16E	HA8U	HA8H
Citrate:6-N-acetyl-6-N-hydroxy-L-lysine ligase, alpha subunit (EC 6.3.2.27), aerobactin biosynthesis protein IucA	0	0	0	0	1	1
N6-hydroxylysine O-acetyltransferase (EC 2.3.1.102), aerobactin biosynthesis protein IucB	0	0	0	0	1	1
Citrate:6-N-acetyl-6-N-hydroxy-L-lysine ligase, alpha subunit (EC 6.3.2.27), aerobactin biosynthesis protein IucA	0	0	0	0	1	1
L-lysine 6-monooxygenase [NADPH] (EC 1.14.13.59), aerobactin biosynthesis protein IucD	0	0	0	0	1	1
Aerobactin siderophore receptor	0	0	0	0	1	1
Ferric aerobactin ABC transporter, ATPase	0	0	0	0	1	1
Ferric aerobactin ABC transporter, periplasmic substrate binding	0	0	0	0	1	1
Ferric aerobactin ABC transporter, permease	0	0	0	0	1	1
Hypothetical protein in aerobactin uptake	1	1	1	1	0	0

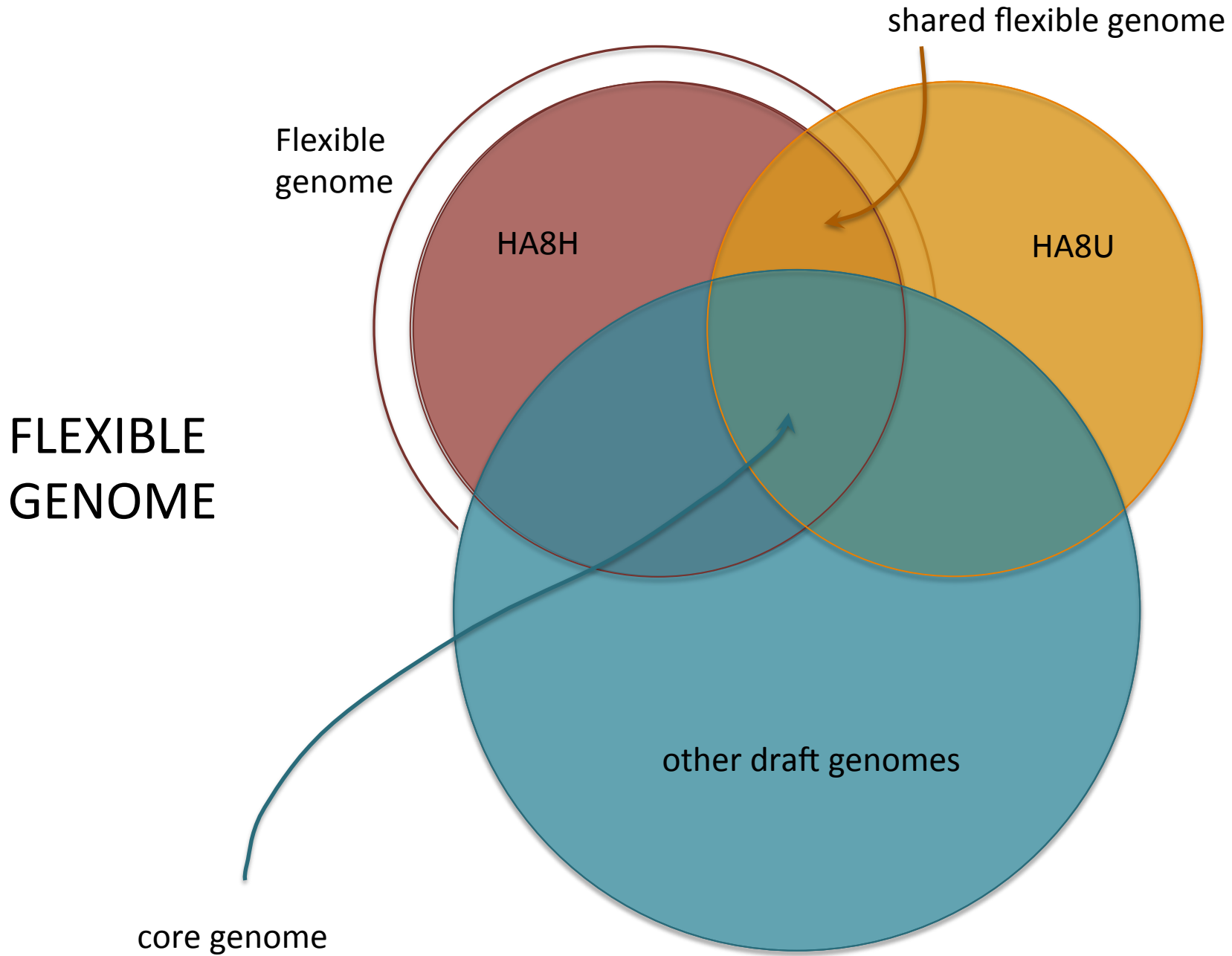




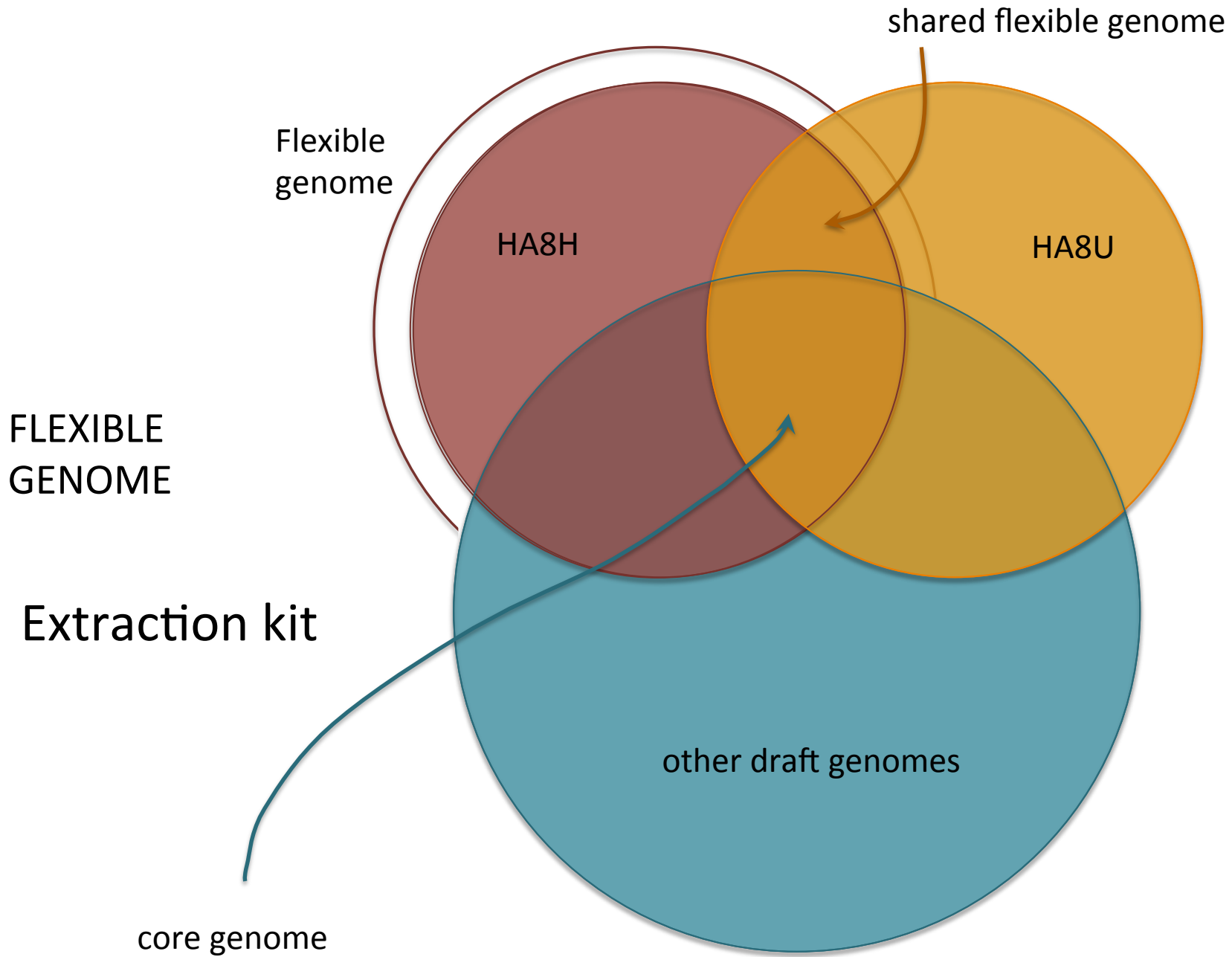
**Gene**

**Strain**

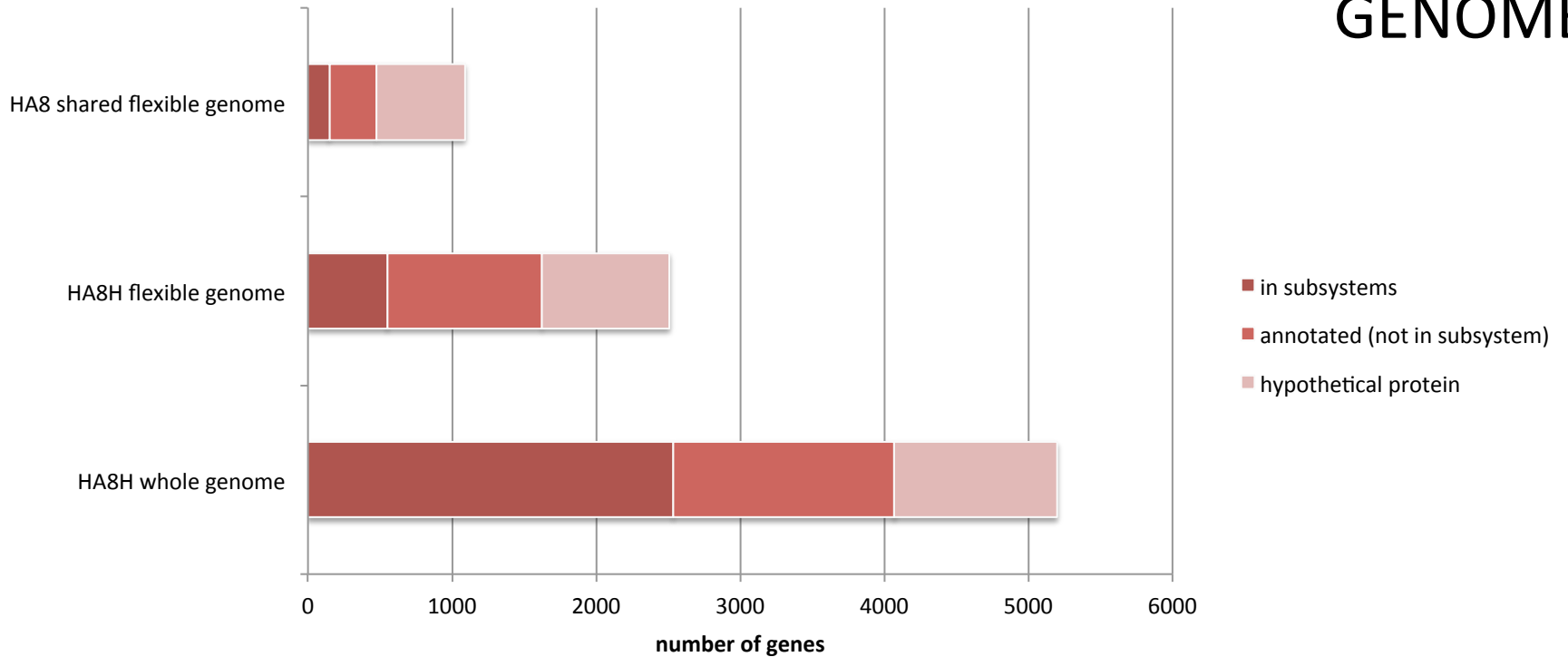
	PA2G	HA7E	PA1E	PA2D	PA16E	HA8U	HA8H	splendidus	harveyi	cholerae	parahaemolyticus	anguillarum
Predicted L-lactate dehydrogenase, Iron-sulfur cluster-binding subunit	0	0	0	0	0	1	1	0	0	0	0	0
Predicted L-lactate dehydrogenase, Fe-S oxidoreductase subunit	0	0	0	0	0	1	1	0	0	0	0	0
Fe-S	0	0	0	0	0	1	1	0	0	0	0	0







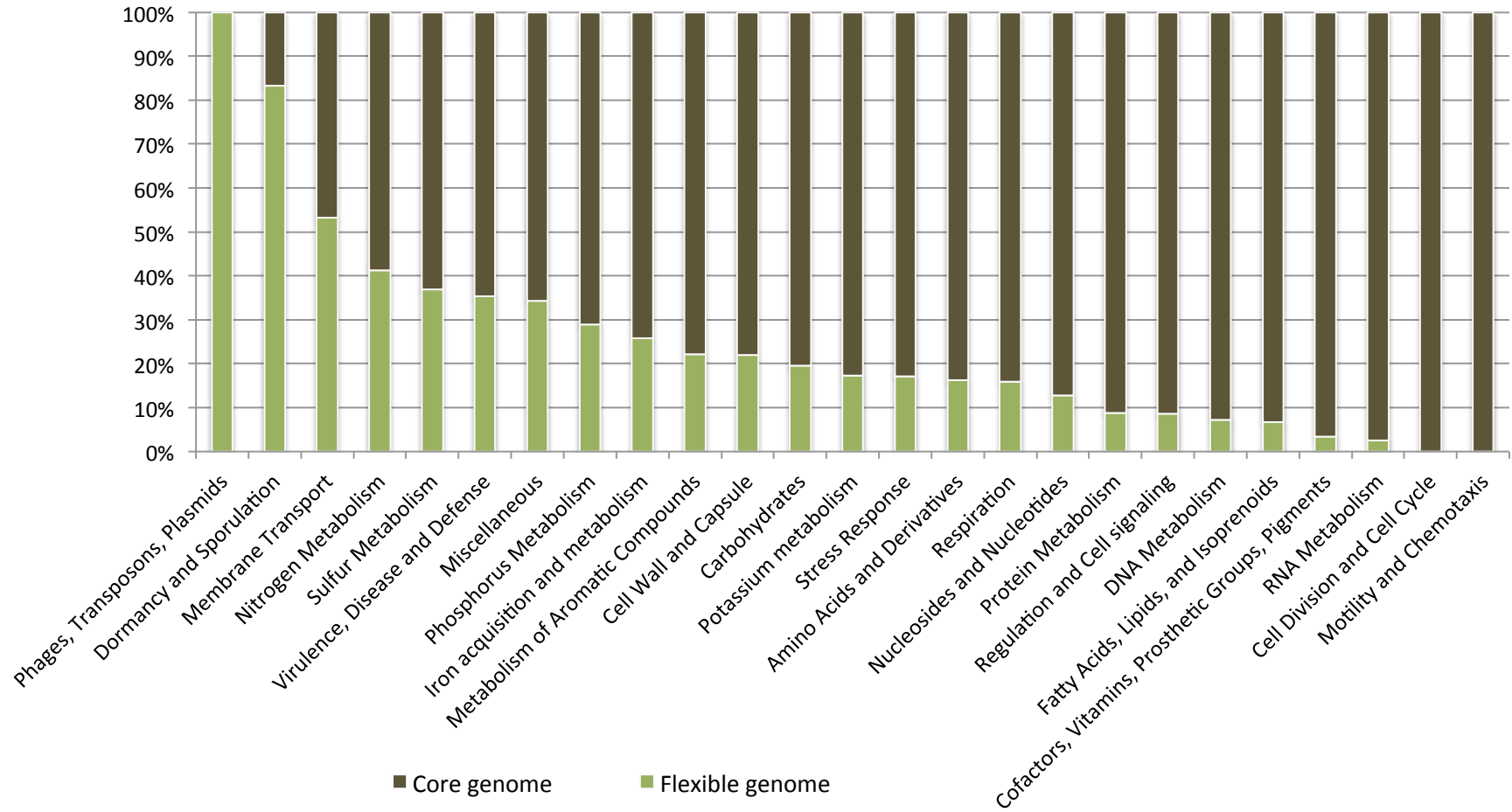
# GLIMPSE OF A FLEXIBLE GENOME



FUNCTION	SUBSYSTEM	SUBCATEGORY	CATEGORY
TesB-like acyl-CoA thioesterase 1	Acyl-CoA thioesterase II	Fatty acids	Fatty Acids, Lipids, and Isoprenoids
Cardiolipin synthetase	Cardiolipin synthesis	Phospholipids	Fatty Acids, Lipids, and Isoprenoids
TonB-dependent hemin , ferrichrome receptor	Hemin transport system	Iron acquisition and metabolism	Iron acquisition and metabolism
TonB-dependent hemin , ferrichrome receptor	Siderophore assembly kit	Siderophores	Iron acquisition and metabolism
...	...	...	...

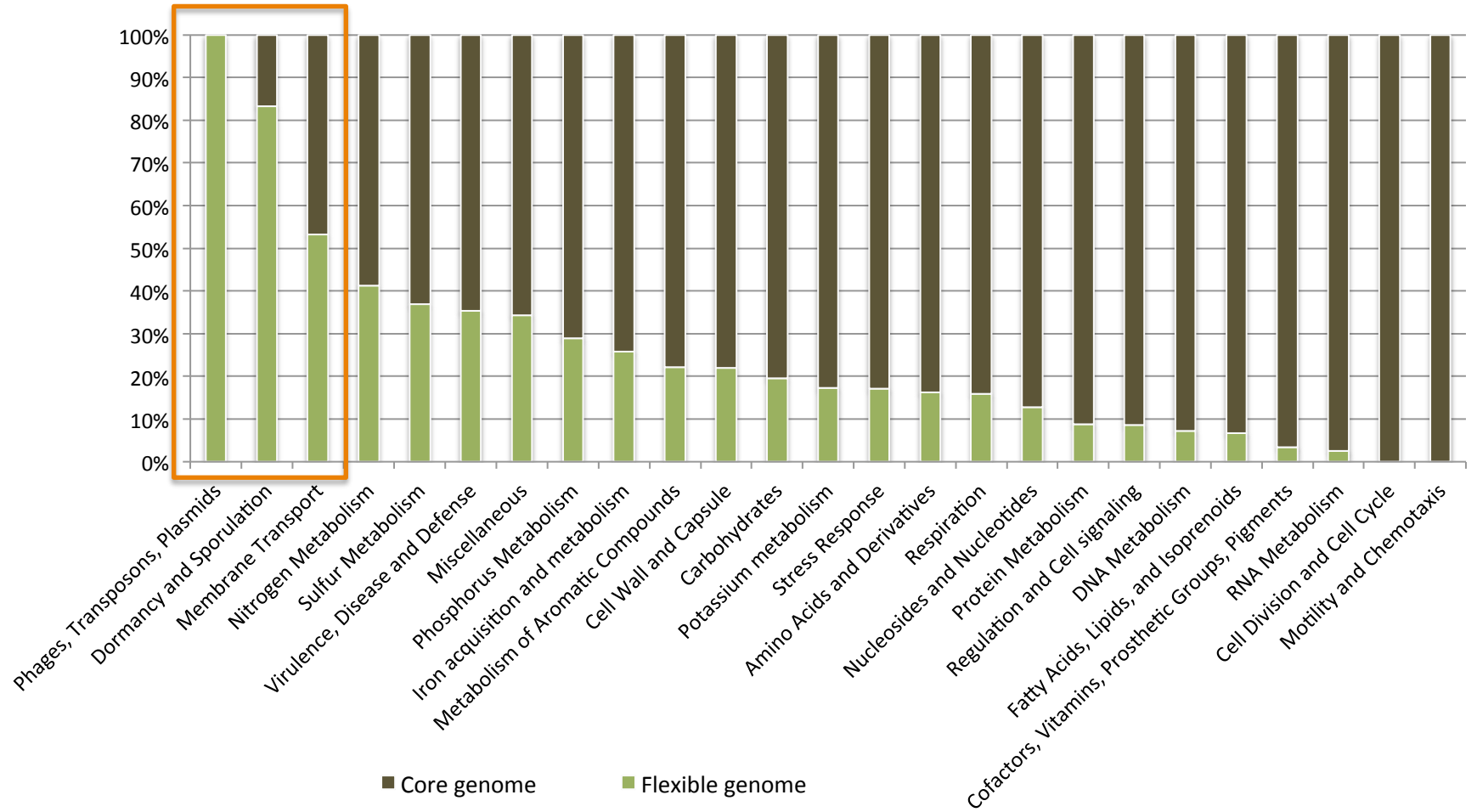
# SUBSYSTEMS ENRICHMENT IN FLEXIBLE GENOME

Fractions of functional subcategories - core vs. flexible genomes



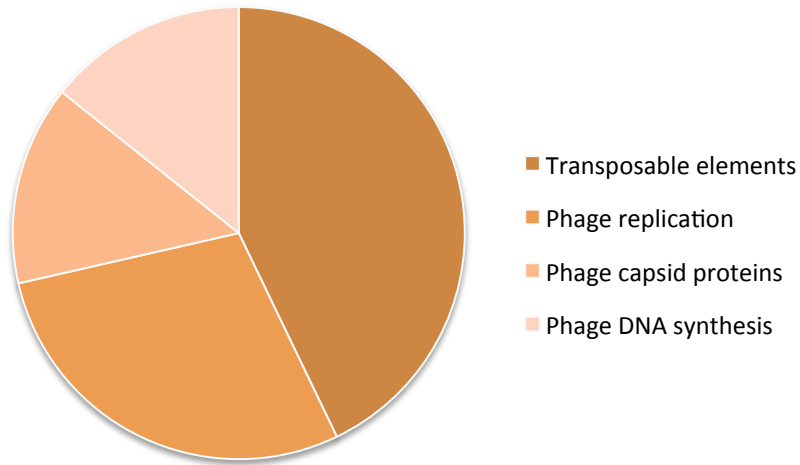
# SUBSYSTEMS ENRICHMENT IN FLEXIBLE GENOME

Fractions of functional subcategories - core vs. flexible genomes

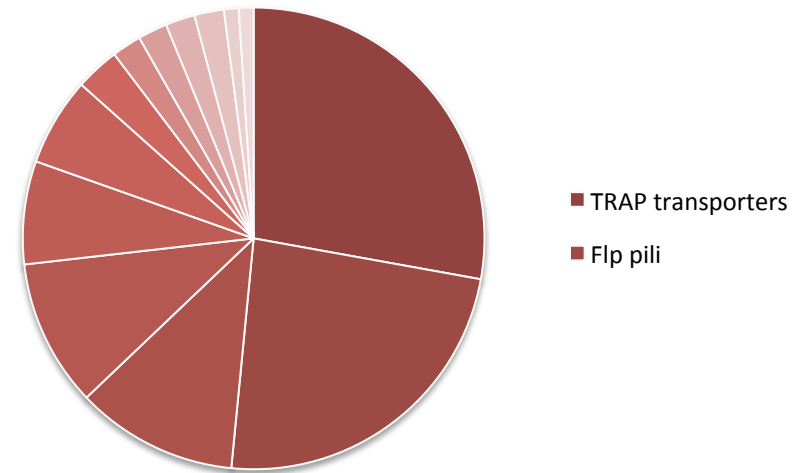


# SUBSYSTEMS ENRICHMENT IN FLEXIBLE GENOME

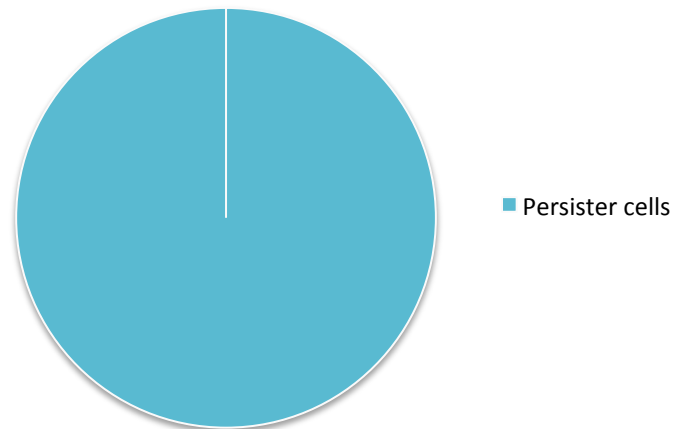
## Phages, Prophages, Transposable elements, Plasmids



## Membrane Transport

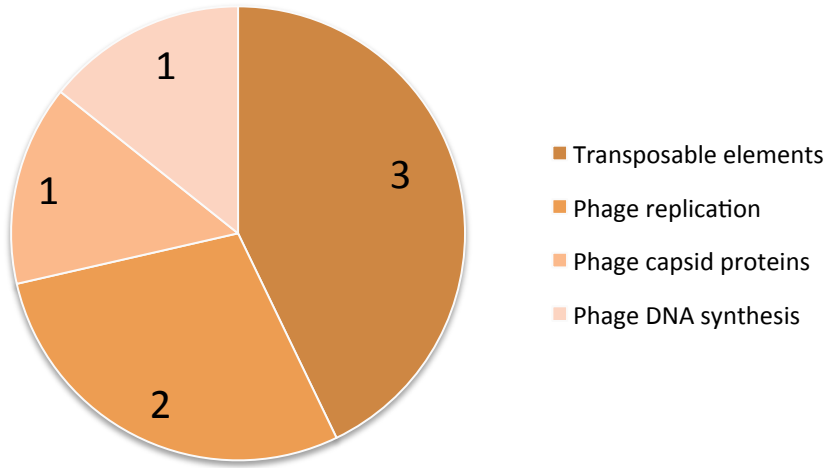


## Dormancy and sporulation

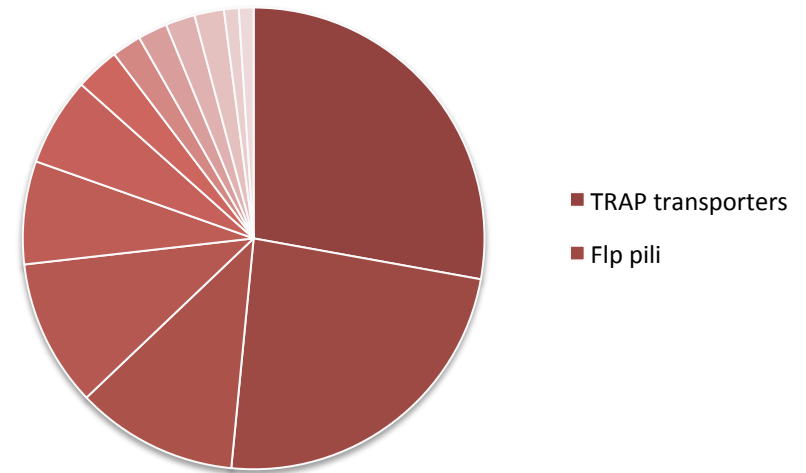


# SUBSYSTEMS ENRICHMENT IN FLEXIBLE GENOME

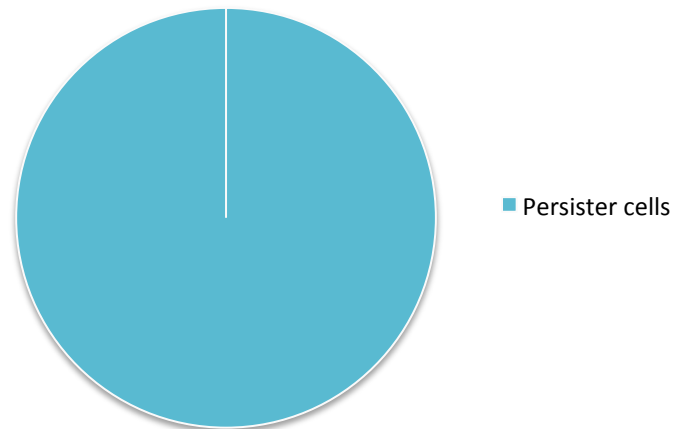
## Phages, Prophages, Transposable elements, Plasmids



## Membrane Transport

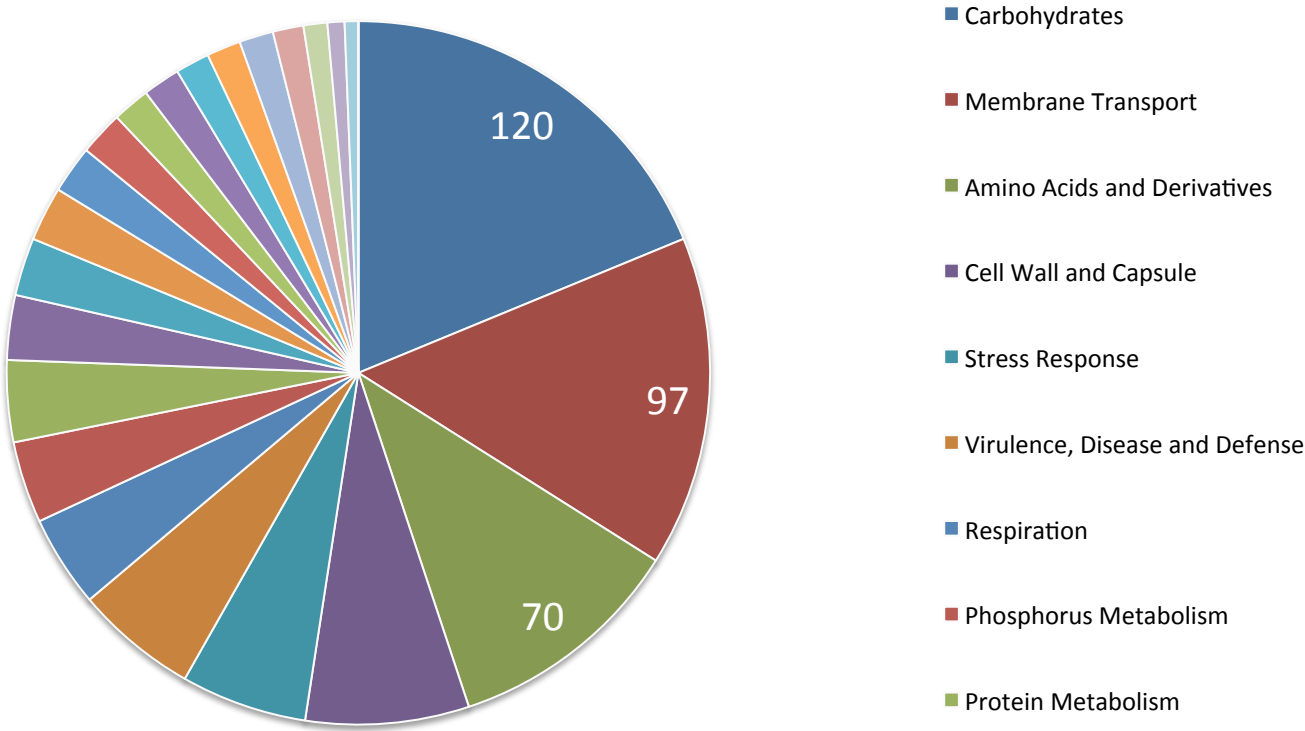


## Dormancy and sporulation



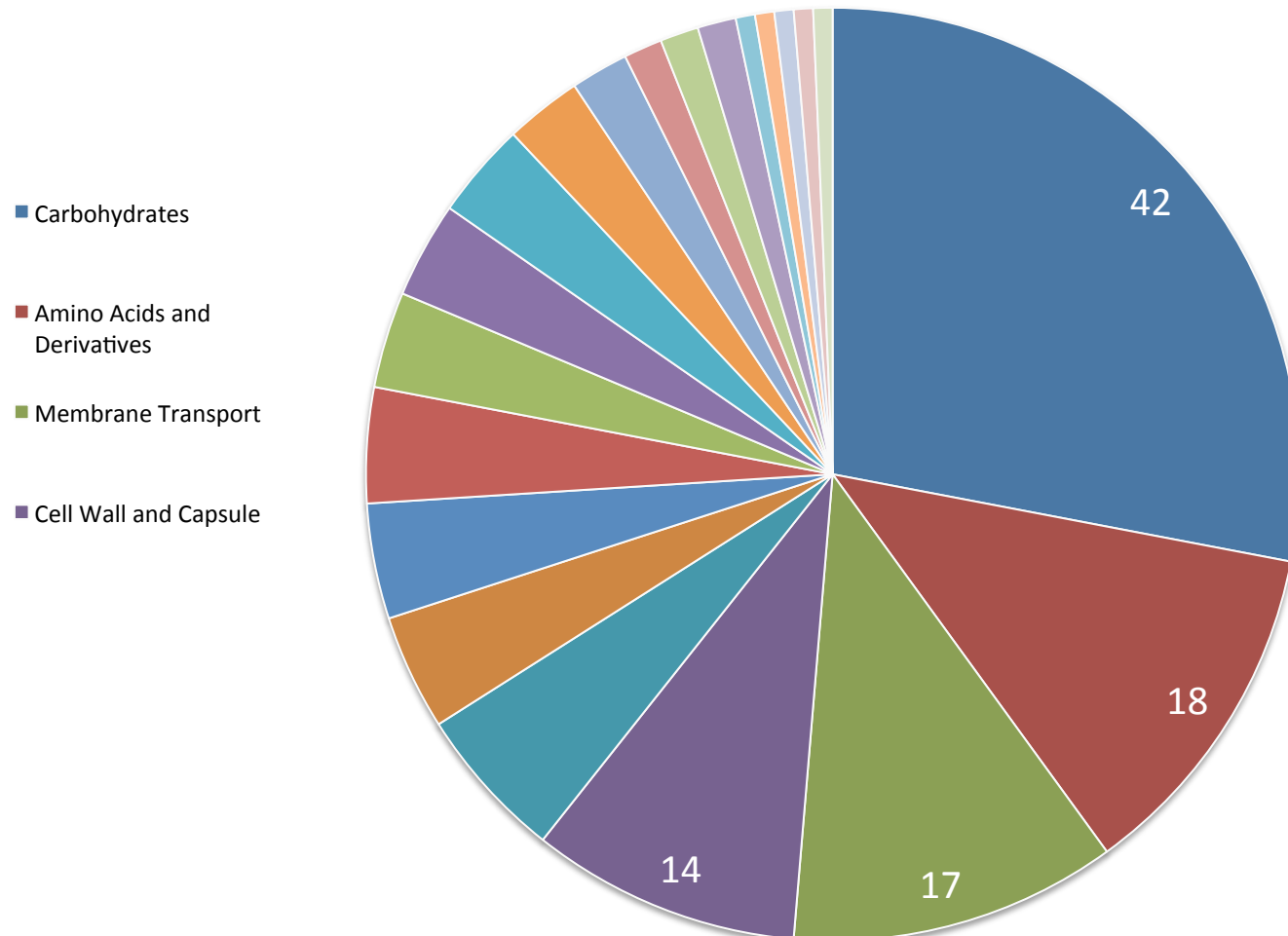
# THE MASTERPLAN OF HA8H

HA8H flexible genome: subcategory distribution



# COALITION: HA8\* SHARED FLEXIBLE GENOME

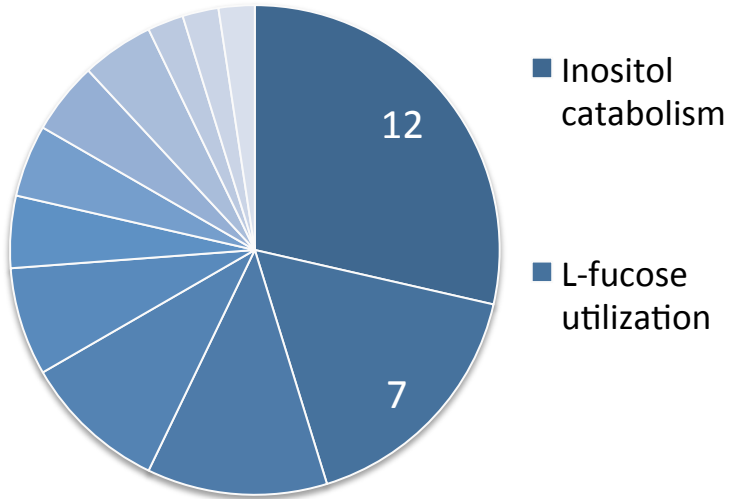
Subcategory distribution



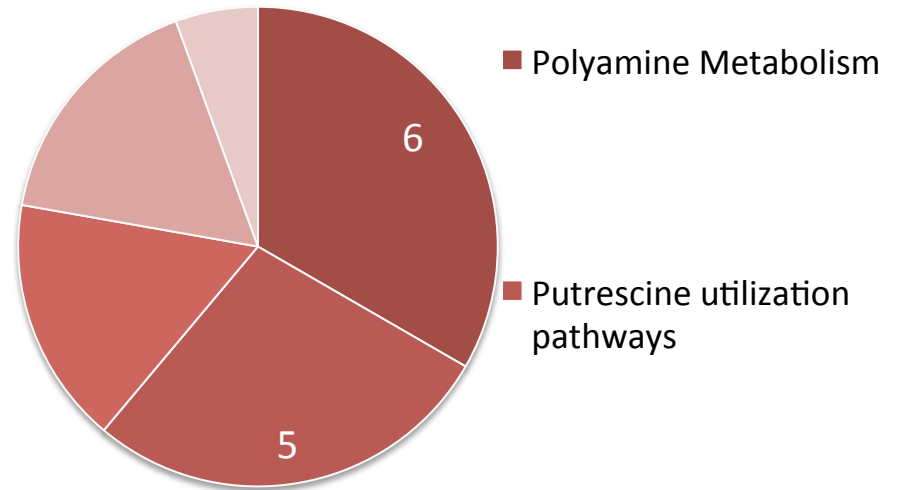


# RIDDLE: FIND THE NICHE

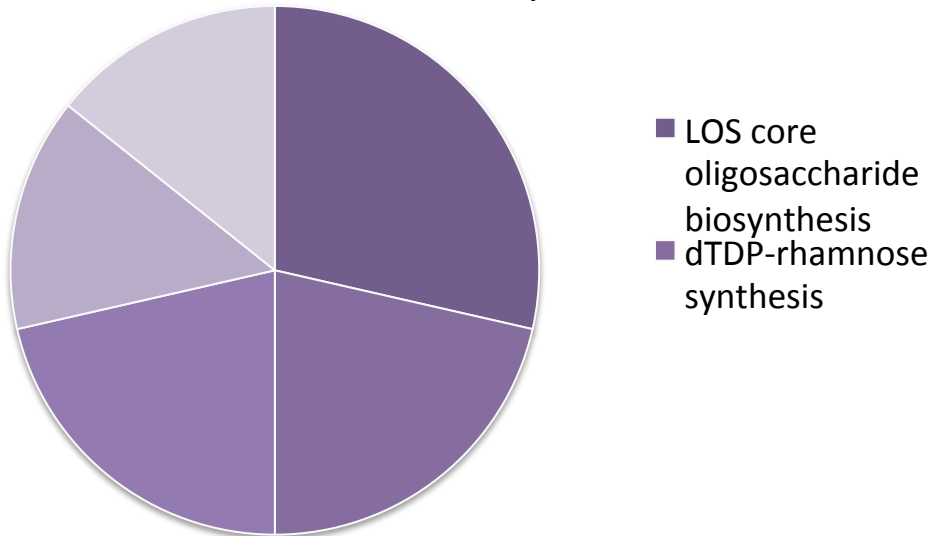
## Carbohydrates



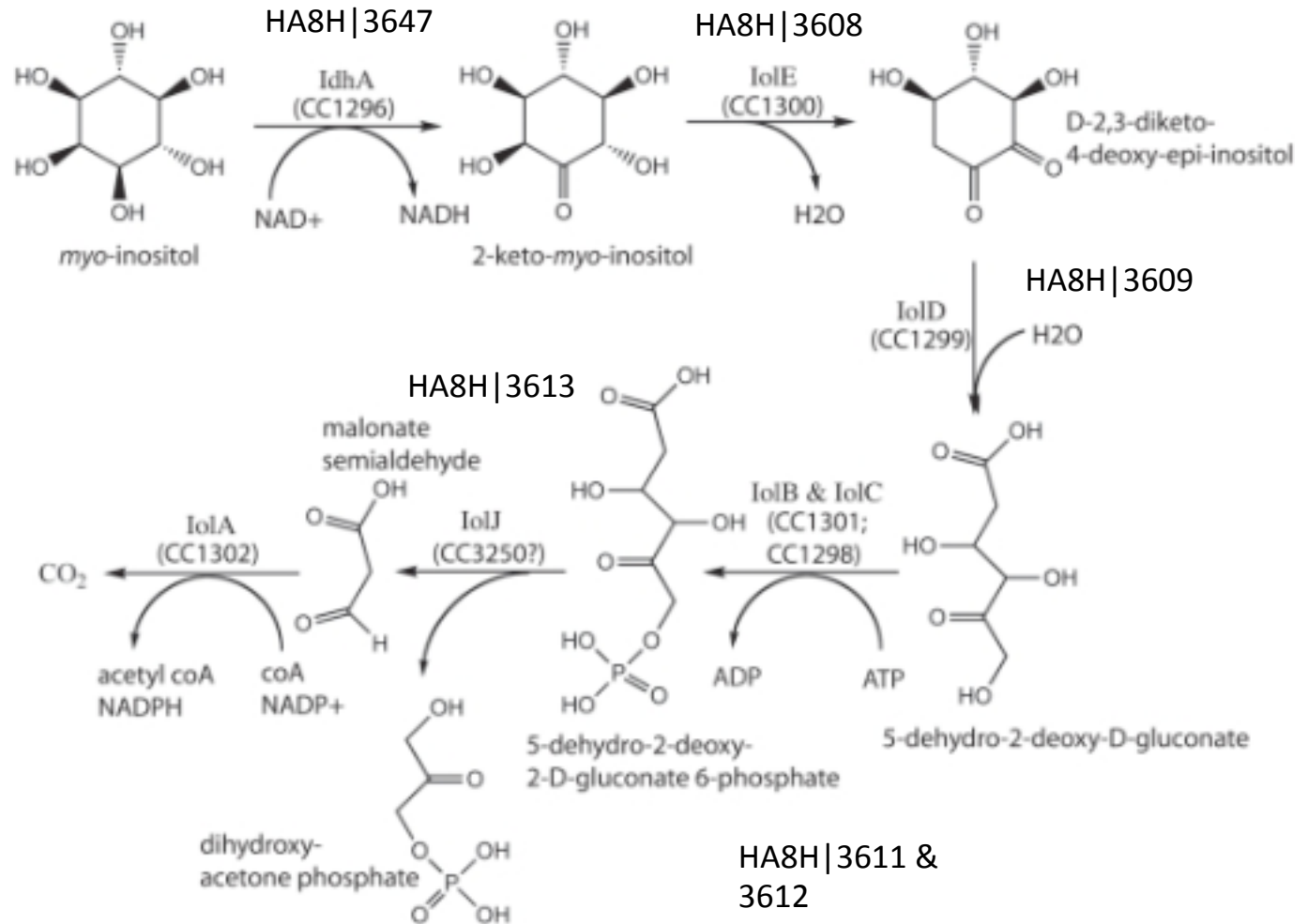
## Amino acids and derivatives



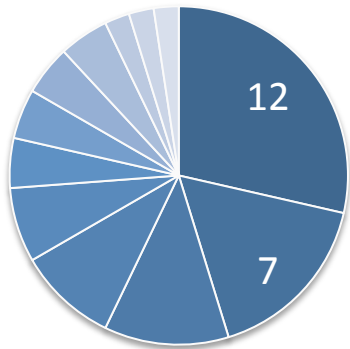
## Cell wall and capsule



# RIDDLE: FIND THE NICHE – inositol as a catabolic substrate



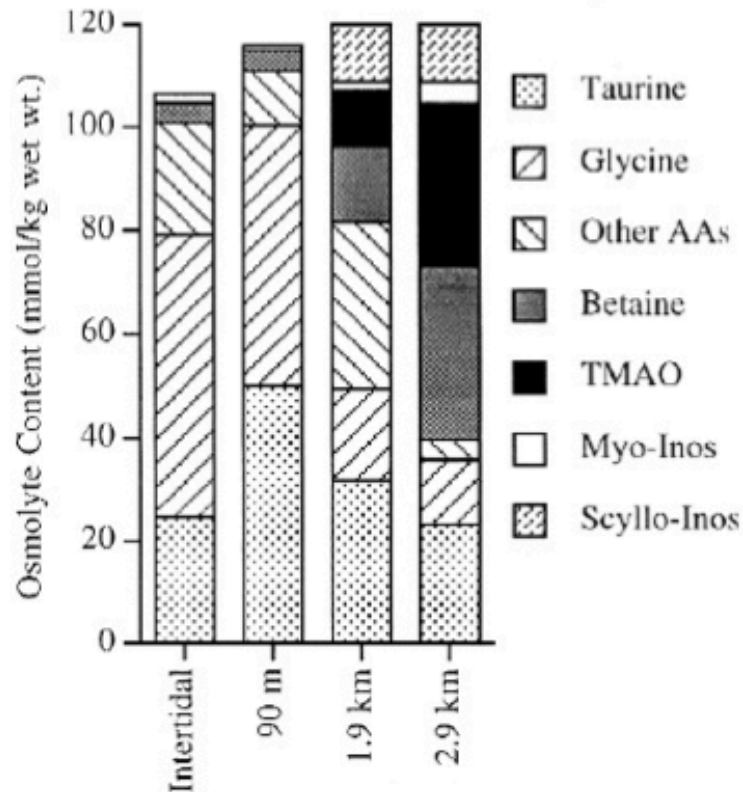
Carbohydrates



# RIDDLE: FIND THE NICHE

## TRIMETHYLAMINE OXIDE, BETAINE AND OTHER OSMOLYTES IN DEEP-SEA ANIMALS: DEPTH TRENDS AND EFFECTS ON ENZYMES UNDER HYDROSTATIC PRESSURE

P.H. YANCEY<sup>1,4</sup>, M.D. RHEA<sup>1</sup>, K.M. KEMP<sup>2</sup> and D.M. BAILEY<sup>2,3</sup>



**Fig. 2** Contents of organic osmolytes in septal tissues of sea anemones (data from Table 1). Inos: inositol

# RIDDLE: FIND THE NICHE – fucose utilization

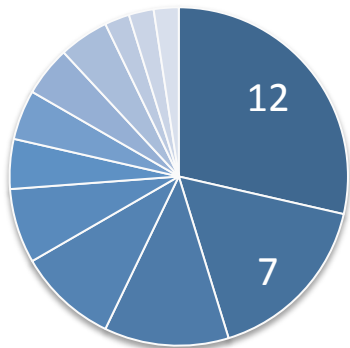
## The Presence of Fucose, Mannose, and Glucosamine-containing Heteropolysaccharide in Collagen from the Sea Anemone *Metridium dianthus*\*

(Received for publication, June 2, 1972)

RICHARD L. KATZMAN<sup>†</sup> AND ANDREW H. KANG<sup>†</sup>, §

From the Department of Biological Chemistry, Harvard Medical School at the McLean Hospital Research Laboratory, Belmont, Massachusetts 02178 and the Developmental Biology Laboratory, Department of Medicine, Harvard Medical School at the Massachusetts General Hospital, Boston, Massachusetts 02114

### Carbohydrates



### SUMMARY

Pepsin-solubilized collagen was prepared from the body-wall connective tissue of the sea anemone *Metridium dianthus*. It was judged to be collagen on the basis of its precipitability with 10% NaCl solution, its segment long spacing pattern in the electron microscope, its amino acid analysis, its melting behavior, its optical rotation, and its behavior on gel filtration chromatography, ion exchange chromatography, and disc electrophoresis. Pepsin-solubilized *Metridium* collagen contained two components, an  $\alpha$  chain and a  $\beta$  component, when examined by polyacrylamide gel disc electrophoresis. These components were separated by agarose 1.5 m gel filtration and further purified by ion exchange chromatography on carboxymethylcellulose. The highly purified  $\alpha$  and  $\beta$  chains each contained almost 1% by weight of the sugars fucose, mannose, and glucosamine, in the form of heteropolysaccharide. These sugars have not been previously reported to be present in highly purified, soluble collagen chains.

which a partial amino acid sequence was obtained, was relatively rich in hydroxyproline but contained slightly less than one-third mole per cent (27%) of glycine. The yield of the peptide obtained, 3.8%, was also quite low. Thus it was possible that Peptide C was derived from a collagenase-sensitive glycoprotein contaminant present in the pepsin-solubilized collagen previously used (2). In this paper we present evidence, based on salt solubility, segment long spacing pattern, amino acid analysis, melting behavior, optical rotation, and behavior on gel filtration chromatography, ion exchange chromatography, and disc electrophoresis, that the pepsin-solubilized preparation from *M. dianthus* described herein is indeed collagen. The purified  $\alpha$  and  $\beta$  collagen chains contained glucose and galactose, in the form of *O*- $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 2)-*O*- $\beta$ -D-galactopyranosyloxy-(1 $\rightarrow$ 5)-L-lysine, (1, 3), and fucose, mannose, and glucosamine, in the form of heteropolysaccharide (2). To the best of our knowledge this is the first report of sugars other than glucose and galactose occurring in highly purified, soluble collagen chains.

### EXPERIMENTAL PROCEDURE

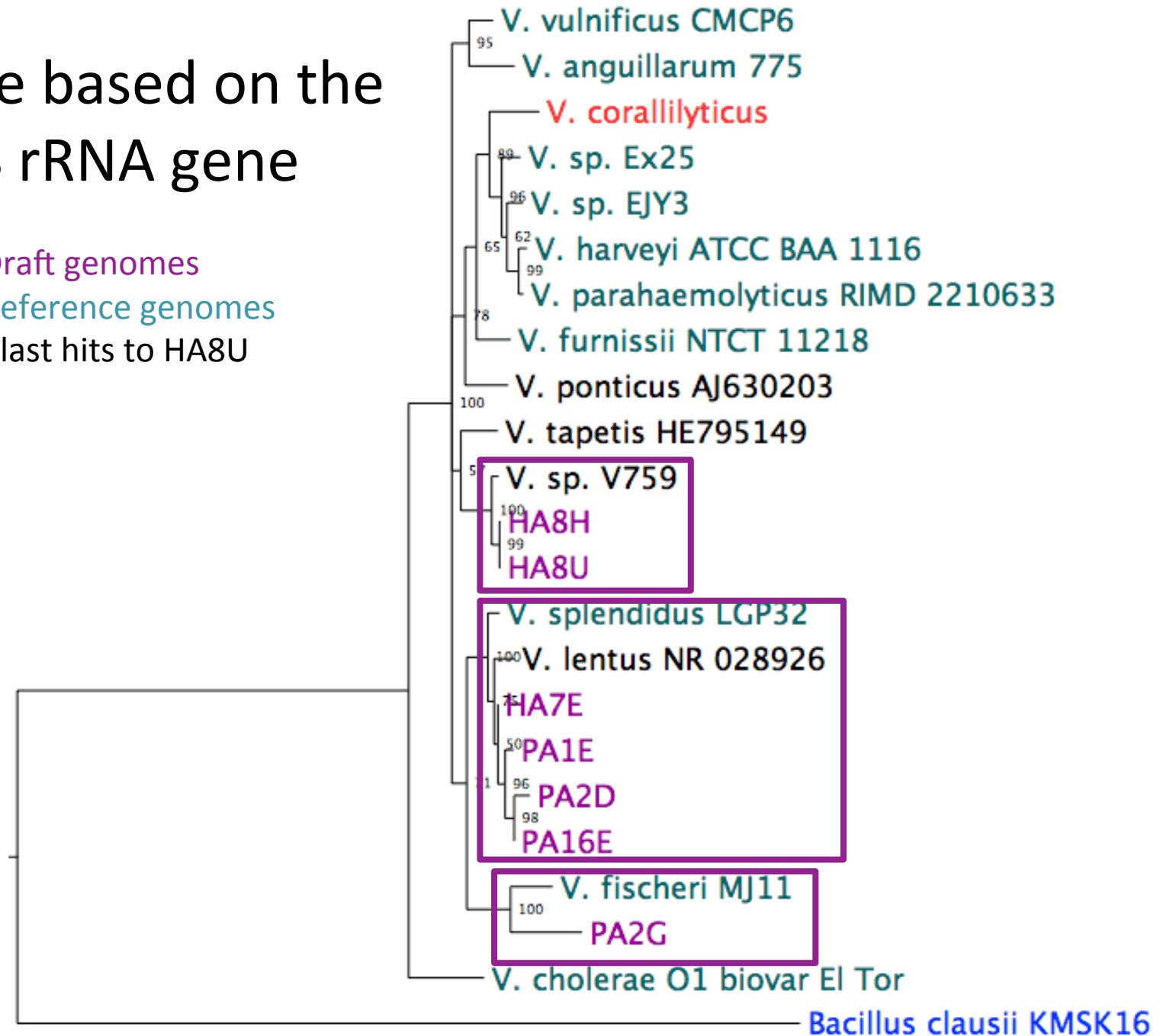
All procedures in the preparation of *Metridium* pepsin-solu-

# Tree based on the 16S rRNA gene

Draft genomes

Reference genomes

Blast hits to HA8U



0.04

# SOLVING THE RIDDLE: FIND THE NICHE

- HA8\* schedule:
  - Found in starburst anemone
  - Possess extra surface adhesion capability (pili, LOS core)
  - Unique catabolic substrate found in anemone's cell and anemone's mucus
  - Degradation pathways for anemone's phospholipids



# SOLVING THE RIDDLE: FIND THE NICHE

HA8\* : parasitic vibrios  
specialized in eating  
anemone's mucus



FOOD !!!!

# CONCLUSION

- What makes a vibrio a vibrio ?

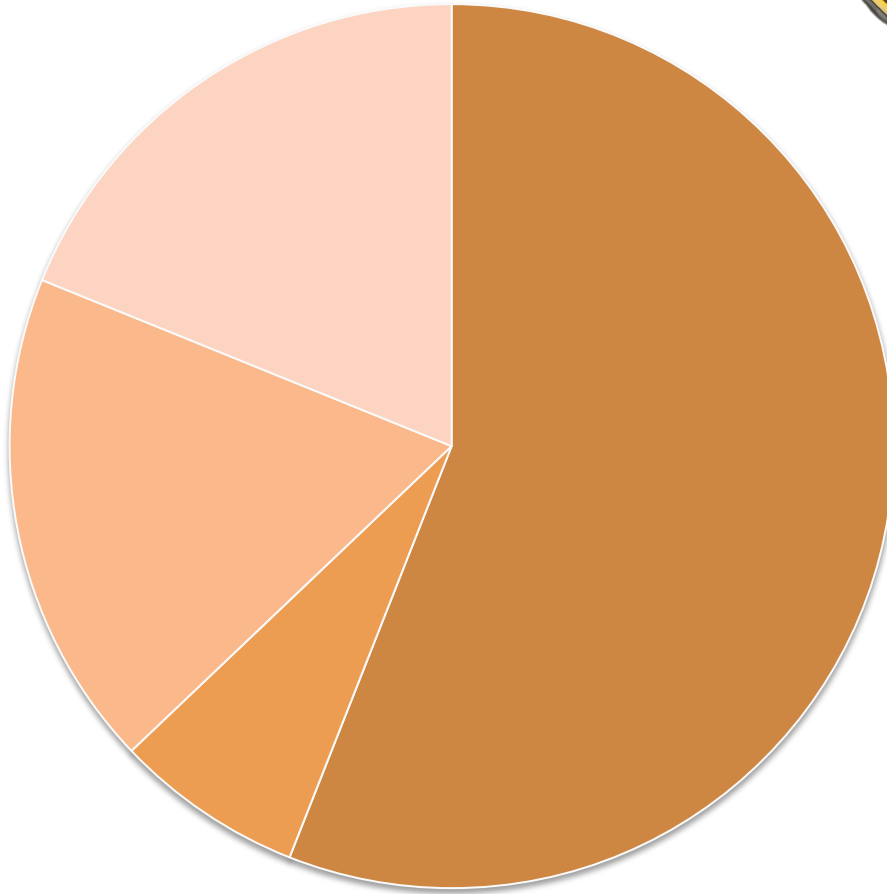
We don't know !



# CONCLUSION

- What makes a vibrio a vibrio ? But :
  - All vibrio share core genome
  - Core genome: loose concept but very useful
  - Hypothesis on the environment can lead to the identification of differences at the genetic level.
  - Polysaccharide genes are conserved across in the genes living in sea organisms for better adhesion.
  - Fitness in different niche are correlated with gene content.
  - HA8\* have specific iron acquisition genes.
  - Vibrios' tree are congruent with each other.
  - No apparent complicated evolution.
  - Quirks and tricks of an organism are present in the flexible genome.
  - Comparing flexible genomes give insight on the organism's niche.
  - HA8\* are anemone's parasite that eat anemone's cell compound.

## Greatness in test according to Vibrio's tasting



- Caramel 690
- ChemoTaxis
- Red Light 4 IPA
- Row 2 Beer

