

全程硝化细菌（Comammox） 富集培养物的元基因组分析

12307110136 沈聪

背景介绍

- 全程硝化细菌（Comammox）

硝化过程整合：
氨氧化细菌/古菌

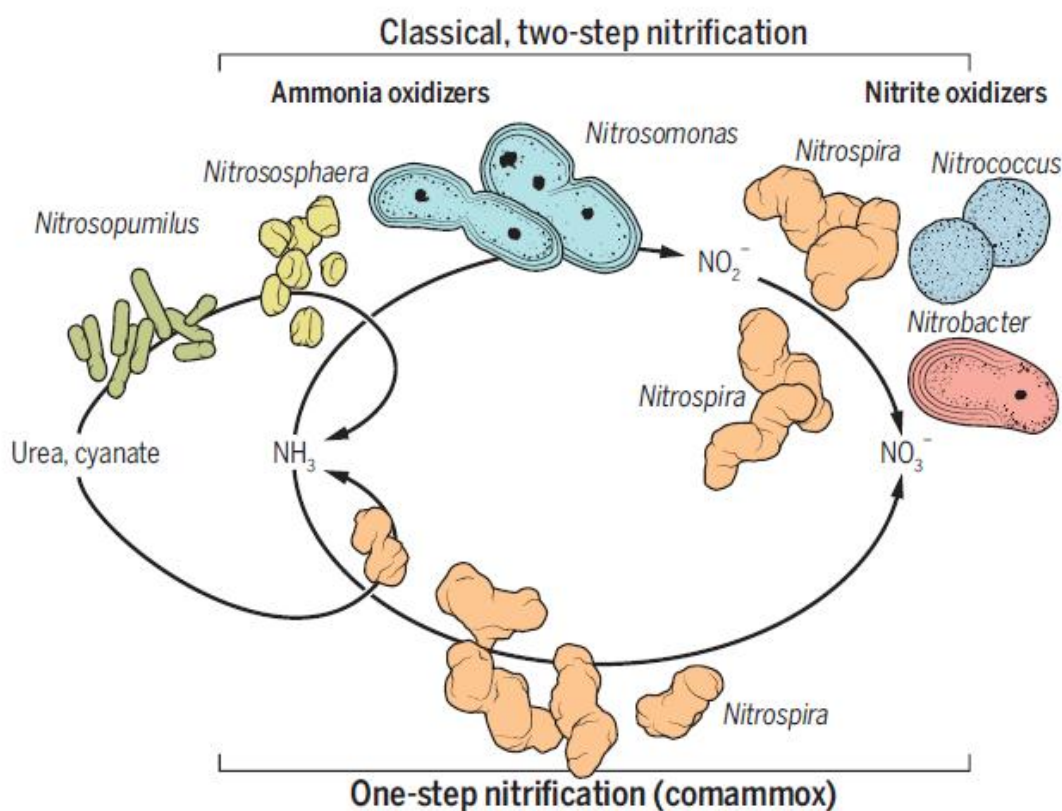
- 元基因组分析微

- 聚类：

- 序列组成性质：GC

- 测序丰度性质：覆盖
groups, CAGs;

- 聚类质量检验：单



本文分析元基因组的总体框架

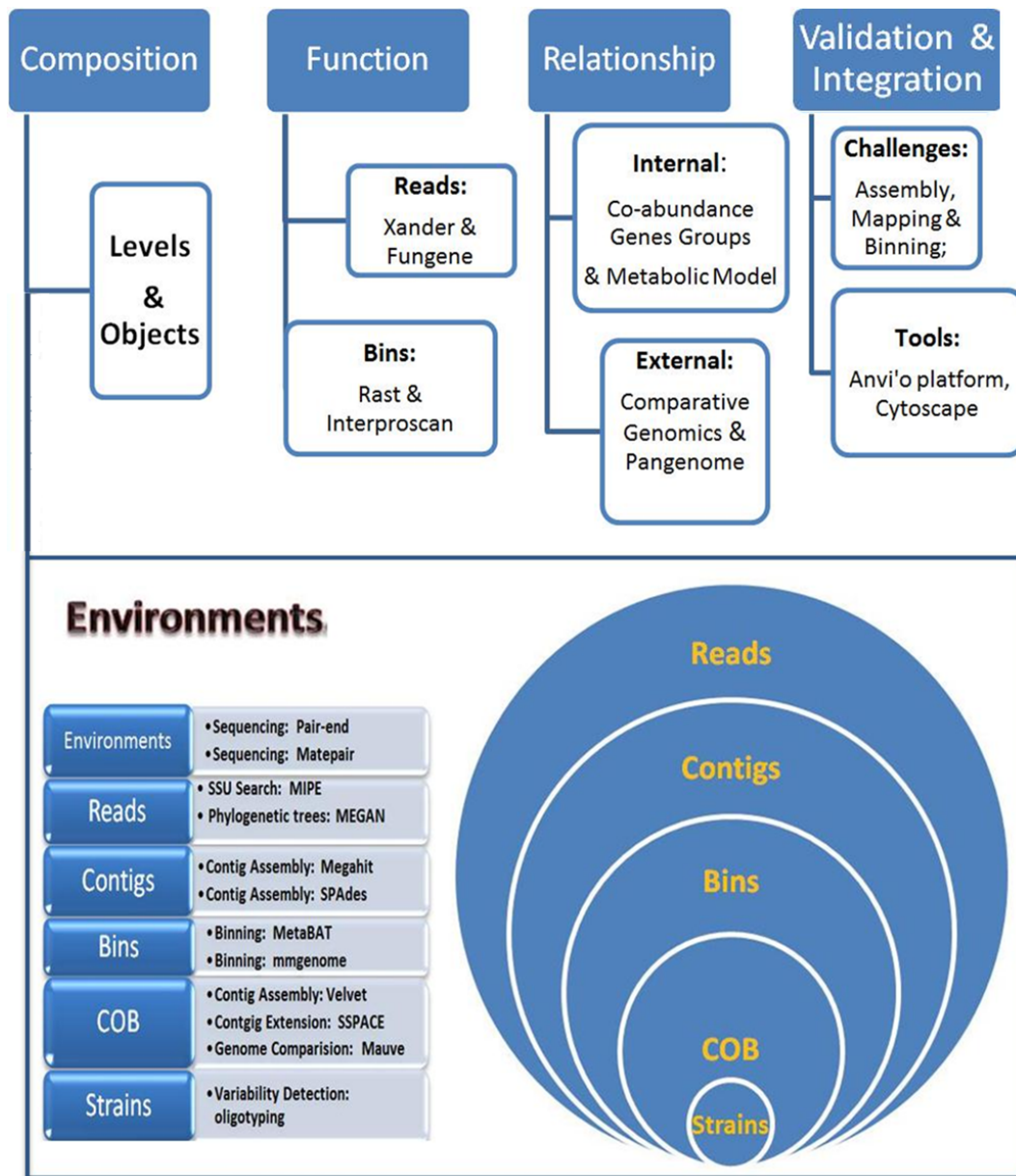
核心问题:

群落组成 (Composition)

功能活性 (Function)

互作关系 (Relationship)

验证整合 (Validation)



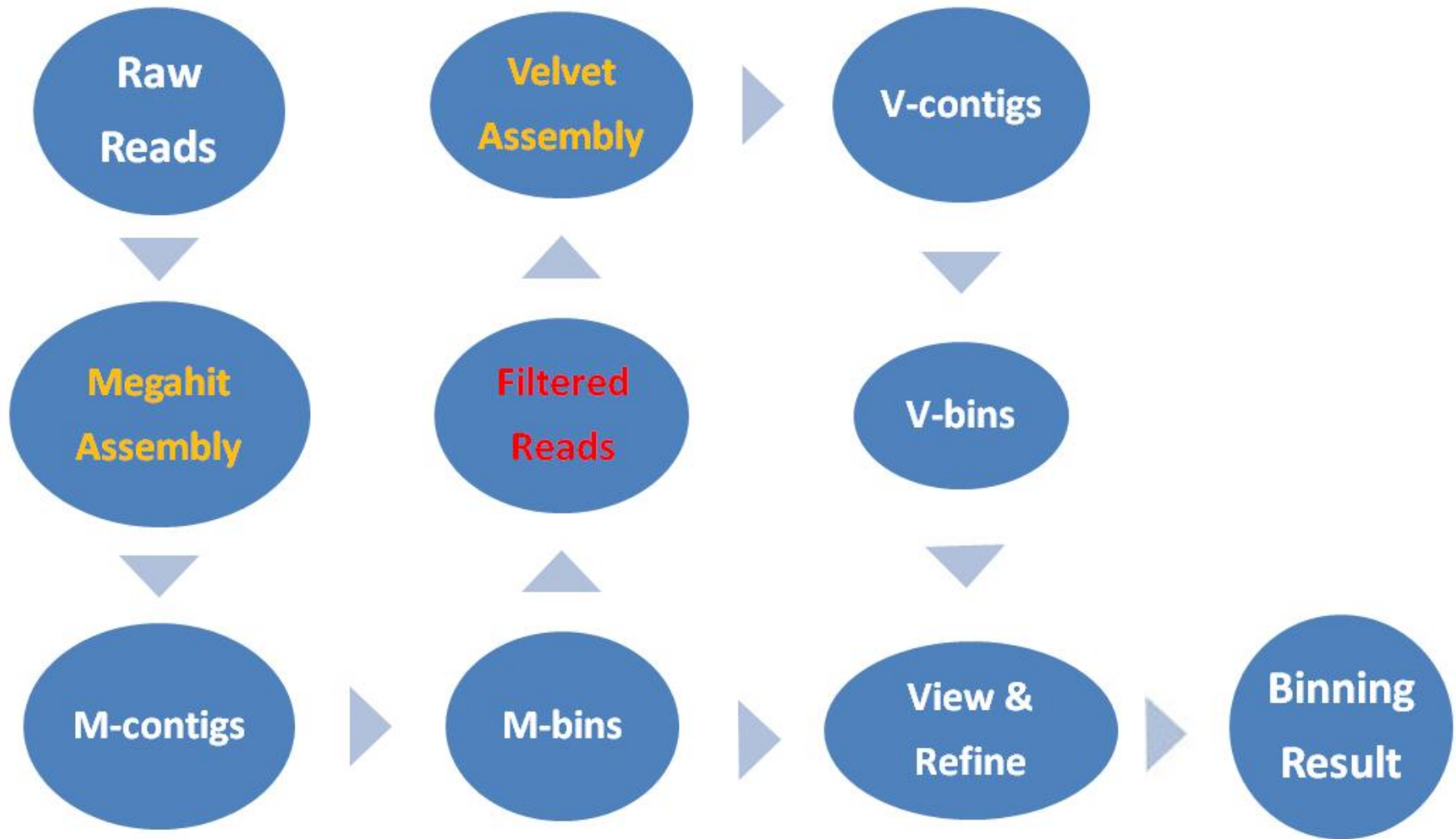
1.群落组成

1.1 测序序列层面（Reads）

Comammox富集培养样品的SSU rRNA门水平的分布表			
分类单元	匹配条数	分类单元	匹配条数
SILVA	0	Fibrobacteres	3
Bacteria	13294	Firmicutes	3767
Acidobacteria	216	Fusobacteria	112
Actinobacteria	1447	Gemmatimonadetes	20
Armatimonadetes	4	JL-ETNP-Z39	1
BD1-5	6	NPL-UPA2	1
Bacteroidetes	1234	Nitrospirae	523
Candidate division BRC1	3	Planctomycetes	134
Candidate division OD1	15	Proteobacteria	9866
Candidate division OP11	1	SHA-109	4
Candidate division SR1	3	SM2F11	1
Candidate division TM7	20	Spirochaetae	29
Candidate division WS3	5	TA06	5
Chlamydiae	10	TM6	2
Chlorobi	20	Tenericutes	19
Chloroflexi	136	Verrucomicrobia	48
Cyanobacteria	35	WCHB1-60	1
Deferribacteres	1	Eukaryota	44645
Deinococcus-Thermus	139	Not assigned	2
Elusimicrobia	13	No hits	1027

1.2 拼接结果层面(Contigs & Scaffolds)

1.2.1 聚类流程图

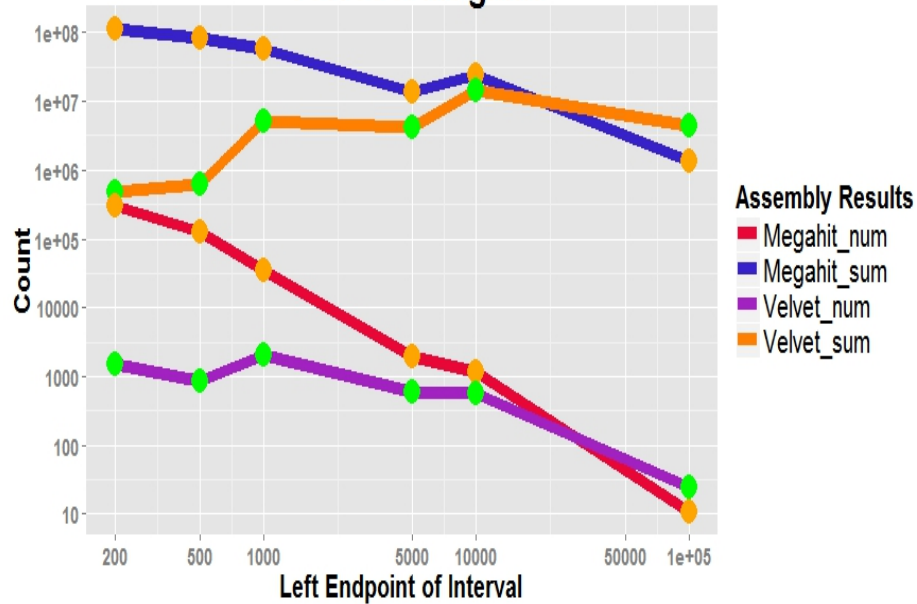


1.2 拼接结果层面(Contigs & Scaffolds)

1.2.2 聚类方法验证

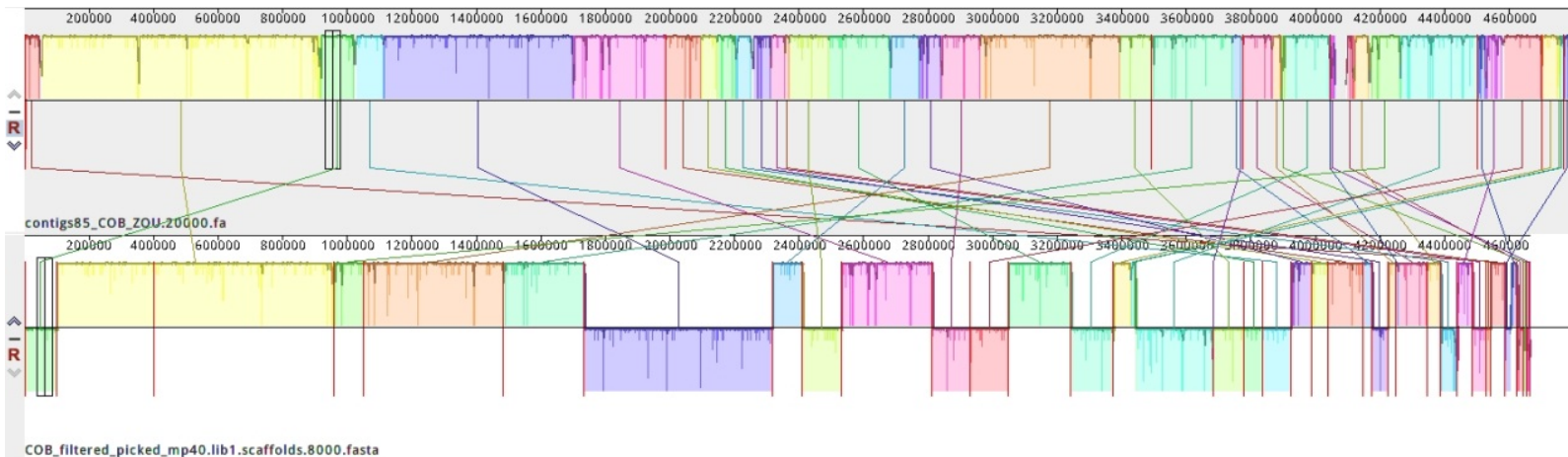
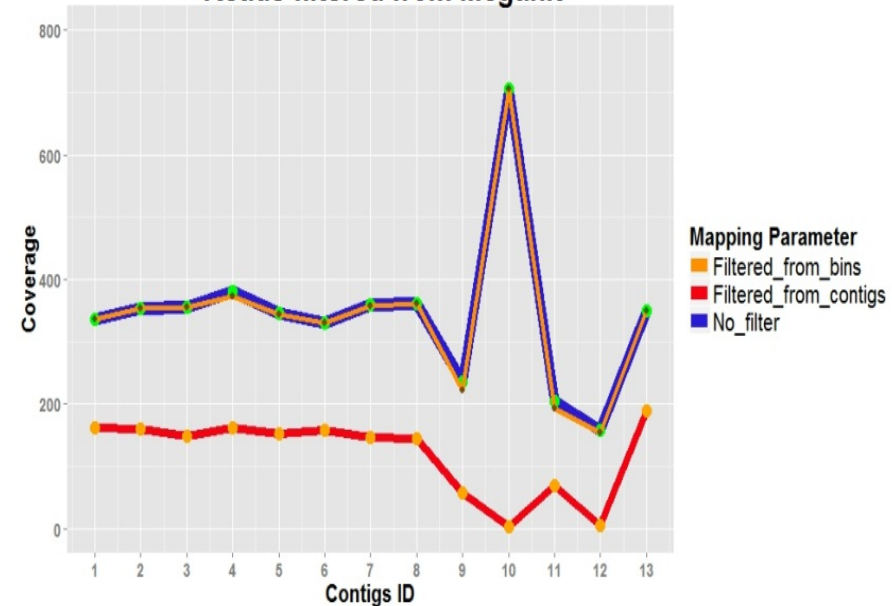
Comparison of contigs' number and total length

-- Velvet & Megahit



Coverage of Comammox Contigs

-- Reads filtered from Megahit



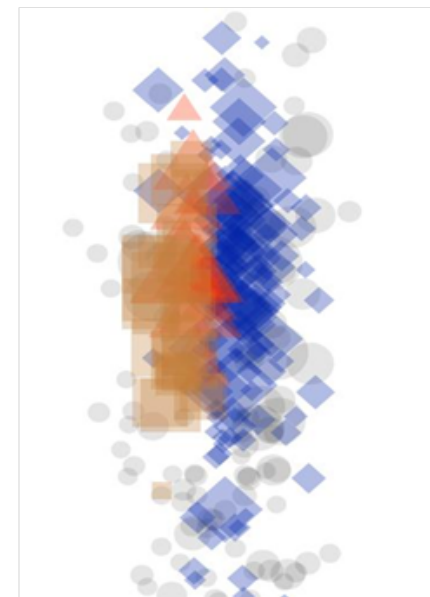
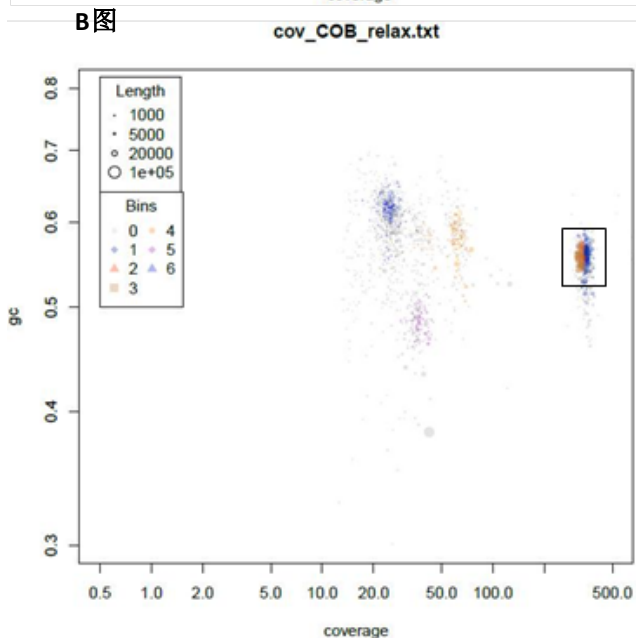
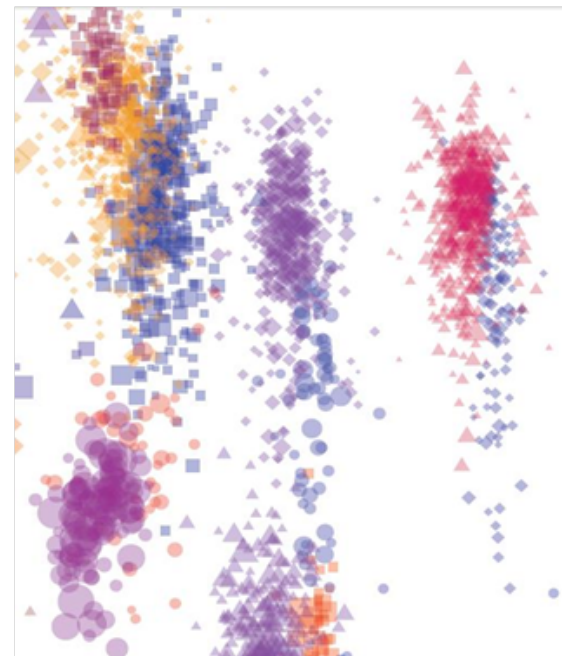
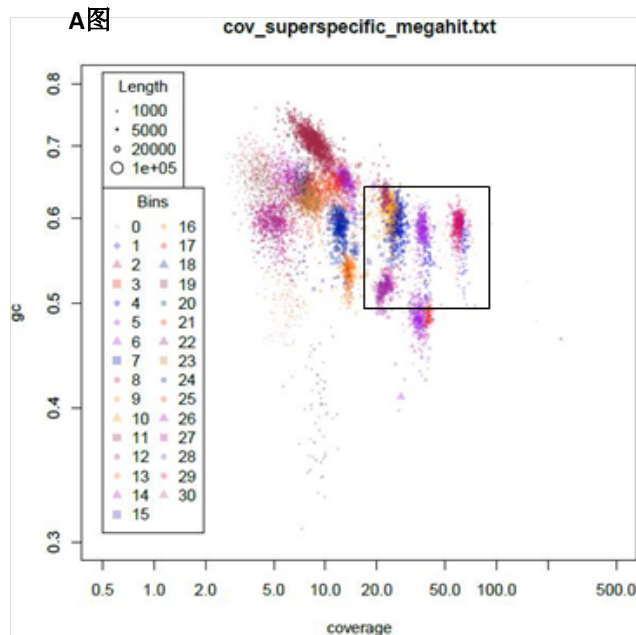
1.2 拼接结果 层面(Contigs & Scaffolds)

1.2.3 聚类修正

A图， Megahit v1.0
拼接结果；

B图， Velvet v1.2.01
对Megahit v1.0步骤
所滤出的序列的拼接
结果；

每张图左侧图为右侧
图中黑框的放大图



1.2 拼接结果 层面(Contigs & Scaffolds)

1.2.3 聚类修正

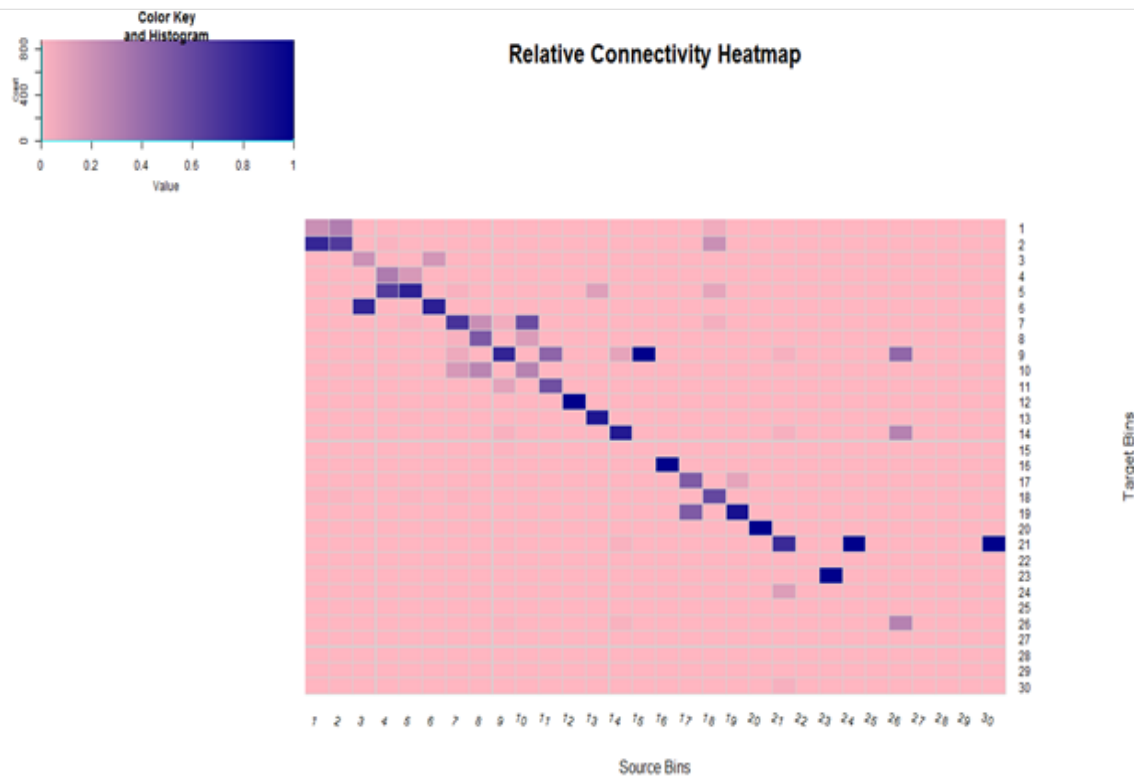
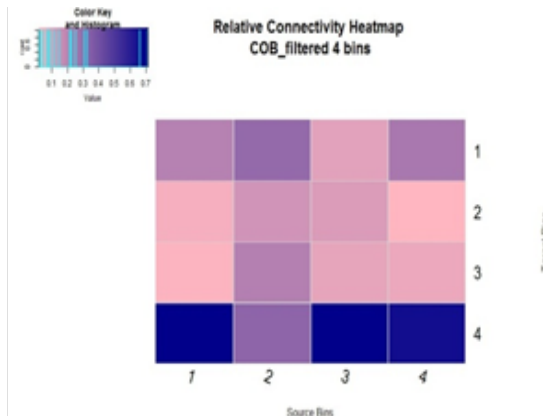
左上：Comammox聚类结果之间的连接值和相对连接值分布表

右上：Comammox聚类结果的相对连接值热图

下图：Megahit v1.0 30个聚类结果的相对连接值热图

Weighted Counts of COB binning				
Node1\Node2	Bin_0	Bin_1	Bin_2	Bin_3
Bin_0	7536	2863	187	785
Bin_1	2863	865	178	101
Bin_2	187	178	84	129
Bin_3	785	101	129	90
Sum	11371	4007	578	1105

Relative Connectivity of COB Binning				
Node1\Node2	Bin_0	Bin_1	Bin_2	Bin_3
Bin_0	66.27%	71.45%	32.35%	71.04%
Bin_1	25.18%	21.59%	30.80%	9.14%
Bin_2	1.64%	4.44%	14.53%	11.67%
Bin_3	6.90%	2.52%	22.32%	8.14%
Total	100.00%	100.00%	100.00%	100.00%



1.2 拼接结果层面(Contigs & Scaffolds)

1.2.4 修正后的聚类结果

Comammox富集培养样品聚类结果汇总

Bins	Taxon	Total length(bp)	Num of contigs	N50	GC content	Completeness(%)	Redundancy(%)	Coverage	Variability
Bin 0	Nitrospira defluvii	4804323	13	1503535	55.43	97.26	6.13	352.75	5.88
Bin 1	Nitrospira defluvii	858186	140	7080	57.36	3.10	0.00	57.62	0.16
Bin 2	Nitrospira defluvii	2461616	409	6999	59.12	37.71	3.60	53.34	0.16
Bin 3	Unknown(Bacteroidetes;Chitinophagaceae, 科)	360266	45	13282	48.83	2.65	0.00	35.36	1.90
Bin 4	Nitrospira defluvii	640098	72	11741	54.24	4.34	0.00	35.21	0.23
Bin 5	Nitrospira defluvii	3516854	413	11628	58.46	61.26	8.48	32.71	0.30
Bin 6	Unknown	2251741	288	10163	48.24	65.70	2.21	31.03	2.04
Bin 7	Nitrospira defluvii	2200054	351	7363	59.20	45.55	4.00	22.51	1.05
Bin 8	Unknown (Nitrospira, 属)	384793	46	12219	53.11	2.30	0.00	21.37	0.40
Bin 9	Bradyrhizobium japonicum	3045532	564	5861	60.97	54.74	4.70	19.89	1.84
Bin 10	Nitrospira defluvii	557864	88	7581	60.60	3.82	0.63	19.98	0.90
Bin 11	Bradyrhizobium japonicum	462360	98	5180	63.09	10.42	0.00	18.95	1.83
Bin 12	Unknown (Flavobacteriaceae, 科)	4168615	113	54165	51.61	94.59	4.59	20.18	0.09
Bin 13	Nitrosospira multiformis	3351085	136	35291	53.42	97.79	4.67	12.44	0.04
Bin 14	Sphingopyxis alaskensis	3702291	223	26138	65.08	84.08	4.69	11.82	0.06
Bin 15	Bradyrhizobium japonicum	331795	36	23855	56.53	0.00	0.00	11.31	0.18
Bin 16	Unknown	252925	21	23842	61.73	0.18	0.00	11.78	0.05
Bin 17	Unknown	796582	57	27835	65.54	9.29	0.00	10.75	0.12
Bin 18	Nitrospira defluvii	4436654	273	25626	58.89	94.30	6.17	10.97	0.05
Bin 19	Unknown (Burkholderiales, 目)	5277716	764	8212	71.10	88.75	8.62	7.70	0.03
Bin 20	Unknown	310141	52	6686	39.67	0.00	0.00	8.16	0.00
Bin 21	Hyphomicrobium zavarzinii	5274288	992	5860	63.47	93.16	38.61	7.55	0.69
Bin 22	Unknown	256566	28	11262	65.11	0.00	0.00	7.18	0.02
Bin 23	Unknown (未定)	2260362	500	4777	62.95	57.22	2.64	6.36	0.58
Bin 24	Hyphomicrobium zavarzinii	840024	188	4645	65.40	11.19	0.00	6.15	0.32
Bin 25	Nitrosomonas sp.	411506	120	3389	50.53	5.67	2.21	4.89	0.04
Bin 26	Sphingopyxis alaskensis	1925899	492	3812	65.50	36.82	1.66	5.24	0.07
Bin 27	Unknown(Planctomycetaceae, 科)	1430240	414	3321	59.18	30.78	3.09	4.55	0.01
Bin 28	Unknown (Rhizobiales, 目)	808658	238	3388	59.89	12.41	0.63	4.35	0.00
Bin 29	Unknown	370416	122	2959	68.01	2.74	0.00	3.61	0.00
Bin 30	Unknown (未定)	1331574	429	2966	65.33	17.85	6.75	3.78	0.01

2. 检验

2.1 对拼接结果的检验:

以Comammox为例:

变异程度:

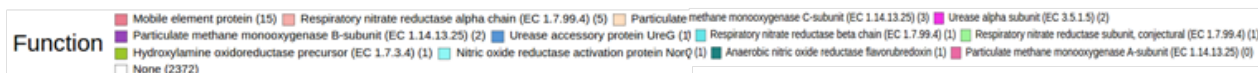
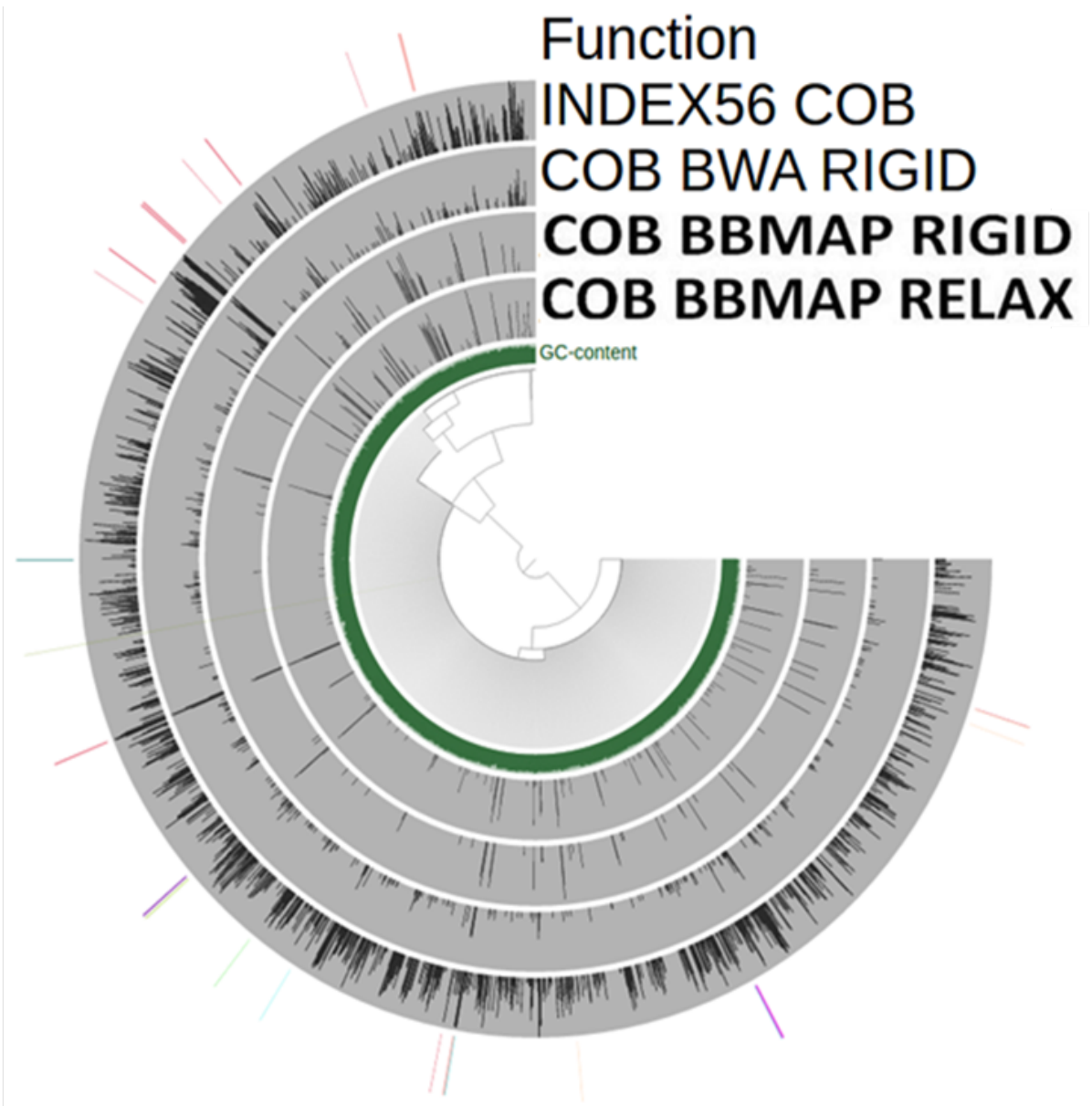
从里到外:

BBmap, relax & rigid;

Bwa, rigid & relax;

特定基因位置:

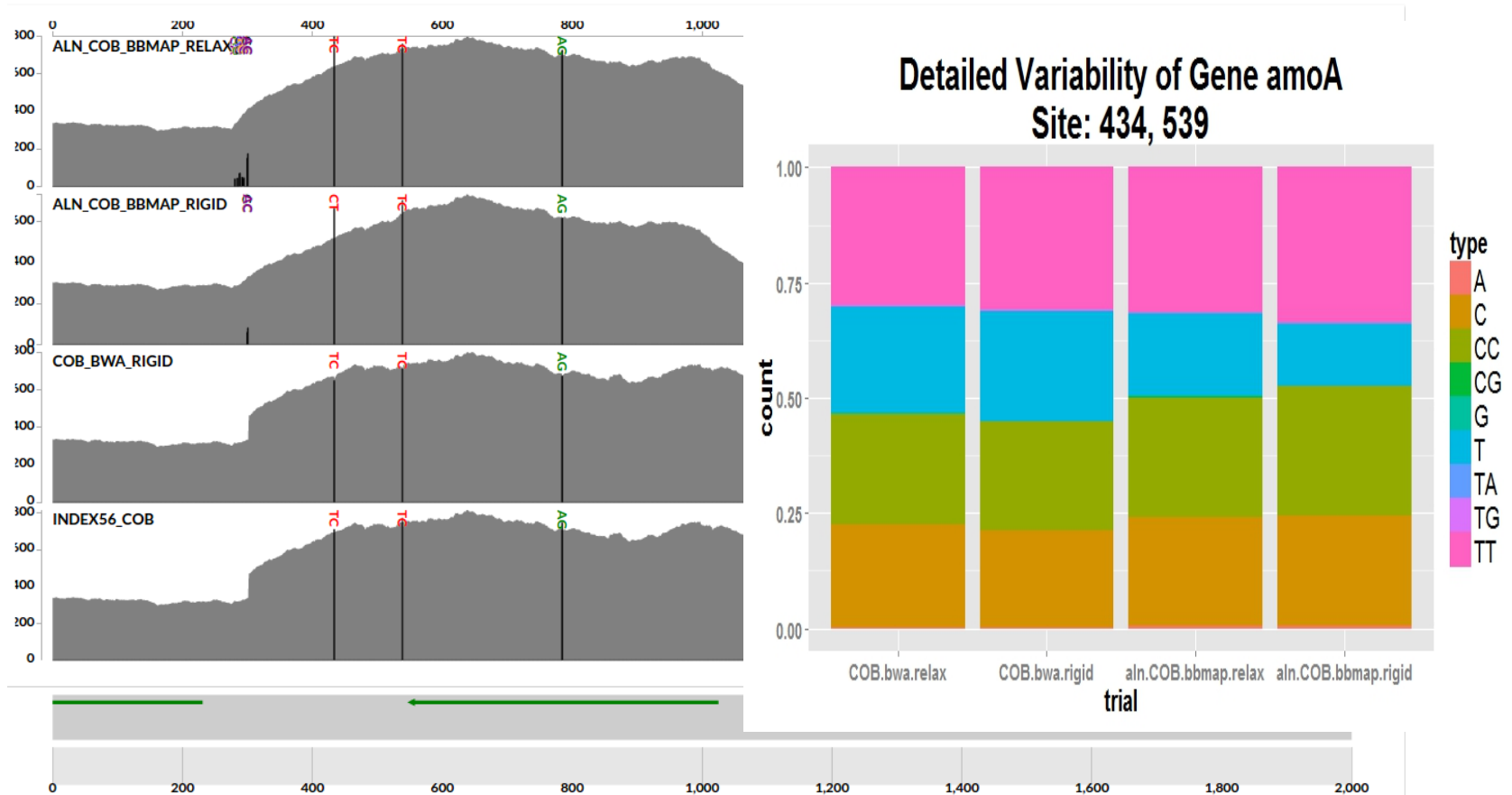
最外圈条块



2.1 对拼接结果的检验： *amoA*基因

scaffold1_size1986790_split_00346 detailed

[<<< prev](#) | [1758 of 2405](#) | [next >>>](#)



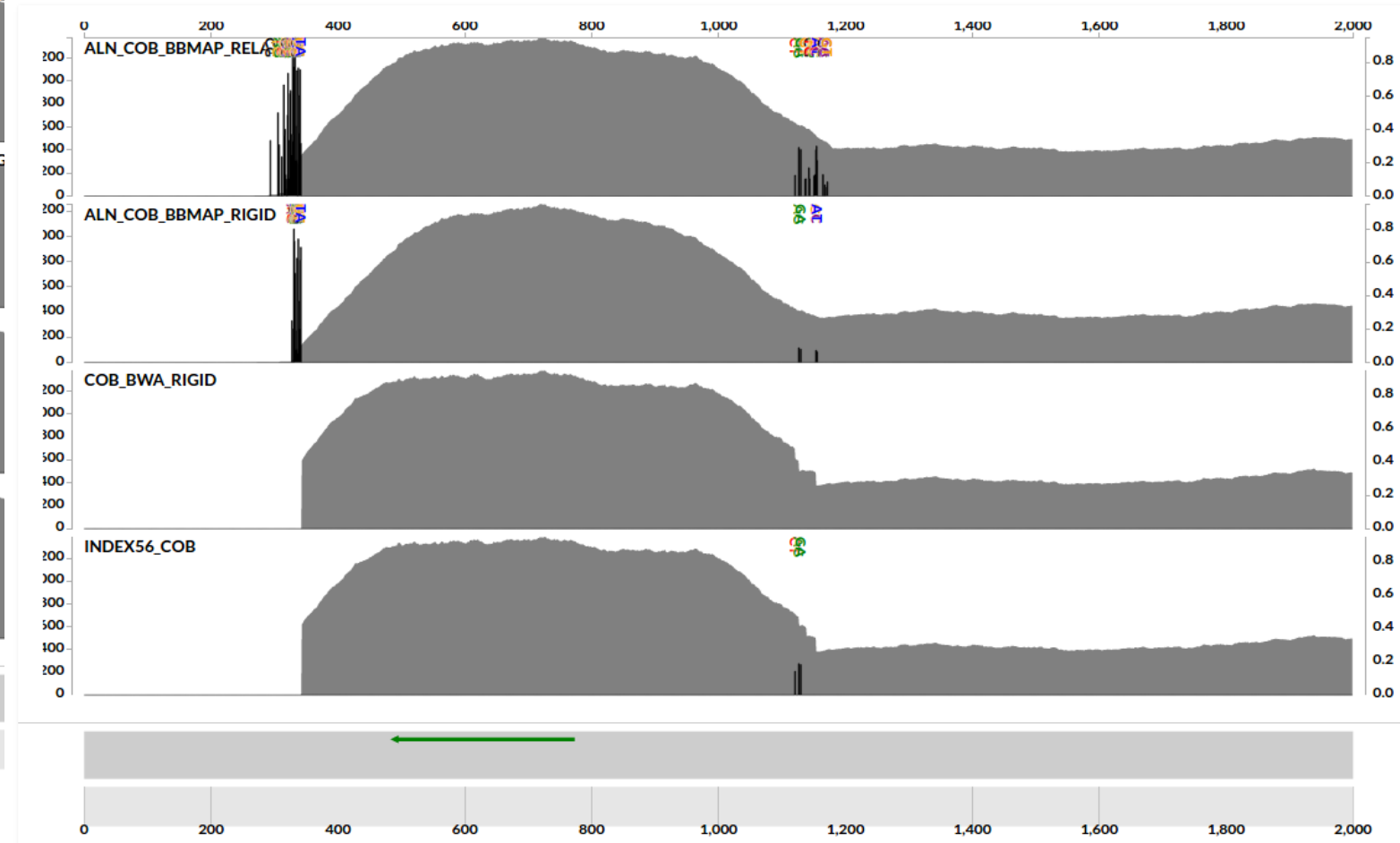
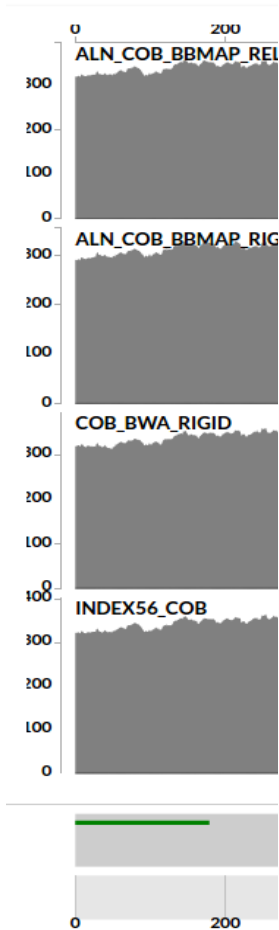
2.1 对拼接结果的检验: *amoB* & *amoC* 基因

scaffold1_size1986790_split_00345 detailed

[<<< prev](#) | [1757 of 2405](#) | [next >>>](#)

scaffold3_size283081_split_00004 detailed

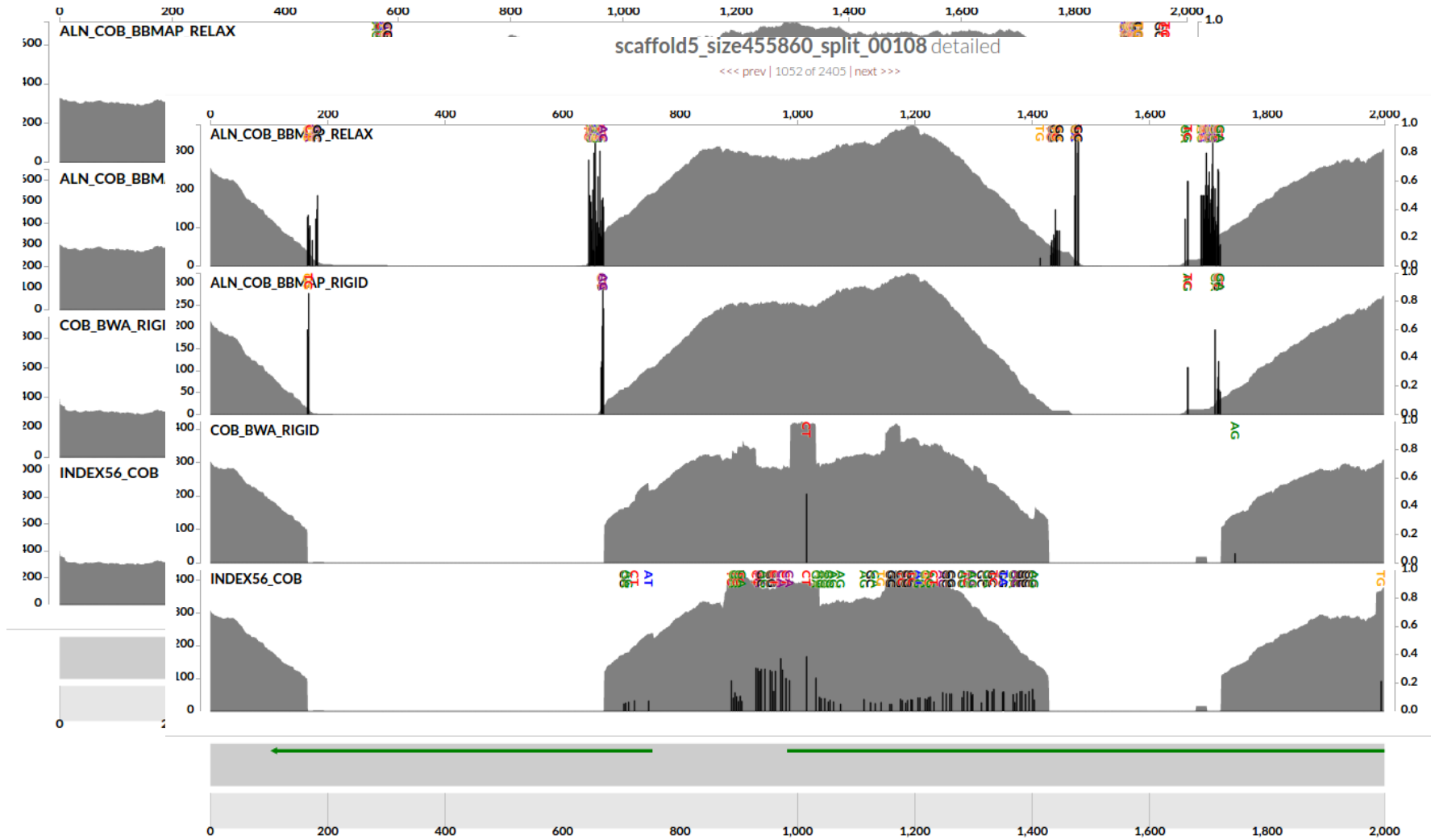
[<<< prev](#) | [791 of 2405](#) | [next >>>](#)



nxrA & *nxrB* 基因

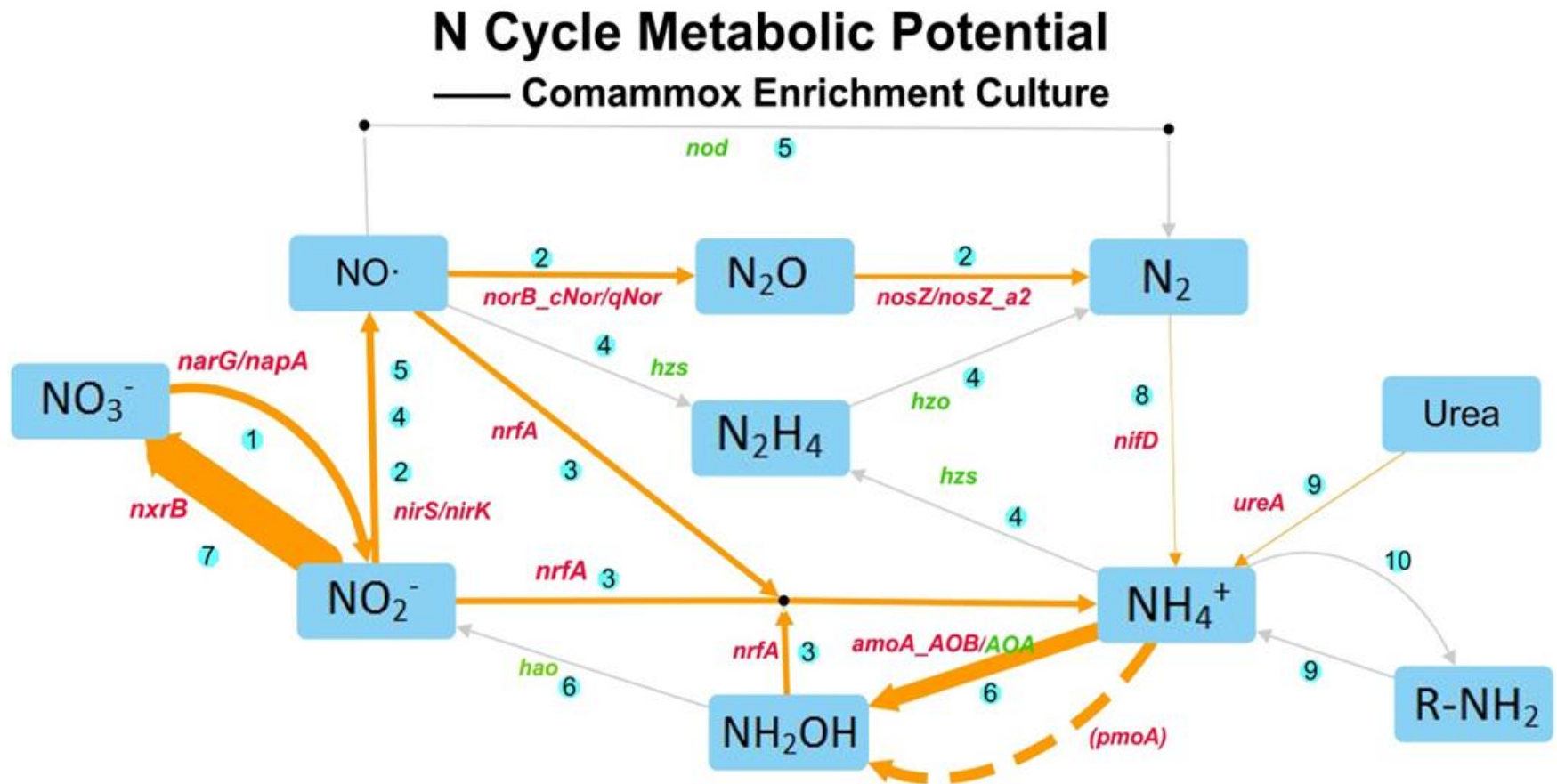
scaffold3_size283081_split_00140 detailed

[<<< prev](#) | [927 of 2405](#) | [next >>>](#)



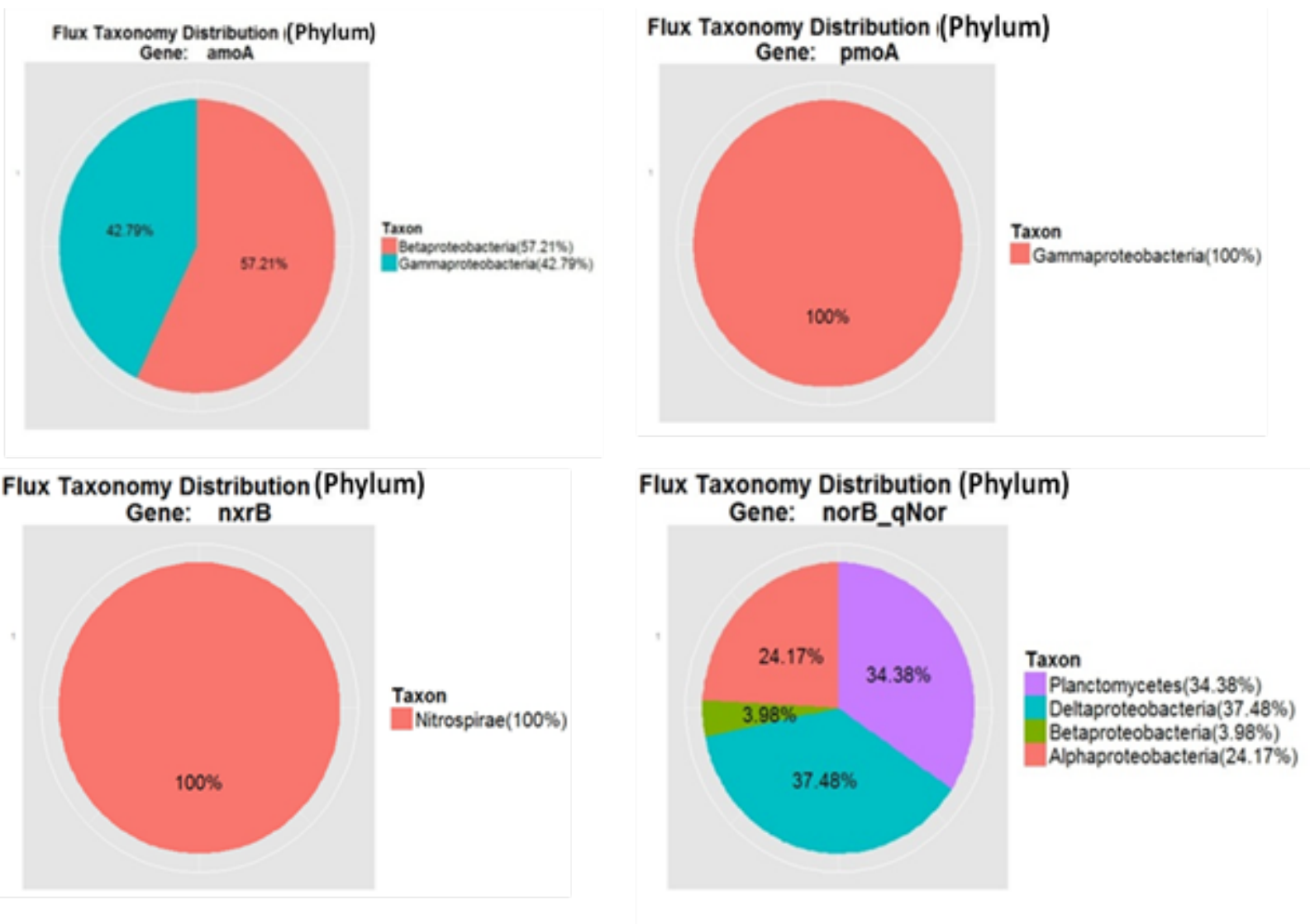
3.功能活性

3.1 测序序列层面 (Reads)



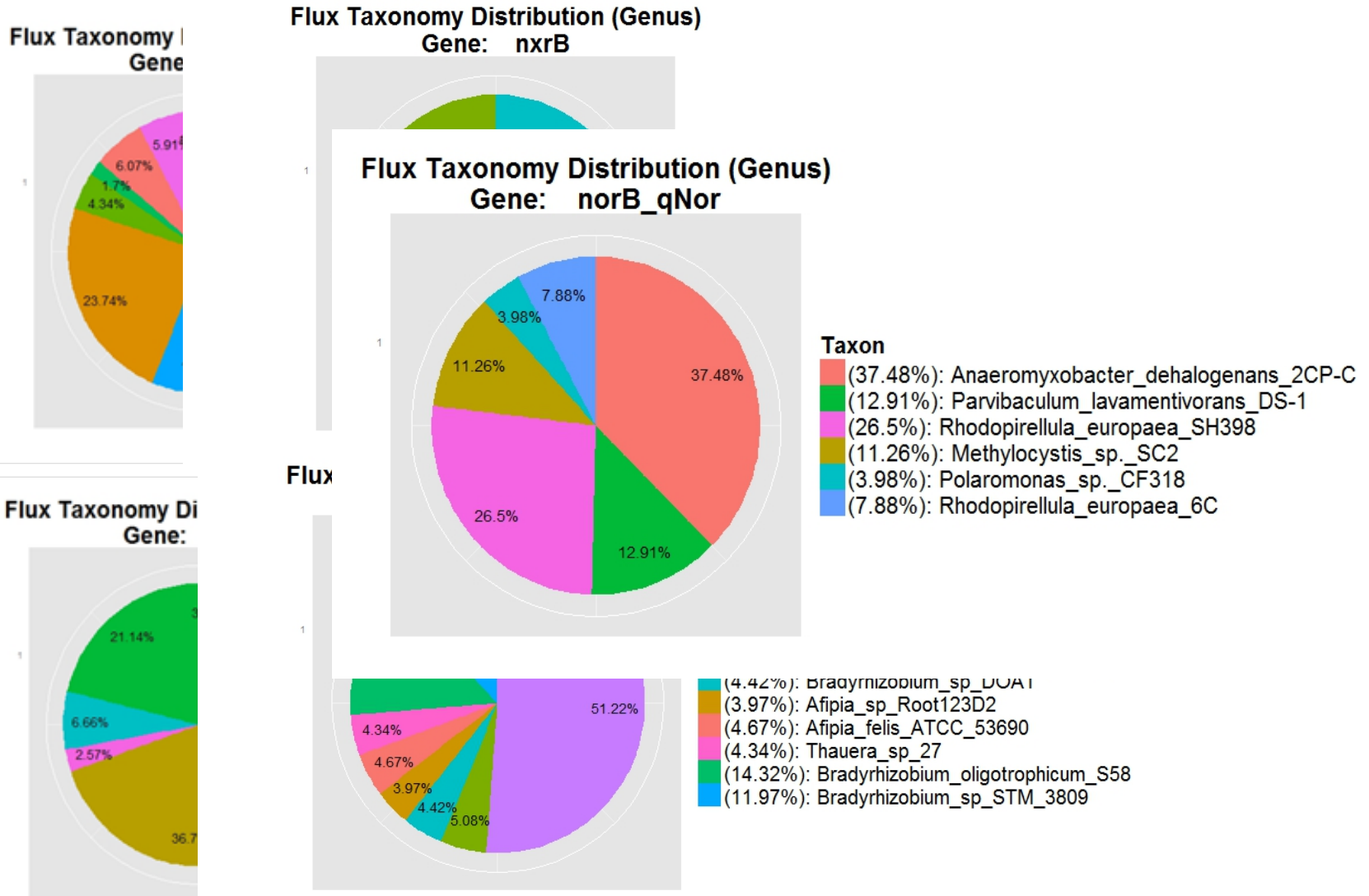
3.1 测序序列层面 (Reads)

氮循环通路标记基因在门水平分布图



3.1 测序序列层面 (Reads)

氮循环通路标记基因在种属水平分布图



3.1 测序序列层面（Reads）

Comammx富集培养物在聚类水平的功能分布

