

# Group #7

## *Vibrio*

*„On the Origins of a Vibrio Species”*

Iliana

Susan s050054

Kasper s052766

Dominika s111948



## On the Origins of a *Vibrio* Species

Tammi Vesth • Trudy M. Wassenaar • Peter F. Hallin •  
Lars Snipen • Karin Lagesen • David W. Ussery

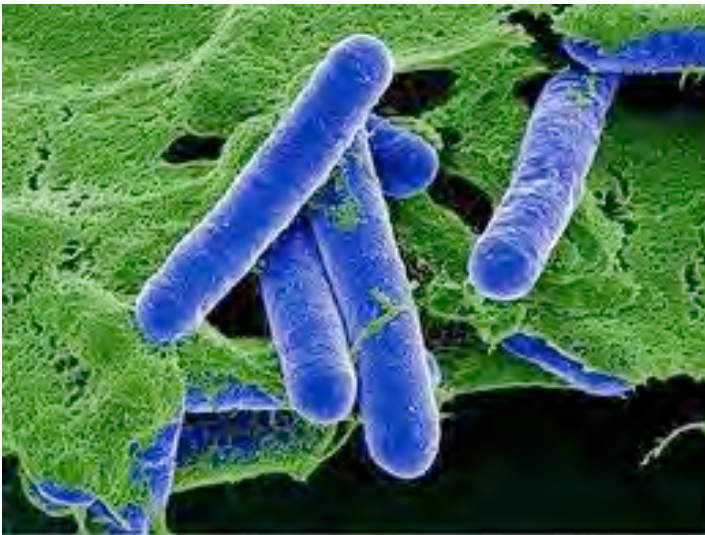
Received: 3 July 2009 / Accepted: 17 September 2009 / Published online: 15 October 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

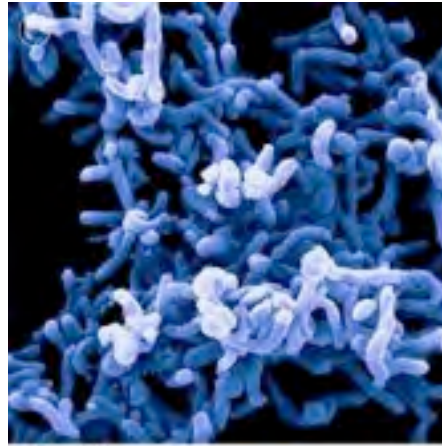
**Abstract** Thirty-two genome sequences of various *Vibrionaceae* members are compared, with emphasis on what makes *V. cholerae* unique. As few as 1,000 gene families are conserved across all the *Vibrionaceae* genomes analysed; this fraction roughly doubles for gene families conserved within the species *V. cholerae*. Of these, approximately 200 gene families that cluster on various locations of the genome are not found in other sequenced *Vibrionaceae*; these are possibly unique to the *V. cholerae* species. By comparing gene family content of the analysed genomes, the relatedness to a particular species is identified for two unsequenced genomes. Conversely, two genomes

presumably belonging to the same species have suspiciously dissimilar gene family content. We are able to identify a number of genes that are conserved in, and unique to, *V. cholerae*. Some of these genes may be crucial to the niche adaptation of this species.

# Bacterial Species

“a category that circumscribes a (preferably) genomically coherent group of individual isolates/strains sharing a high degree of similarity in (many) independent features, comparatively tested under highly standardized conditions”





Independent features = phenotypes

Phenotypes  $\longrightarrow$  can be tested

# Phylogenetic relationships

## Ideally:

16S rRNA genes are identical or almost identical between isolates belonging to the same species and those differ from isolates belonging to different species

## In reality:

different species can share identical rRNA genes, for instance *E. coli* and *Shirella*

isolates from one species can have different rRNA genes (beyond 97 % - „border“ of the species)

# Lateral transfer!!!

Is the phylogenetic tree the most suitable model for bacterial ancestry?

Is it valid for network-like evolutionary structures?



# Hypothesis

Certain genes can be said to be specific and conserved to a particular species but not present in related species



# Bacterial family *Vibrionaceae*

- 51 recognised species
- Found mainly in marine environments
- Mostly symbionts

BUT

- Human pathogens:
  - ✓ *Vibrio cholerae* – producing cholera
  - ✓ *Vibrio parahaemolyticus* – gastroenteritis
  - ✓ *Vi vulfinicus* – wound infections

HEY WALDO-WITCHERS!  
SAW SOME TRULY TERRIFIC  
SIGHTS TODAY—SOMEONE  
BURNING TROUSERS WITH  
AN IRON; A LONG THIN MAN  
WITH A LONG THIN TIE;  
A GLOVE ATTACKING A MAN.  
PHEW! INCREDIBLE!



TO:  
WALDO-WITCHERS  
OVER THE MOON,  
THE WILD WEST  
NOW

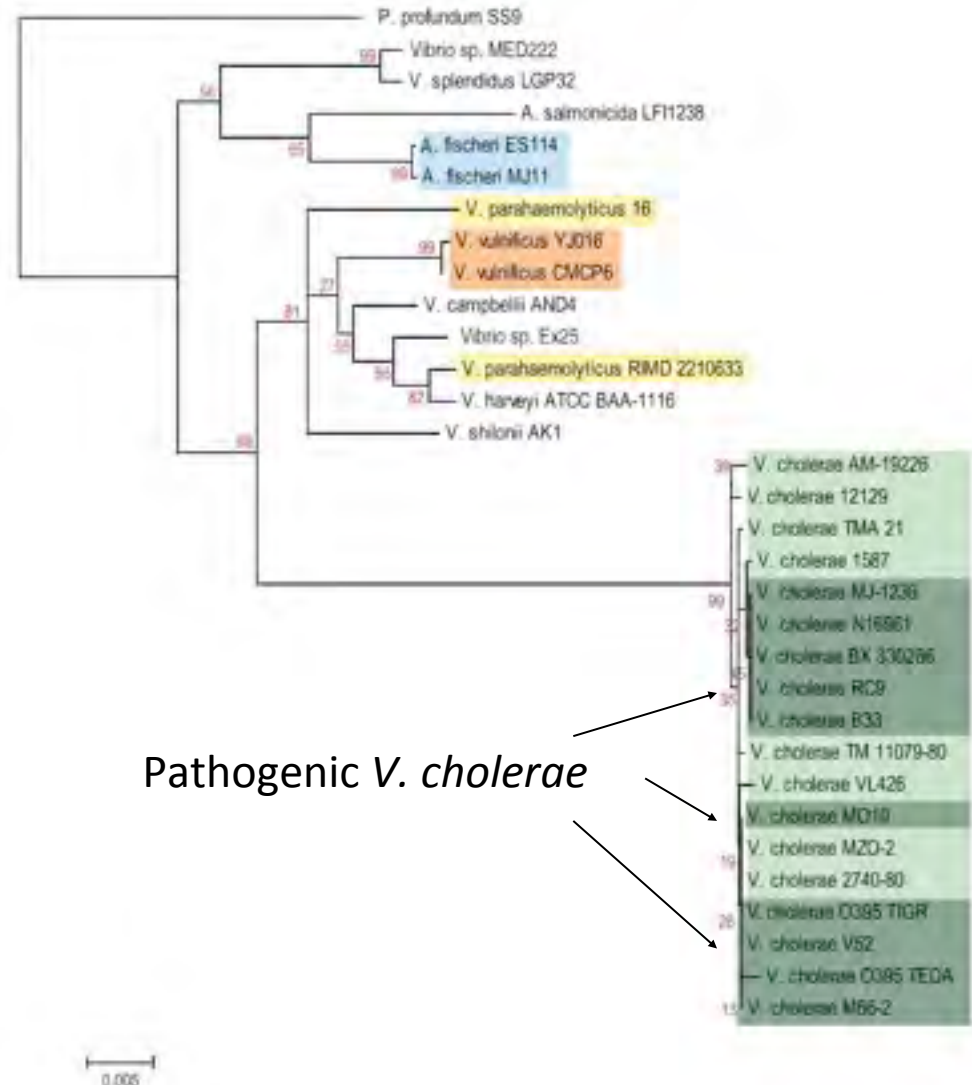
Waldo



# Phylogenetic trees

16s rRNA gene,  
extracted from 32 seq.

- Gene annotations: GenBank, EasyGen
- RNAmmer: score threshold >1.800
- Alignment: PRANK
- Phylogenetic tree: MEGA4, nj (neighbor-joining), Jukes-Cantor, Bootstrap (500 samples)



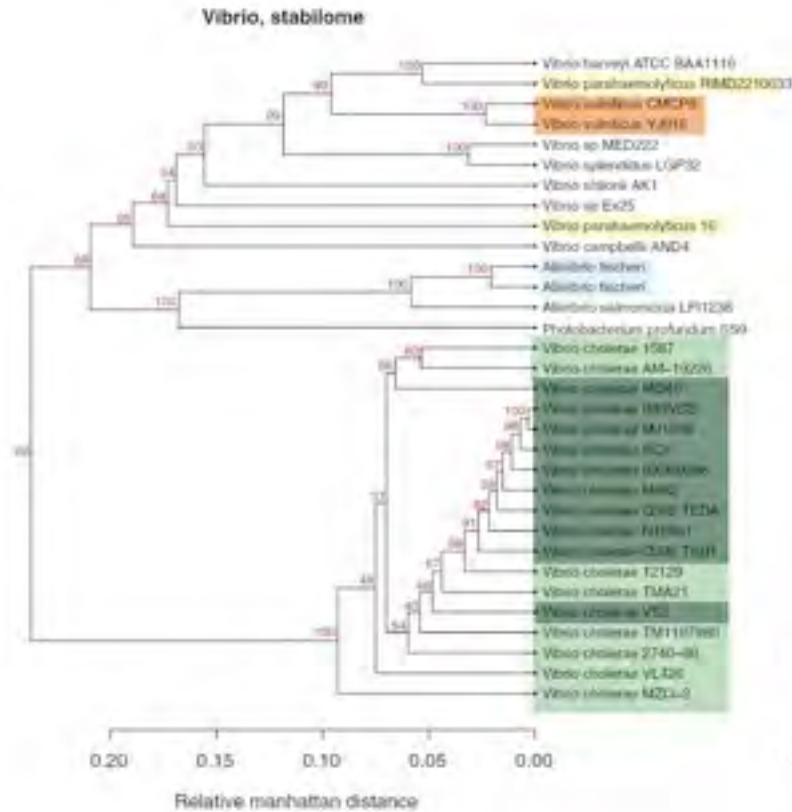
# Phylogenetic trees

More depth in the *V. cholera* cluster?

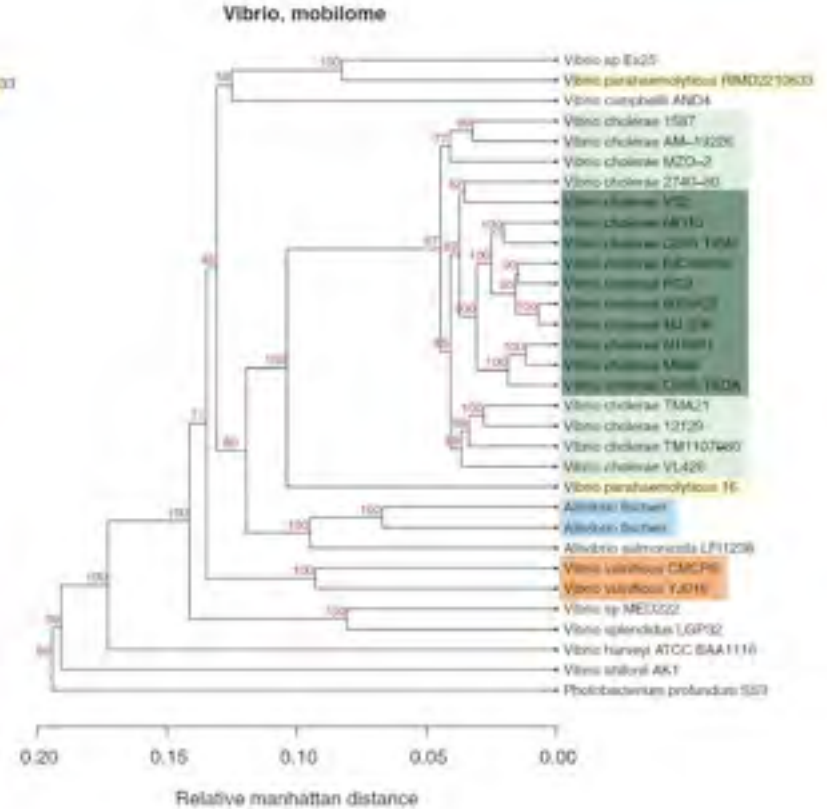
- Matrix based tree (scaled depending on no. of genomes)
- Genes (GeneFam.) found in most genomes – Stabilome
- Genes found in only few genomes – Mobilome
- BLASTP 50/50-rule (alignment produce 50% identity for at least 50% of the length of the longest gene)

	Genome1	Genome2	.....	Genome <i>j</i>
GeneFam. 1	1	1	.....	0
GeneFam. 2	0	1	.....	0
GeneFam. 3	0	1	.....	1
.....	.....	.....	.....	.....
GeneFam. <i>i</i>	0	0	.....	1

# Phylogenetic trees

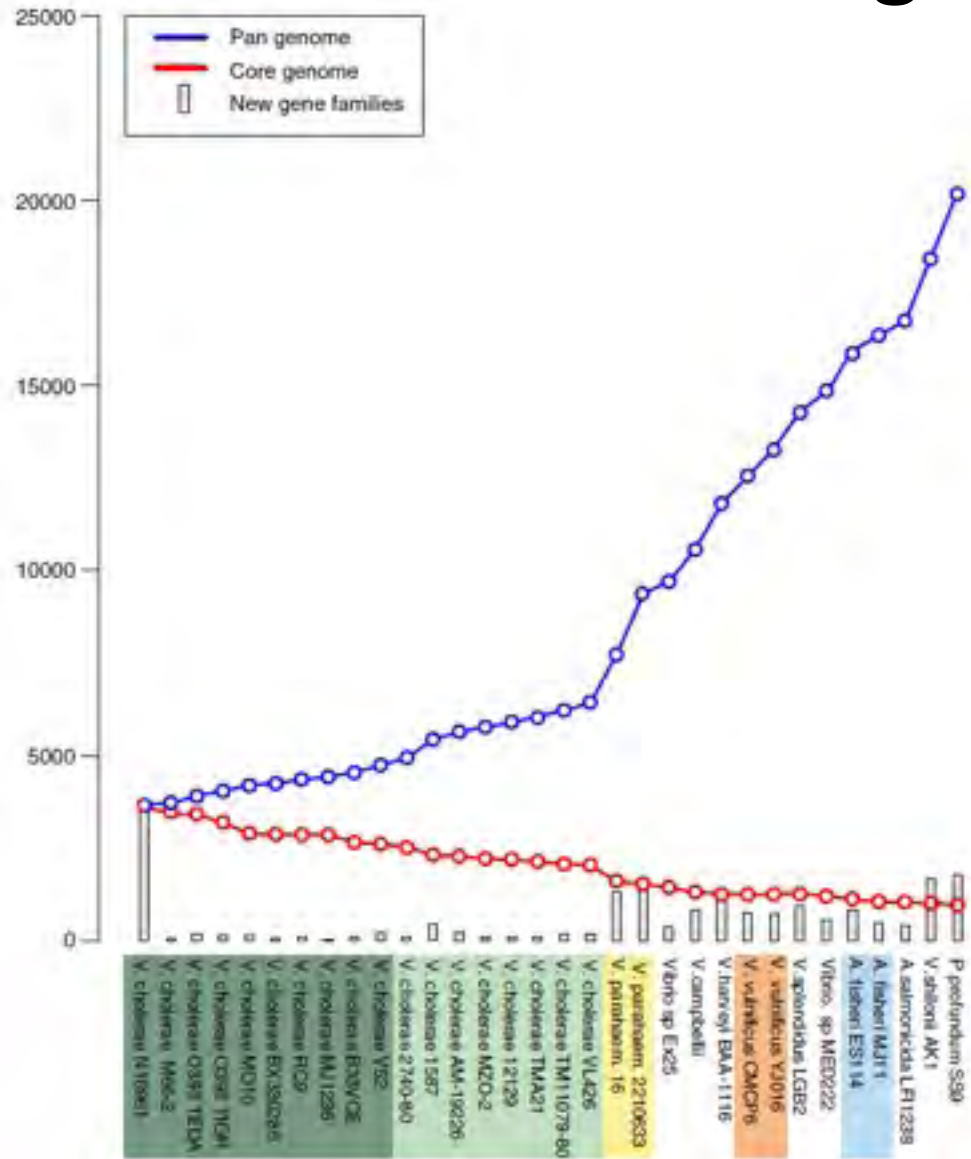


Genes present in at least 90%



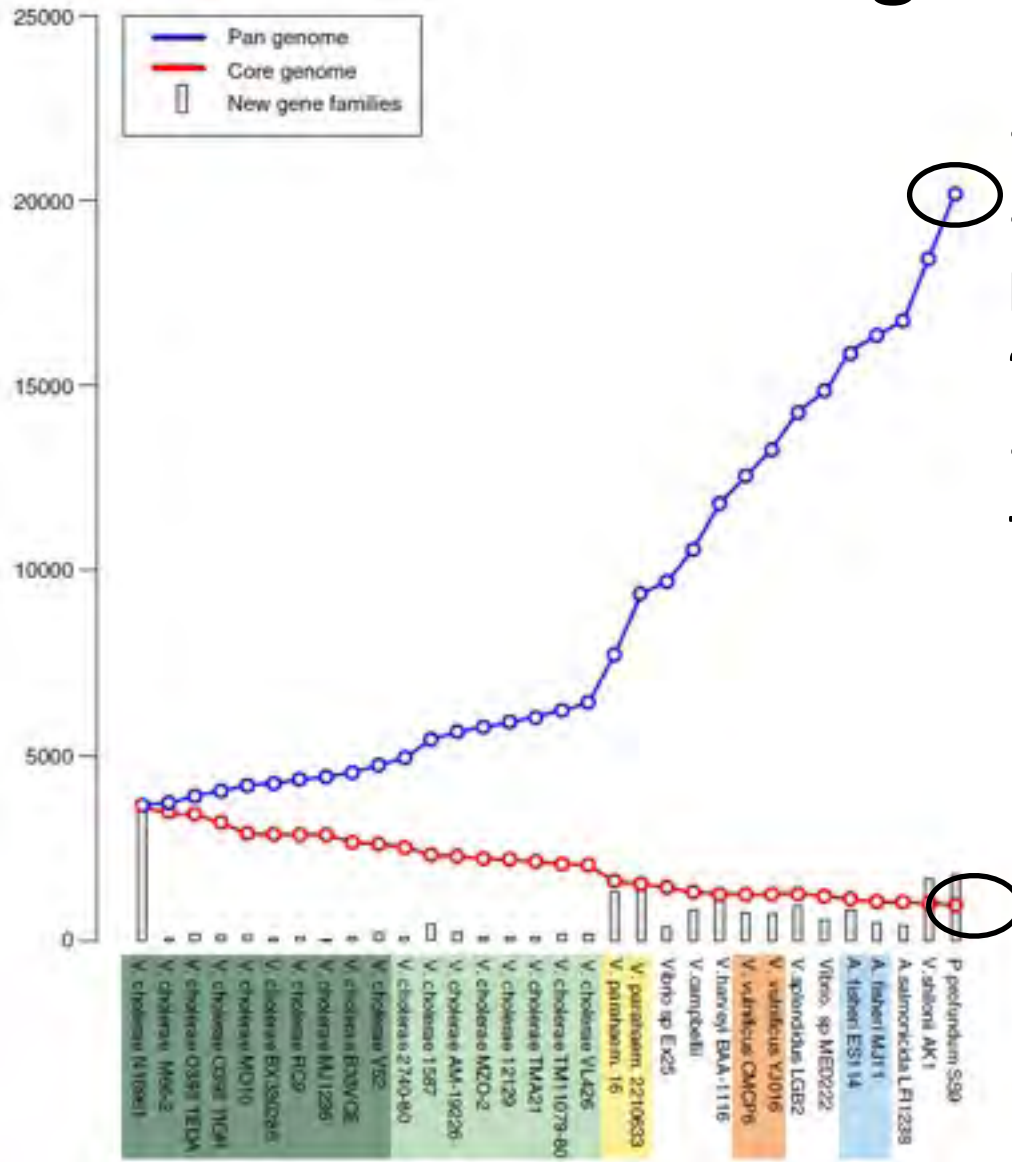
Genes found only in two to four genomes

# Pan- and core genome plot



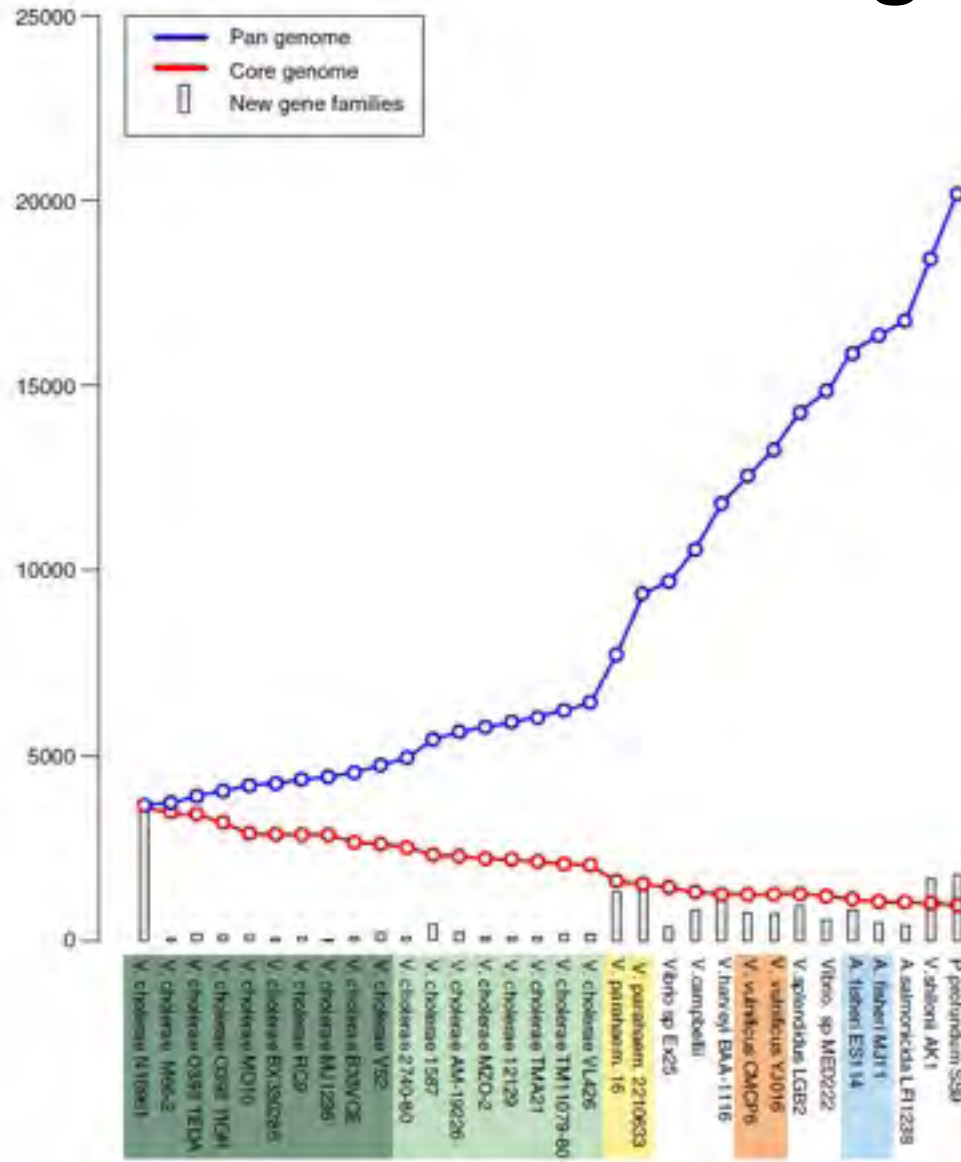
• BLAST analysis

# Pan- and core genome plot



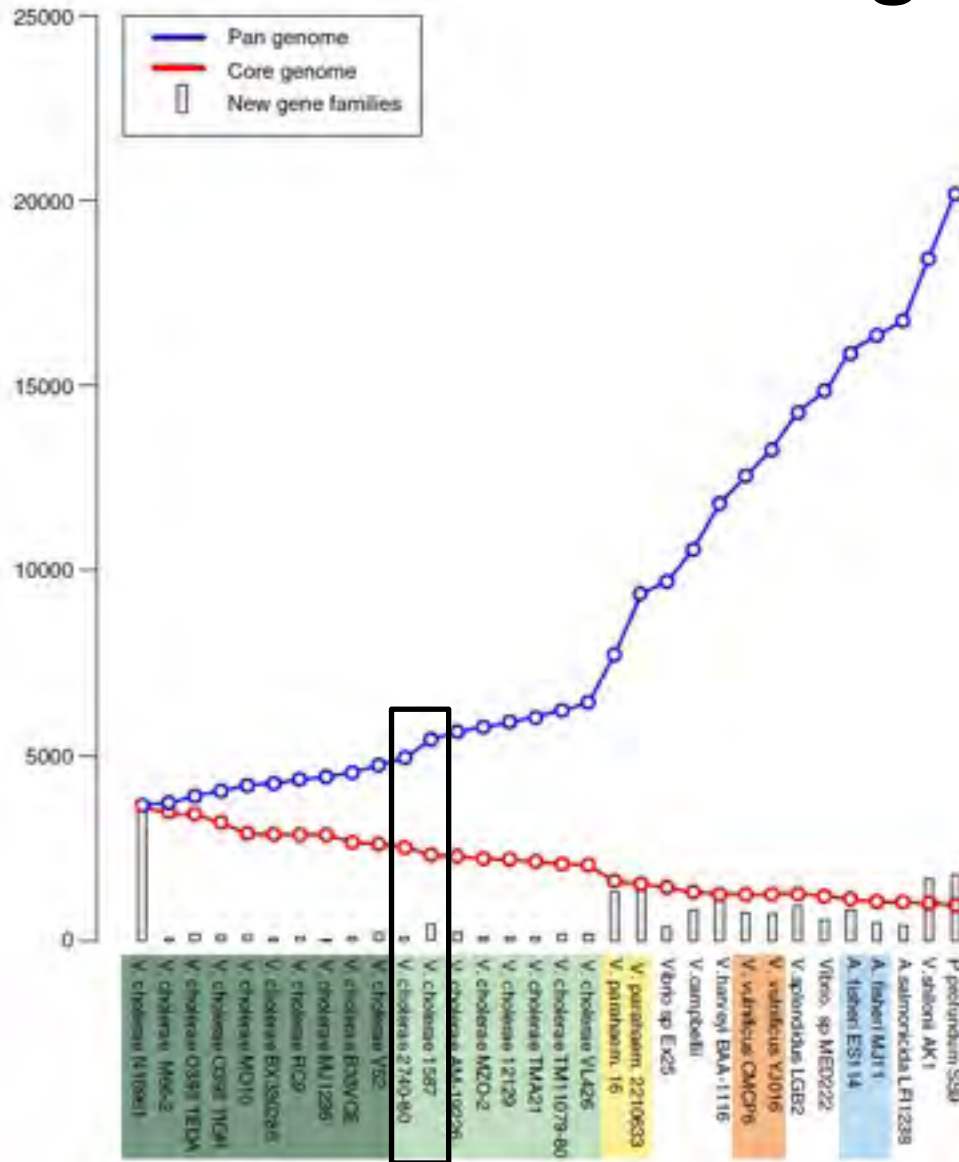
- BLAST analysis
- Pan genome: total hypothetical gene families. ~ 20,200 total gene families
- Core genome: conserved gene families. ~ 1,000 gene families

# Pan- and core genome plot



- BLAST analysis
- Pan genome: total hypothetical gene families. ~ 20,200 total gene families
- Core genome: conserved gene families. ~ 1,000 gene families
- Pan genome plateau at **6,500 gene families**

# Pan- and core genome plot



- BLAST analysis
- Pan genome: total hypothetical gene families. ~ 20,200 total gene families
- Core genome: conserved gene families. ~ 1,000 gene families
- Pan genome plateau at 6,500 gene families
- Increase in pan genomes at *V. cholerae* 1587
- *V. cholerae* 2740-80 behaves atypical in all figures

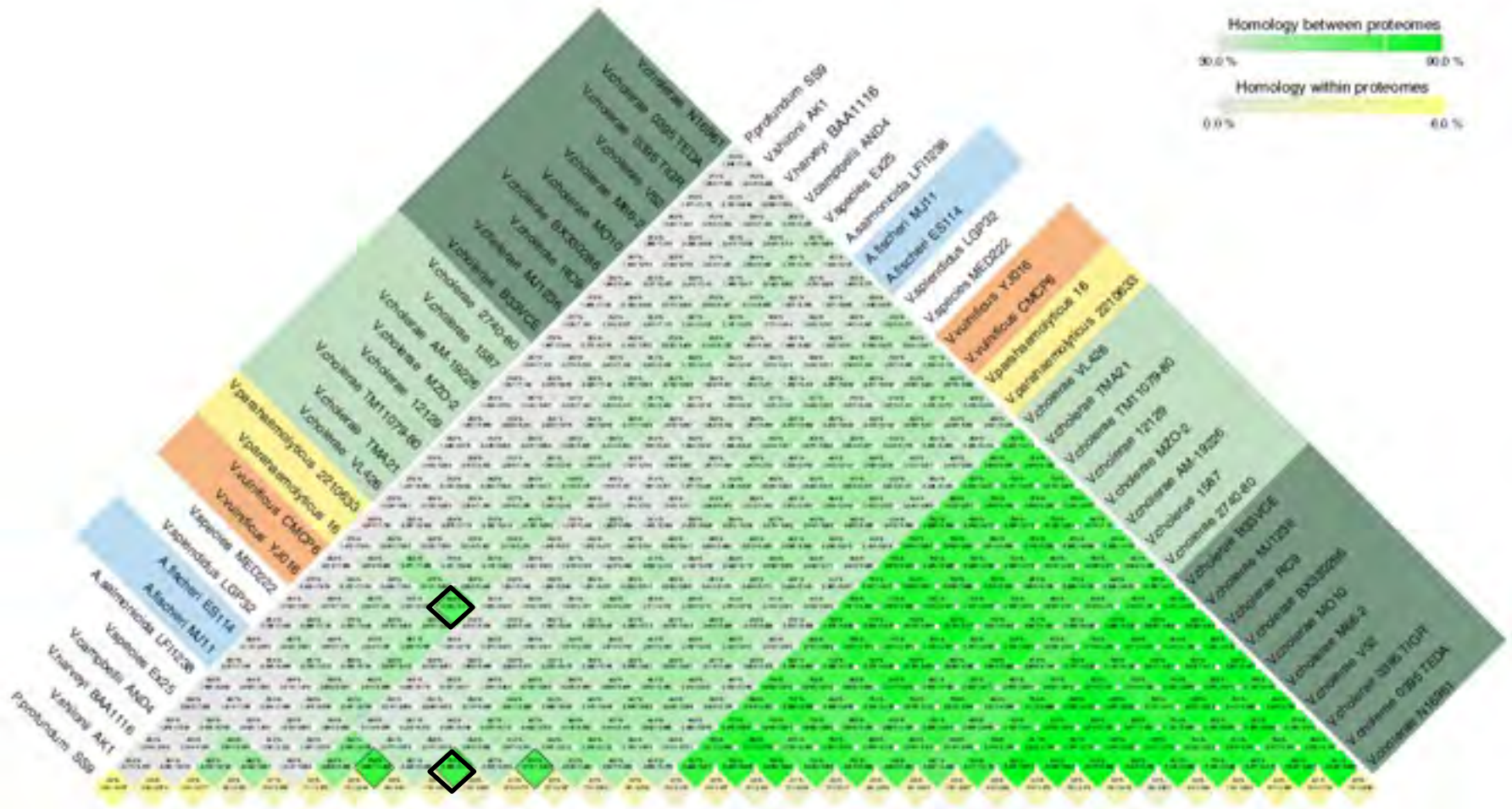




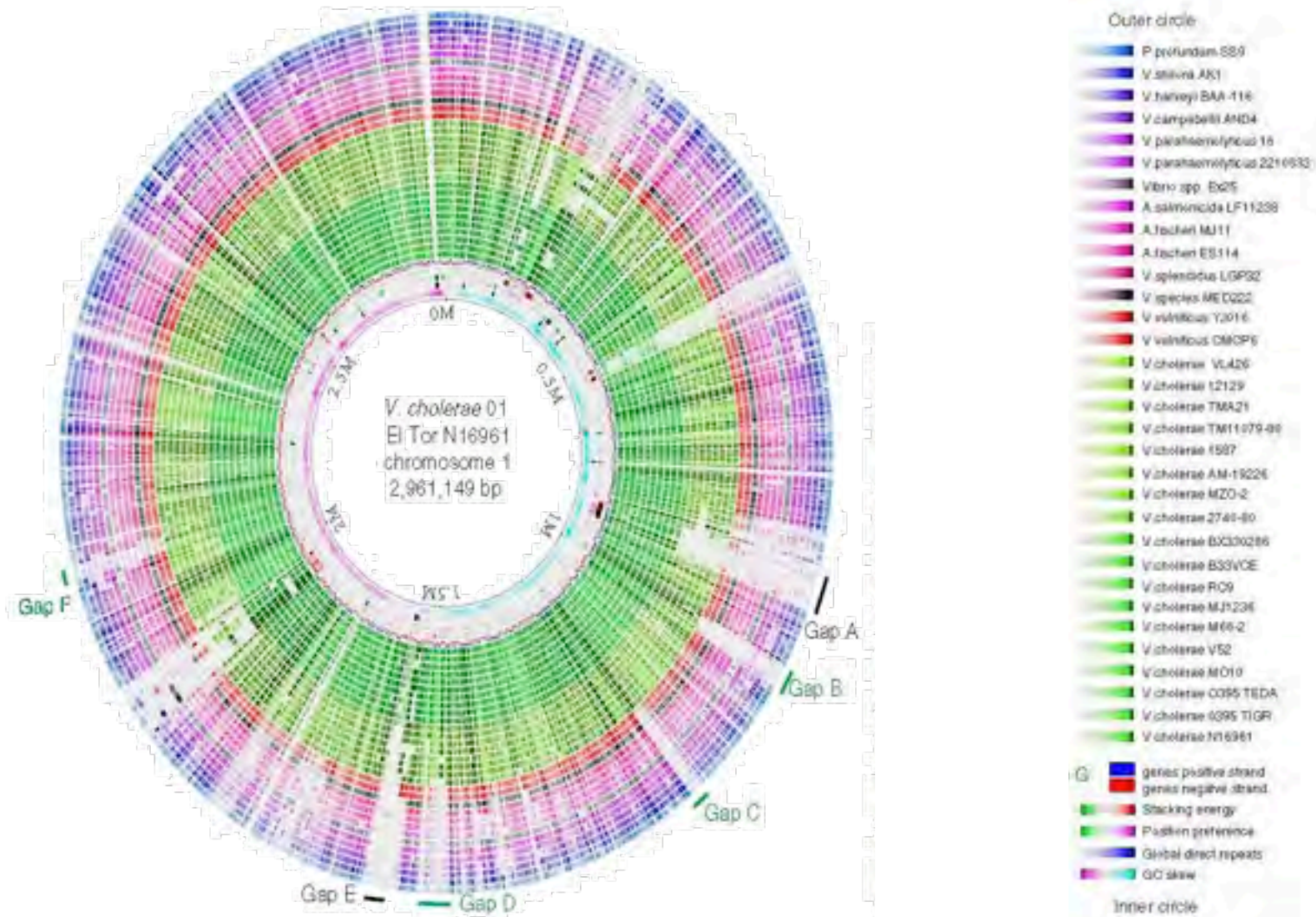




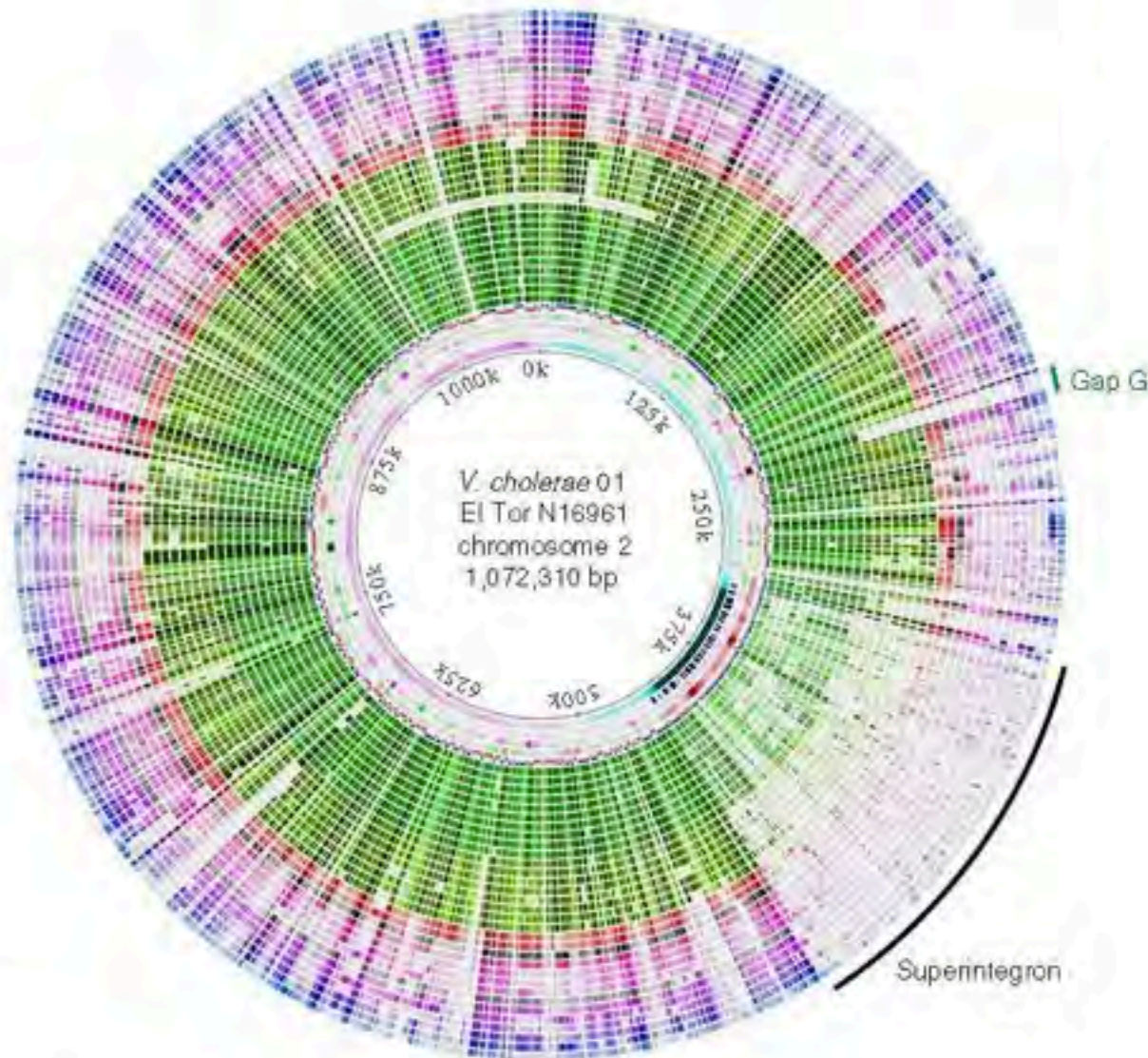
# BLAST matrix



# BLAST Atlas



# BLAST Atlas



Outer circle

- P. profundum SS9
- V. shankai AK1
- V. Harveyi BAA-116
- V. campbelli AND4
- V. parahaemolyticus 16
- V. parahaemolyticus 2210633
- Vibrio spp. Ex25
- A. salmonicida LF11238
- A. fischeri MJ11
- A. fischeri ES114
- V. splendidus LGP32
- V. species MED22
- V. vulnificus YJ016
- V. vulnificus CMOP6
- V. cholerae VL426
- V. cholerae 12129
- V. cholerae TMA21
- V. cholerae TMT 1079-80
- V. cholerae 1597
- V. cholerae AM-18226
- V. cholerae NZO-2
- V. cholerae 2740-80
- V. cholerae BX330286
- V. cholerae B33VCE
- V. cholerae RC9
- V. cholerae MJ1236
- V. cholerae M68-2
- V. cholerae V52
- V. cholerae MD10
- V. cholerae CG95 TEDA
- V. cholerae 0395 TIGR
- V. cholerae N16961

- G genes positive strand
- genes negative strand
- Stacking energy
- Position preference
- Global direct repeats
- GC skew

Inner circle

## Conclusion

- Better resolution with the matrix based method compared to 16s rRNA within *Vibrio cholerae* cluster.
- Observed difference between clinical and environmental strains due to genes related to virulence.
- In *V. cholerae* two-component signal transduction regulatory system was found.

## Future Perspectives

- Possibly, in some time *V. cholerae* would become distinct species
- It is still to be confirmed *V. cholerae* 2740-80 lost the CTX $\Phi$  prophage, rather than being the precursor of pathogen
- Be critical to the 16S rRNA trees when looking into the species, they may not be accurate enough

Dziękuję za uwagę

GRACIAS POR SU ATENCION!

Tak for jeres opmærksomhed

**THANK YOU FOR YOUR ATTENTION**

