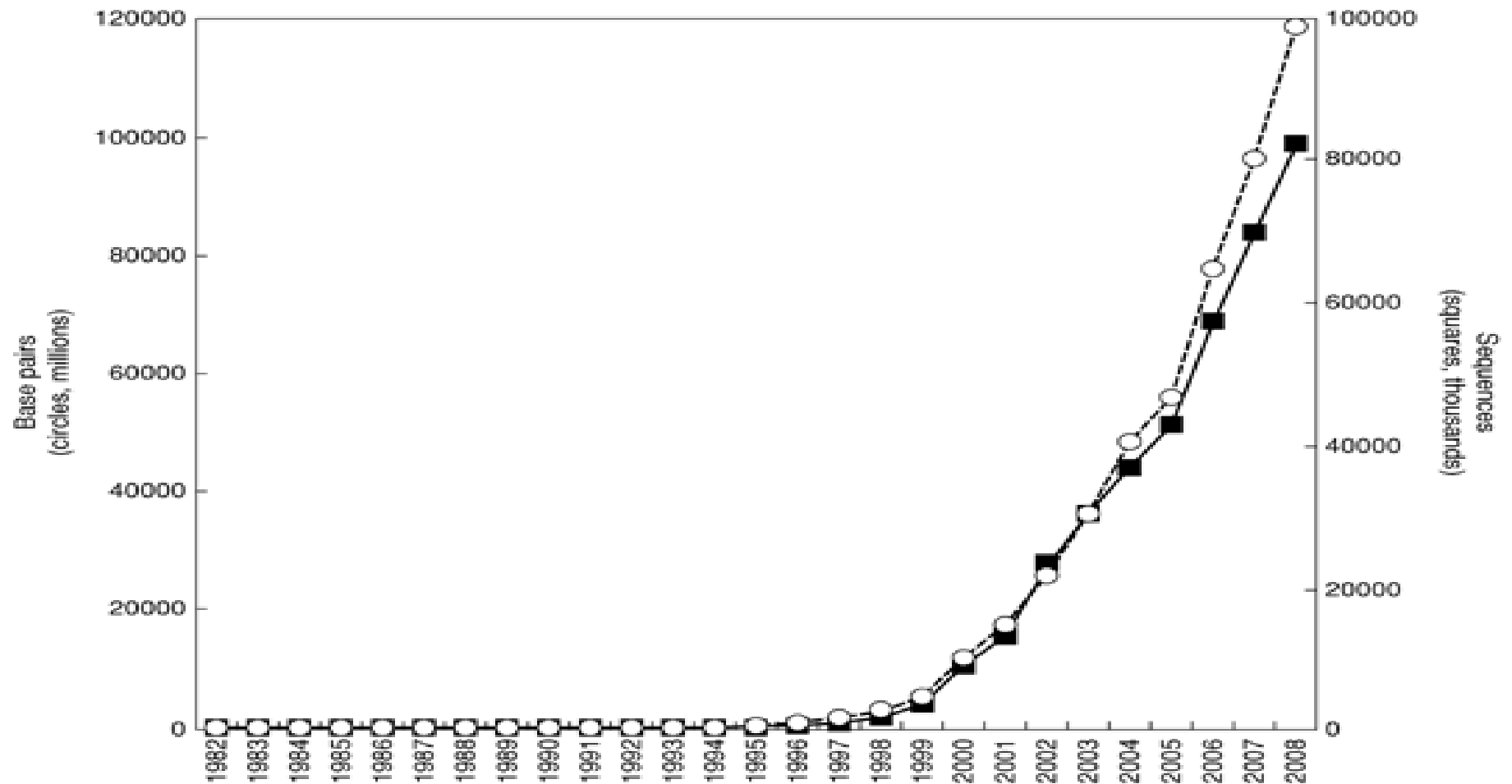# In search of RNase P RNA

Dilmurat Yusuf
Manja Marz
Peter Stadler
Ivo Hofacker

Bled
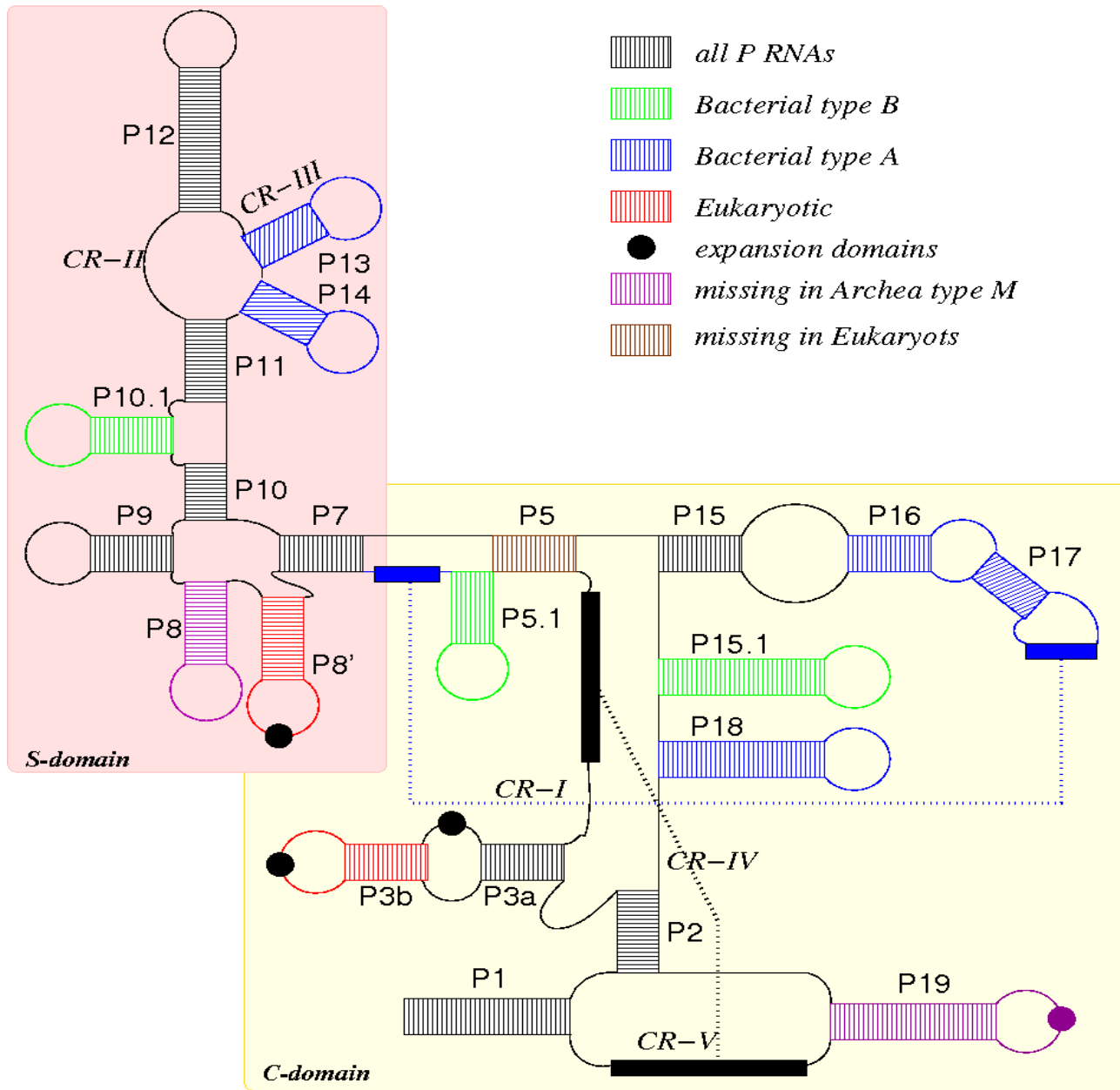2010

Outline

- Background

- Method

  - Bcheck

- Results

- Bcheck webserver

- Conclusion

Background

Exponential growth of database

RNase P RNA (rnpB)

Legend:
- all P RNAs
- Bacterial type B
- Bacterial type A
- Eukaryotic
- expansion domains
- missing in Archea type M
- missing in Eukaryots

S-domain

C-domain

Structure labels: P12, CR-III, CR-II, P13, P14, P11, P10.1, P10, P9, P7, P5, P15, P16, P17, P8, P8', P5.1, P15.1, P18, CR-I, CR-IV, P3b, P3a, P2, P1, CR-V, P19

Element	Bact.		Arch.		Euk.
	A	B	A	M	P
P1	●	●	●	●	●
P2	●	●	●	●	●
P3a	●	●	●	●	●
P3b					●
P5	●	●	●		●
P5.1					
P7	●	●	●		●
P8	●	●	●		●
P8'					●
P9	●	●	●	●	●
P10	●	●	●	●	●
P11	●	●	●	●	●
P10.1		●			
P12	●	●	●	●	●
P13	●				
P14	●				
P15	●	●	●	●	●
P16	●		⌣●		
P17	●				
P15.1		●			
P18	●	●			
P19	●	●	●		●
P4	●	●	●	●	●
P6	●		●		
CR-I	●	●	●	●	●
CR-II	●	●	●	●	●
CR-III	●	●	●	●	●
CR-IV	●	●	●	●	●
CR-V	●	●	●	●	●

Method
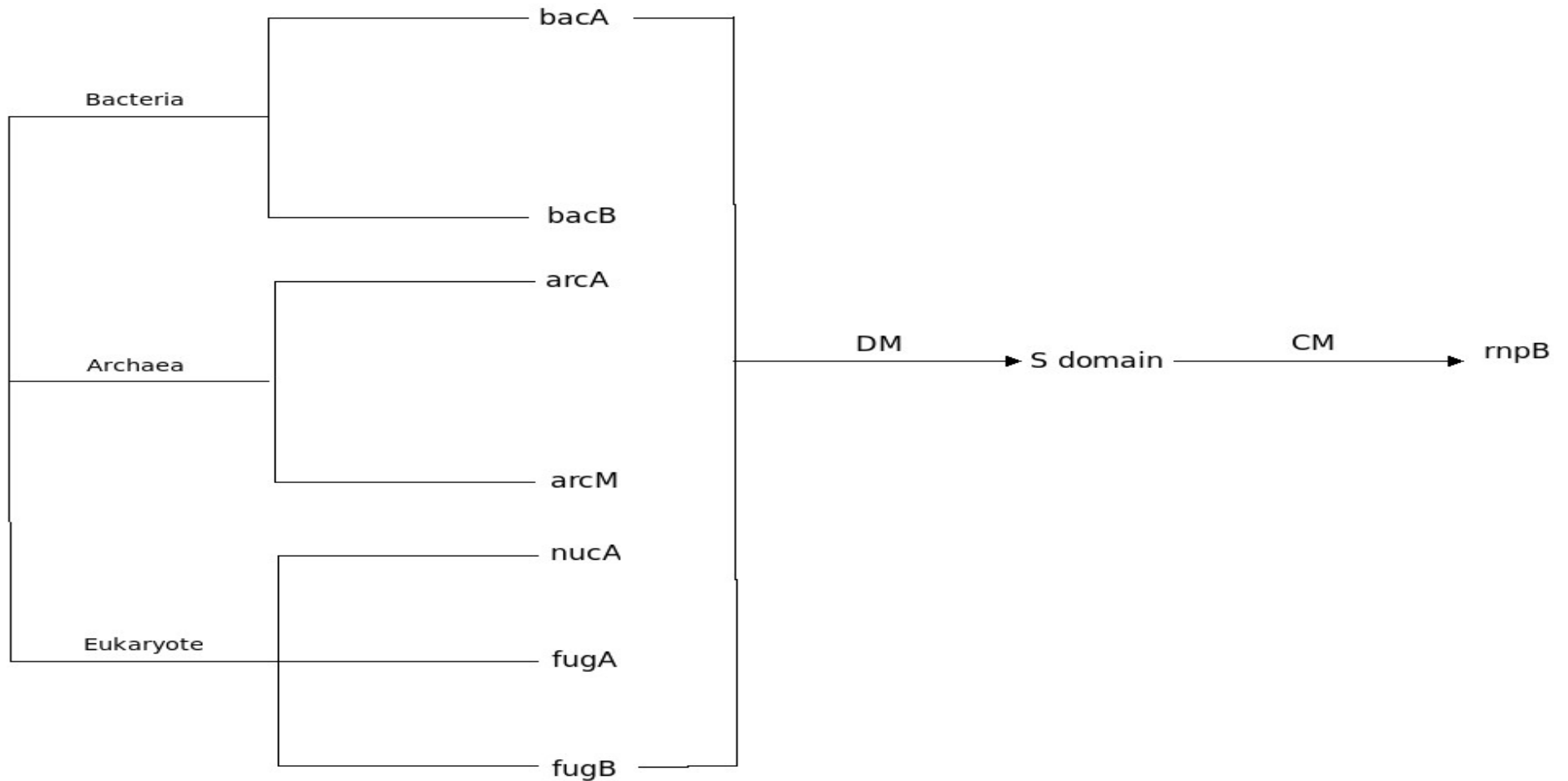
Bcheck summary: RNABOB + INFERNAL

Results

Predictions for microb & eukaryote

Summary of microbial rnpB predictions

	known	unknown	Sum
Bacteria	365 (365)	581 (591)	946 (956)
Archaea	22 (22)	37 (46)	59 (68)
Sum	387 (387)	618 (637)	1005 (1024)

Summary of eukaryotic rnpB predictions

	known	unknown	Sum
Metazoans	84 (85)	4 (13)	88 (98)
Fungi	19 (21)	32 (49)	51 (70)
Heterokonts	0 (0)	0 (6)	0 (6)
Plants	0 (0)	0 (30)	0 (30)
Other Eukaryots	7 (12)	1 (21)	8 (33)
Sum	110 (118)	37 (119)	147 (237)

Predictions for GOS metagenome

Environmental Genomics

Traditional Microbial Genomics

Environmental Genomics
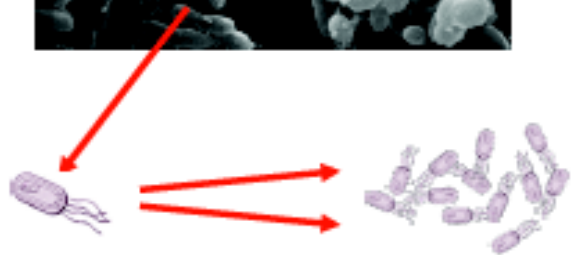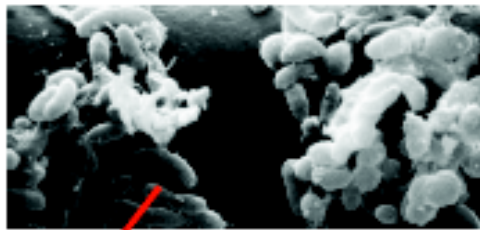
isolate individuum   clonal growth

prepare DNA, sequence randomly
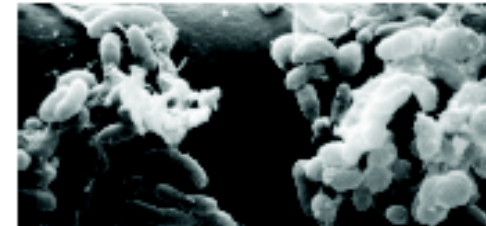
prepare DNA, sequence randomly

assemble genome

assemble partial genomes

Global Ocean Sampling Expedition (GOS)

- Access marine genetic diversity
- Over 10 million sequences with size of 18 Gbp
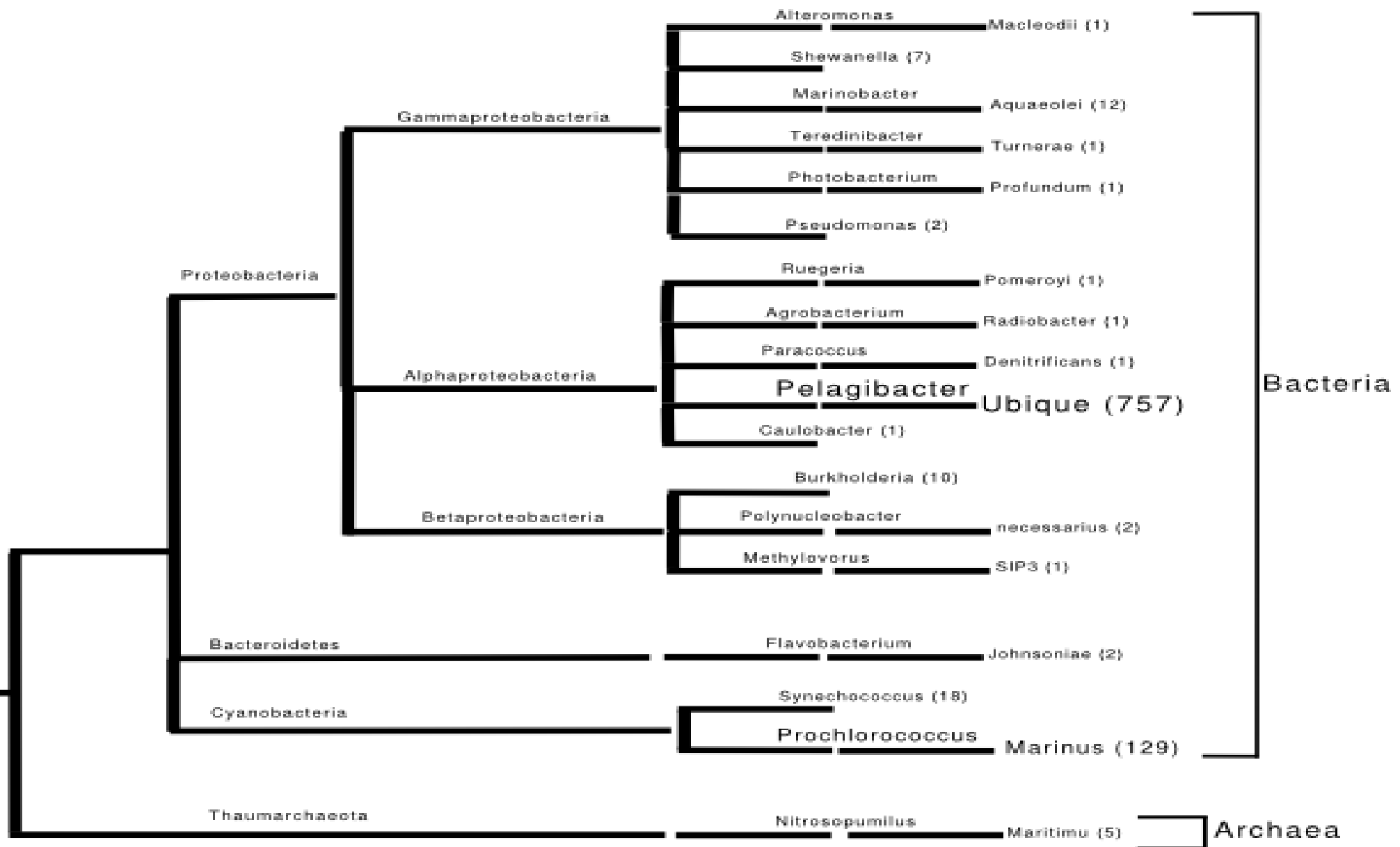
Predictions for GOS metagenome

	arcM	arcA	bacB	bacA	total
GenBank	10	45	101	621	777
GOS	0	49	3	2857	2909

4675 predictions, 1 hit per sequence.

rnpB as marker: Assignment statistics

No hits (66%)

other_assignment (2%)

genus_assignment (1%)

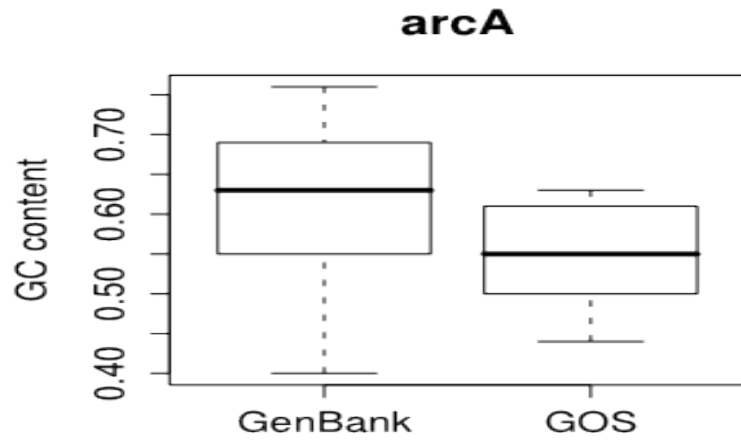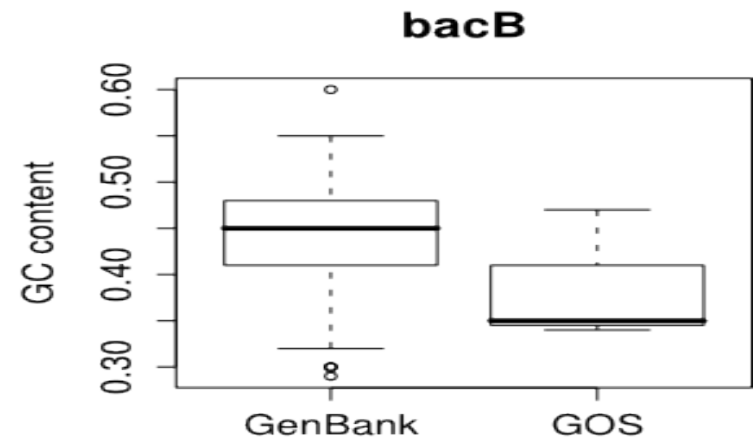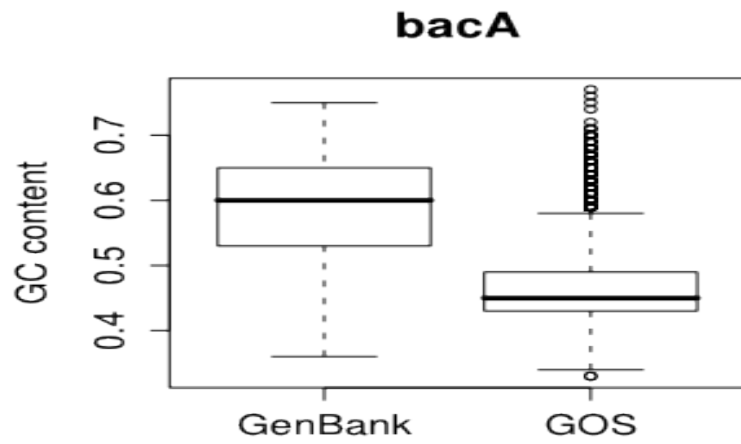species_assignment (31%)
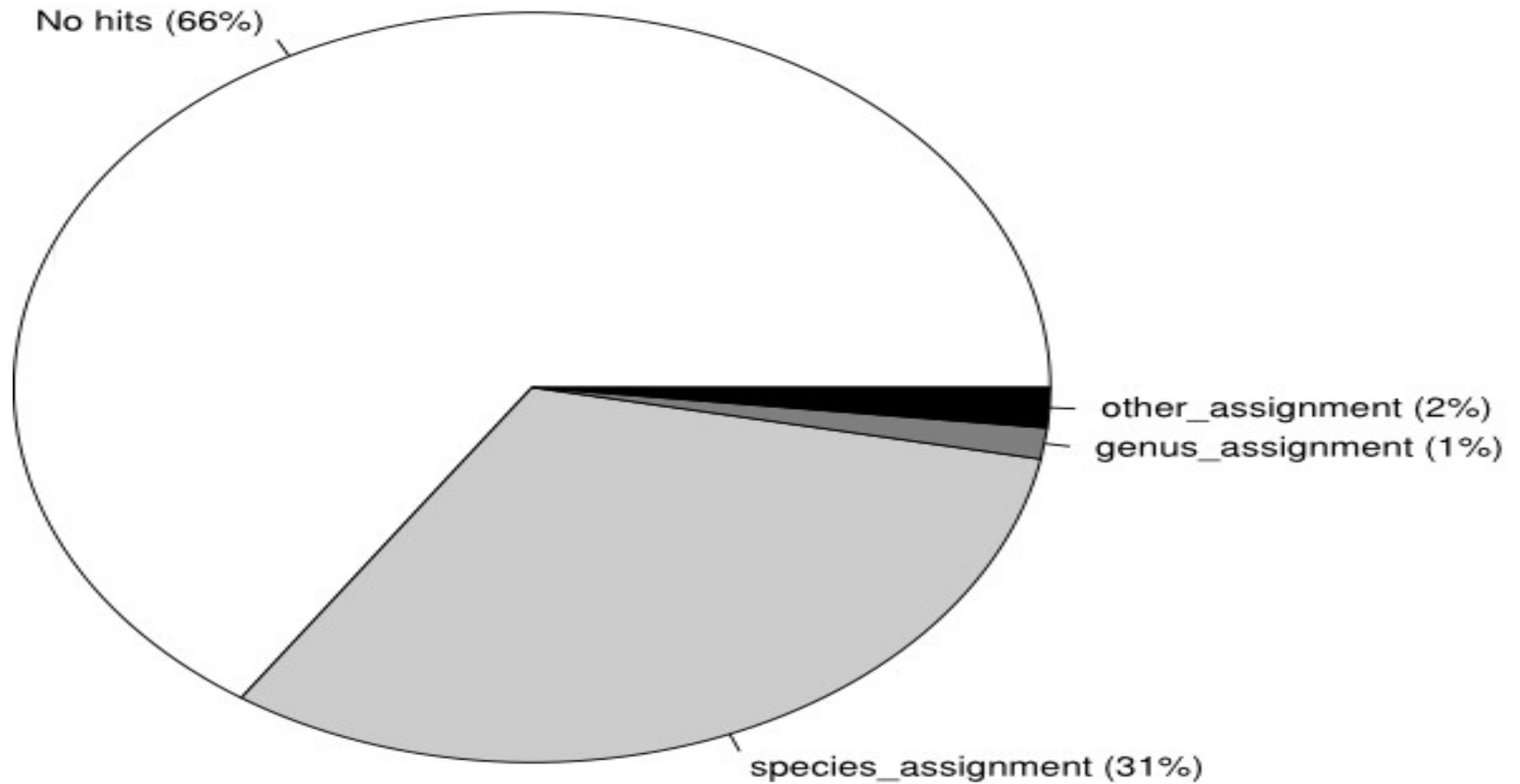
BLAST $(1e^{-50})$: GOS rnpB (2909) vs known rnpB (777)

- 1003 rnpB genes assigned to species of 18 genus
- Pelagibacter ubique: 75% of assignment
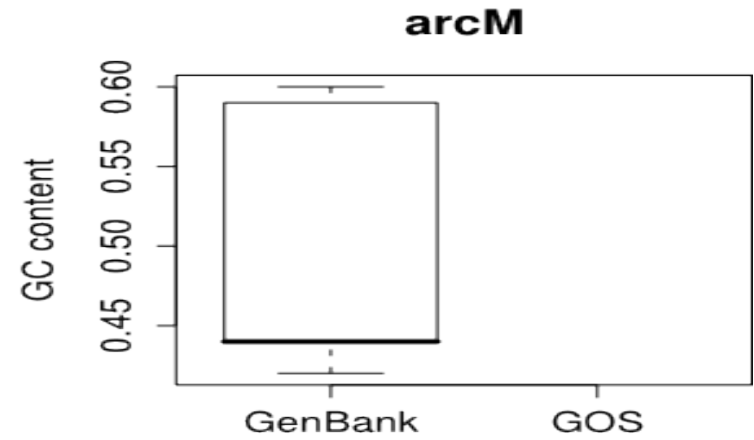
	arcA	arcM	bacA	bacB	total
Origin unknown	44	0	1859	3	1906

Bcheck
rnpB
server

Bcheck
Database
BLAST
Help

MySql database of rnpB genes predicted by Bcheck. Simply enter your query(s) and run *Search*. The size limit of query(s) is 200000 characters.

Expand All | Collapse All

📂 **rnpB**
    📁 **Archeae**
    📁 **Bacteria**
    📁 **Eukaryote**

📂 **GOS_rnpB**
    📁 **Archeae**
    📁 **Bacteria**

Download all rnpB genes in the database.

Search rnpB database:

Search GOS_rnpB database:

[Query Samples] [Clear Query]

[Search]

Conclusion

- ncRNA with strong variation

  - Sub family based approach

- The most abundant rnpB subfamily

  - bacA type

- No homolog found for 66% GOS rnpB

  - May indicate unknown species

- ncRNA prediction

  - Secondary structure based homology search