利用高通量测序检测复杂群落中稀有生物区系的生物多样性

战爱斌

中国科学院生态环境研究中心

2015年1月20日

水生生物多样性丧失

栖息地丧失

生物入侵

污染

过度渔猎

>4600淡水动物濒危或灭绝，占所有淡水动物的25%

Thomsen et al. (2012) Molecular Ecology

我国水生态系统的生物多样性丧失

水生生态系统污染问题严峻

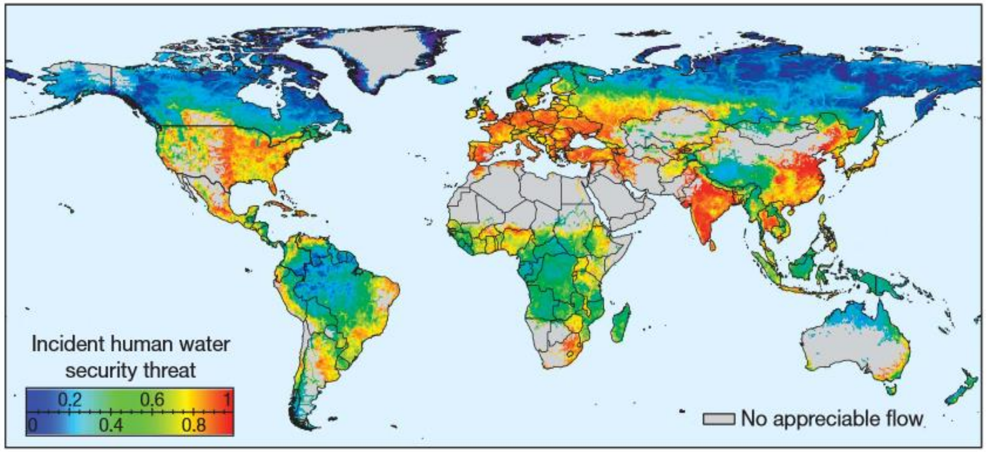
-化学品：中国大量使用>1万种

-工农业、生活污水

nature

Global threats to human water security and river biodiversity

C. J. Vörösmarty[1]*, P. B. McIntyre[2]*†, M. O. Gessner[3], D. Dudgeon[4], A. Prusevich[5], P. Green[1], S. Glidden[5], S. E. Bunn[6], C. A. Sullivan[7], C. Reidy Liermann[8] & P. M. Davies[9]



生物多样性空间分异

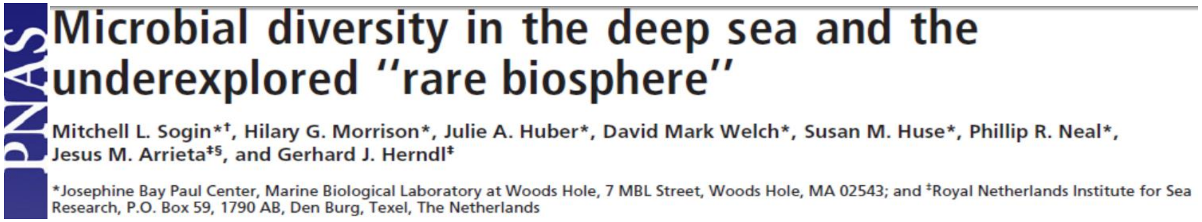
1、驱动因子是什么？

1、过程是什么？

2、机理是什么？

稀有生物区系

稀有生物区系（Rare biosphere）


PNAS

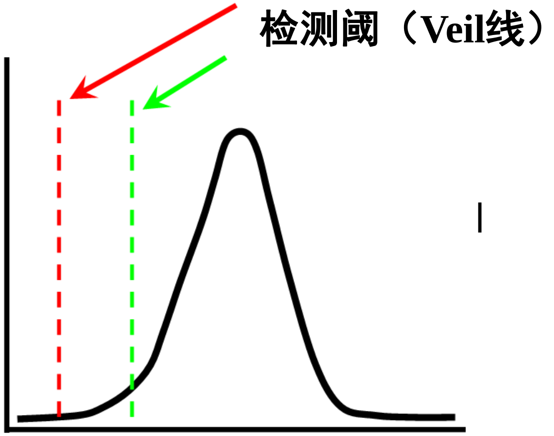
Microbial diversity in the deep sea and the underexplored “rare biosphere”

Mitchell L. Sogin*†, Hilary G. Morrison*, Julie A. Huber*, David Mark Welch*, Susan M. Huse*, Phillip R. Neal*, Jesus M. Arrieta‡§, and Gerhard J. Herndl‡

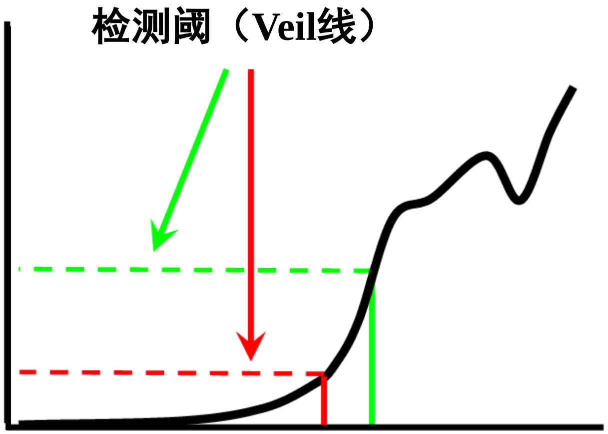
*Josephine Bay Paul Center, Marine Biological Laboratory at Woods Hole, 7 MBL Street, Woods Hole, MA 02543; and ‡Royal Netherlands Institute for Sea Research, P.O. Box 59, 1790 AB, Den Burg, Texel, The Netherlands

群落中所有低丰度生物的集合

污染生态学研究的关键

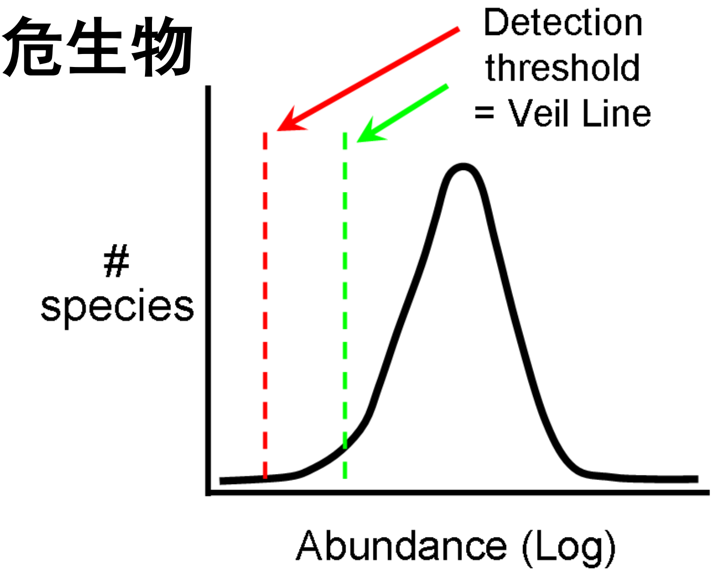


生物入侵早期检测的关键

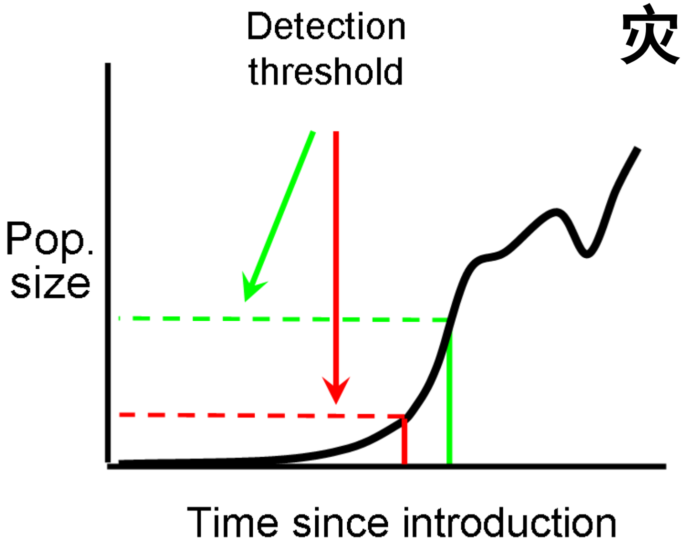


方法学挑战

稀有/濒危生物



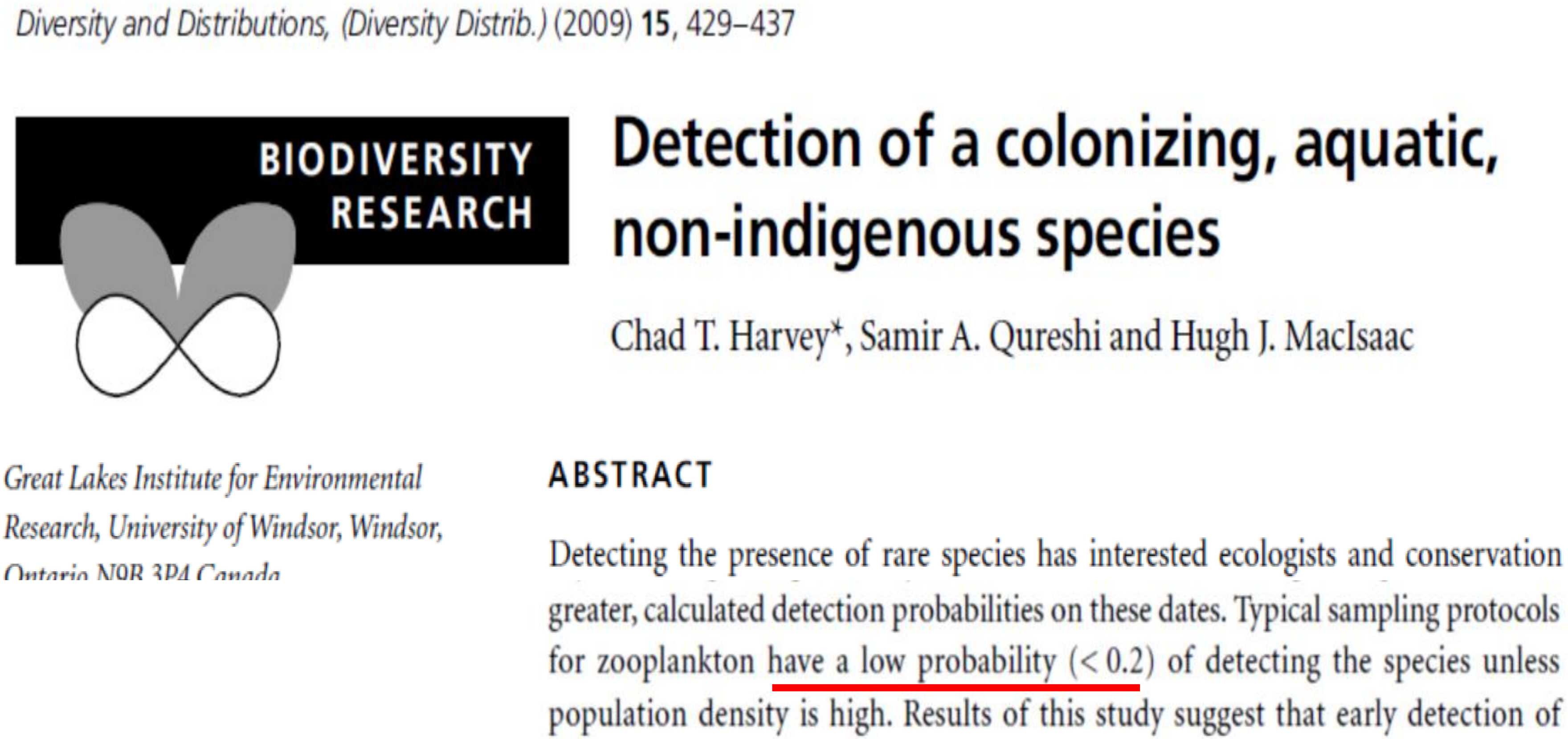
灾害生物



世界性难题

- 一匿藏于水下
- 一水域广袤
- 一绝大数物种个体微小
- 一可利用的外部形态极少
- 一群体数量及密度极低

方法学挑战



Diversity and Distributions, (Diversity Distrib.) (2009) **15**, 429–437

BIODIVERSITY RESEARCH

Detection of a colonizing, aquatic, non-indigenous species

Chad T. Harvey*, Samir A. Qureshi and Hugh J. MacIsaac

Great Lakes Institute for Environmental Research, University of Windsor, Windsor, Ontario N9B 3P4 Canada

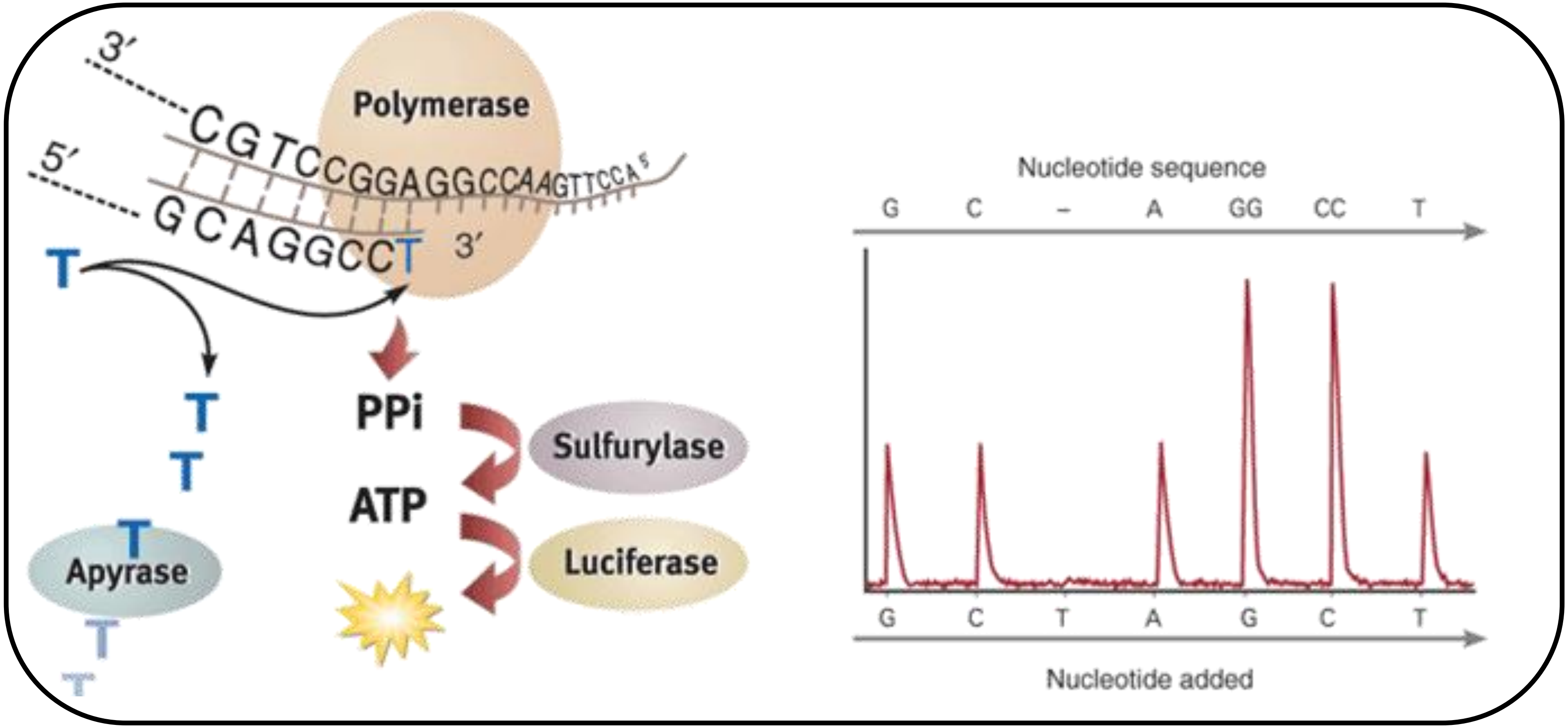
ABSTRACT

Detecting the presence of rare species has interested ecologists and conservation greater, calculated detection probabilities on these dates. Typical sampling protocols for zooplankton have a low probability (< 0.2) of detecting the species unless population density is high. Results of this study suggest that early detection of

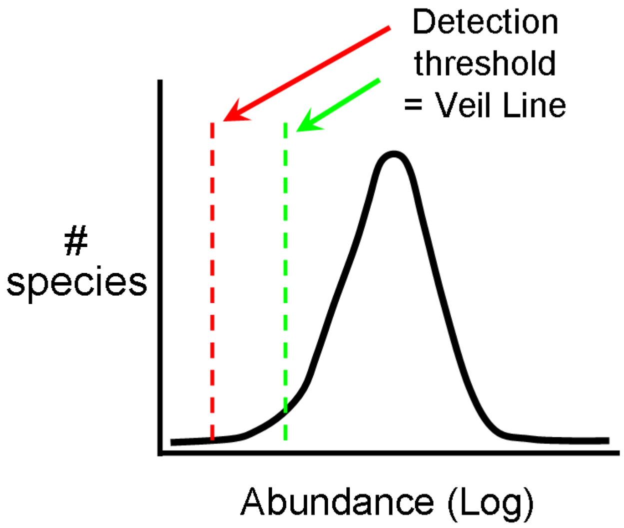
常规方法的检出率<0.2

解决方案

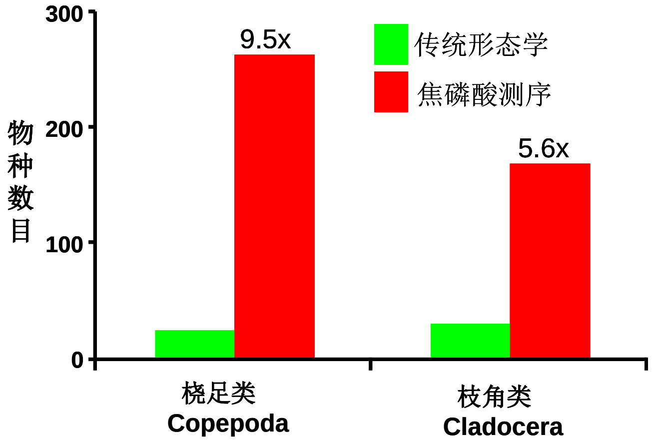
高通量测序

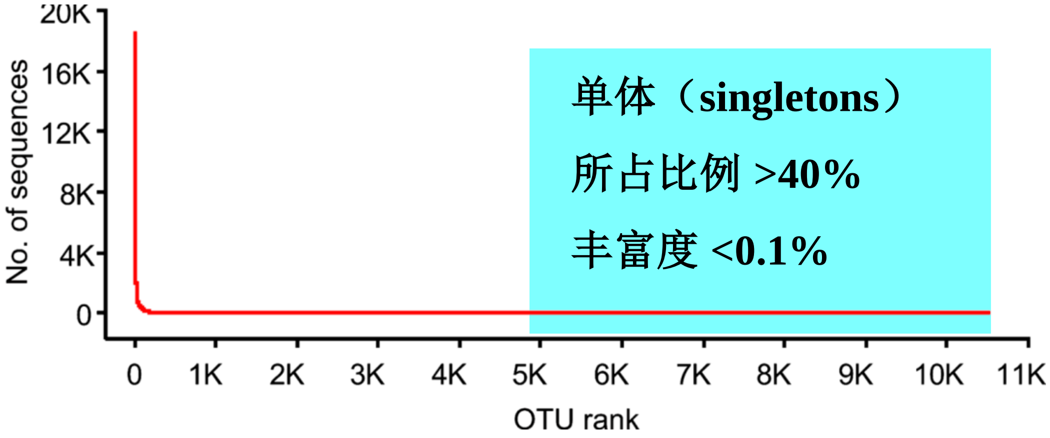

3′
5′
CGTCCGGAGGCCAAGTTCCA 5′
GCAGGCCT 3′
Polymerase
T
T
T
T
Apyrase
PPi
ATP
Sulfurylase
Luciferase
Nucleotide sequence
G C – A GG CC T
G C T A G C T
Nucleotide added

亟待回答的方法学问题



1. 稀有生物检测限问题
2. 稀有生物检出的重复率问题
3. 数据滤错对稀有生物检出的影响
4. 序列丰度与生物量之间的关系





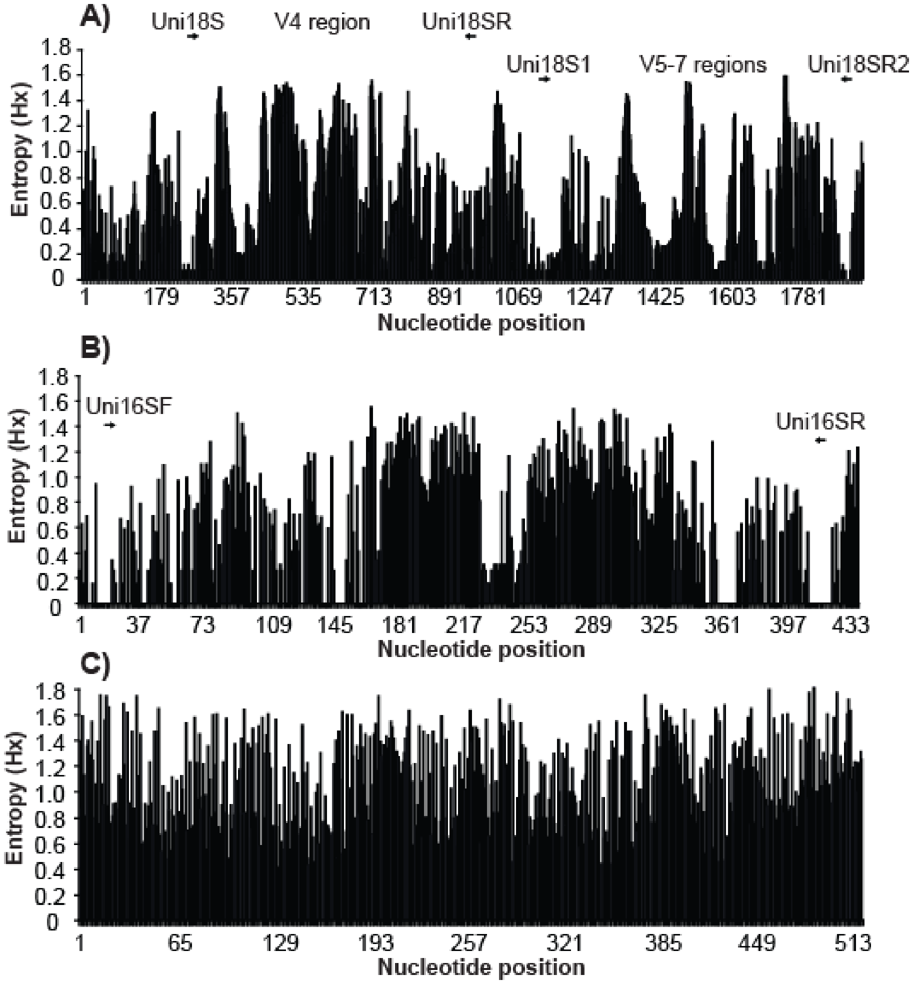
单体（singletons）

所占比例 >40%

丰富度 <0.1%

1. 检测限问题

分子标记选择



18S（Nuclear）

V4：Uni18S-Uni18SR

V5-7：Uni18S1-Uni18SR2

16S（mtDNA）

Uni16SF-Uni16SR

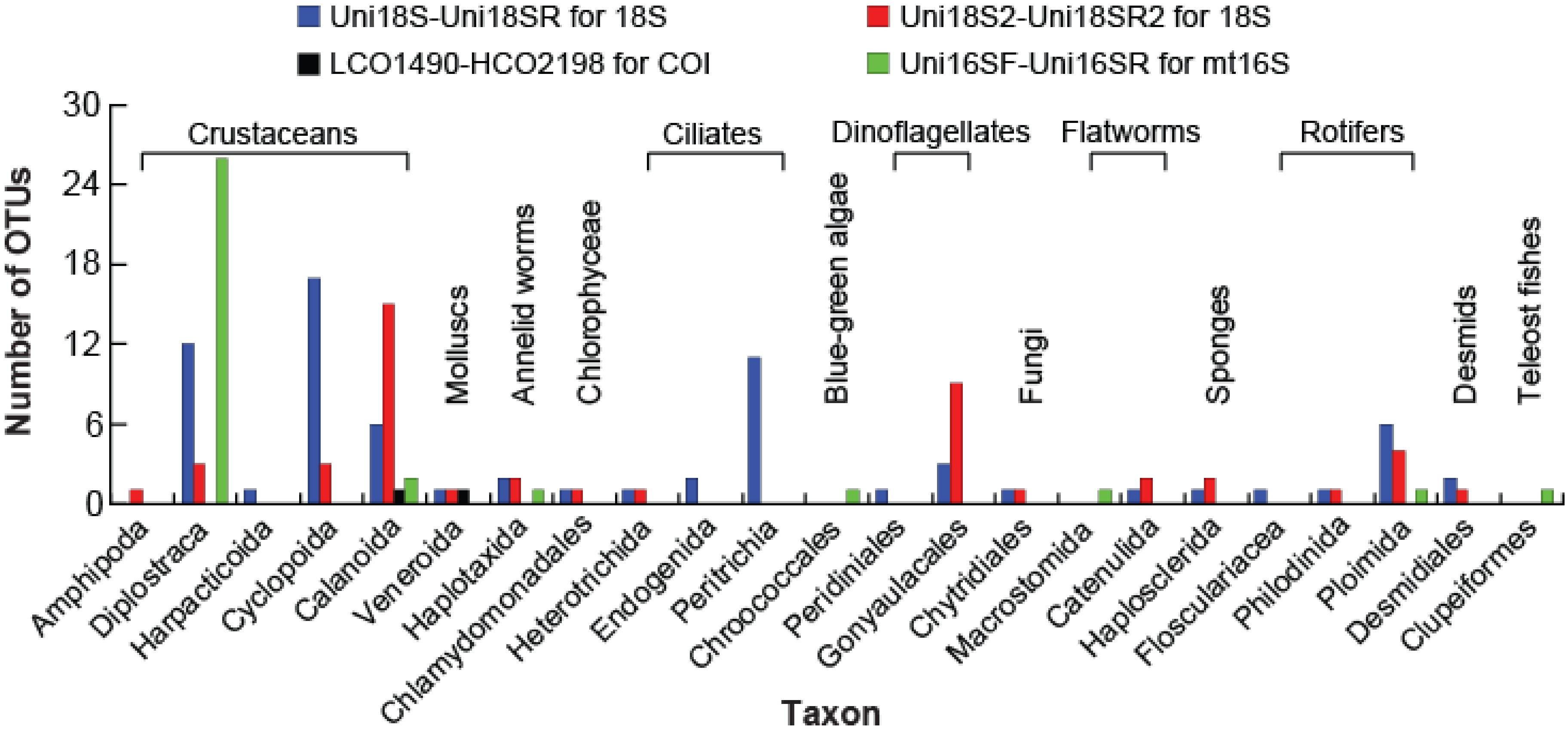
COI（mtDNA）

No new versatile primer detected, using HCO-LCO

Zhan et al. (2014) Molecular Ecology Resources

1. 检测限问题

分子标记选择

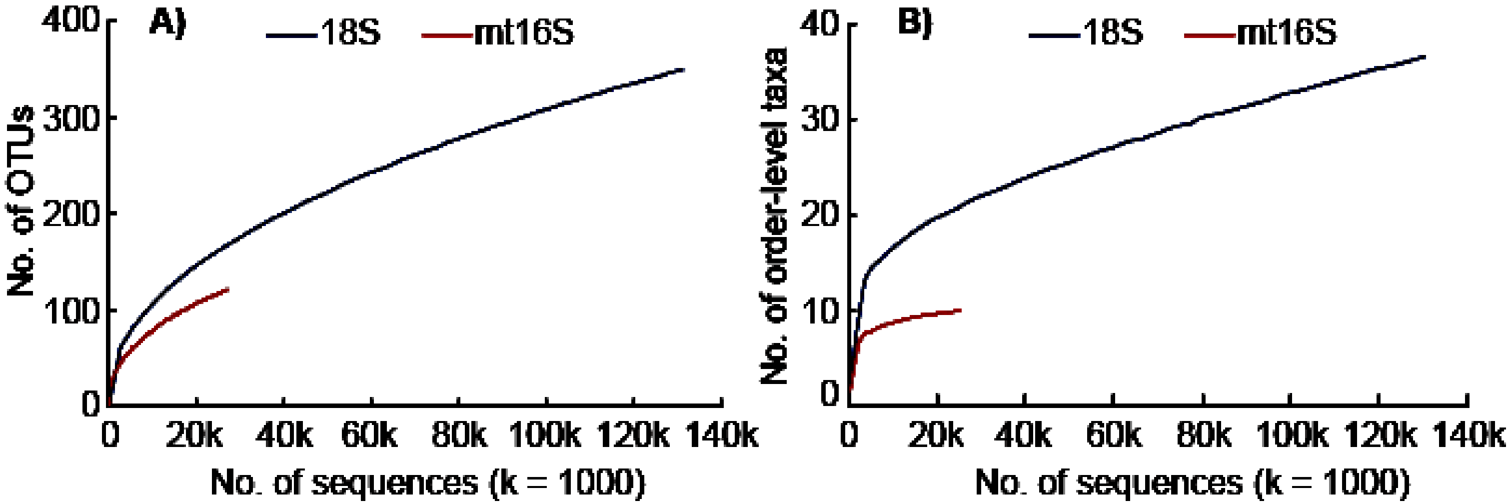


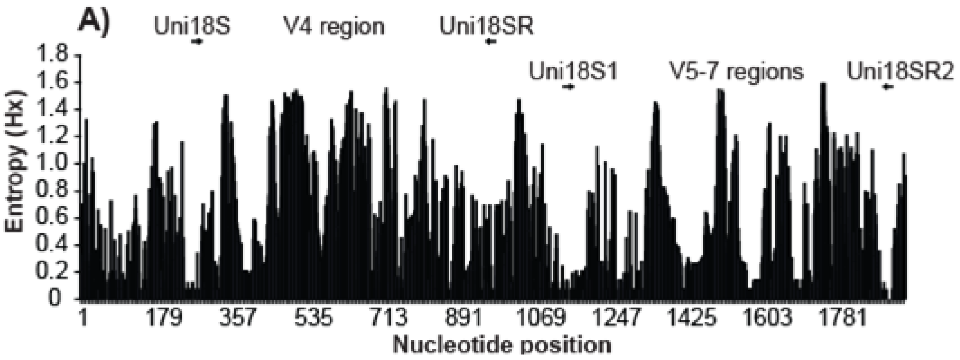
Sequencing depth = 1/2 PicoTiter plate

Zhan et al. (2014) Molecular Ecology Resources

1. 检测限问题

分子标记选择





18S-V4

Uni18S

AGGGCAAKYCTGGTGCCAGC

Uni18SR

GRCGGTATCTRATCGYCTT

Zhan et al. (2014) Molecular Ecology Resources

1. 检测限问题

群落构建

海洋生物

海湾扇贝

©All rights reserved
puffinbytes@yahoo.com.au

仿刺参

↓ 幼虫（实验室）

淡水浮游群落

淡水生物

沼蛤

水虱

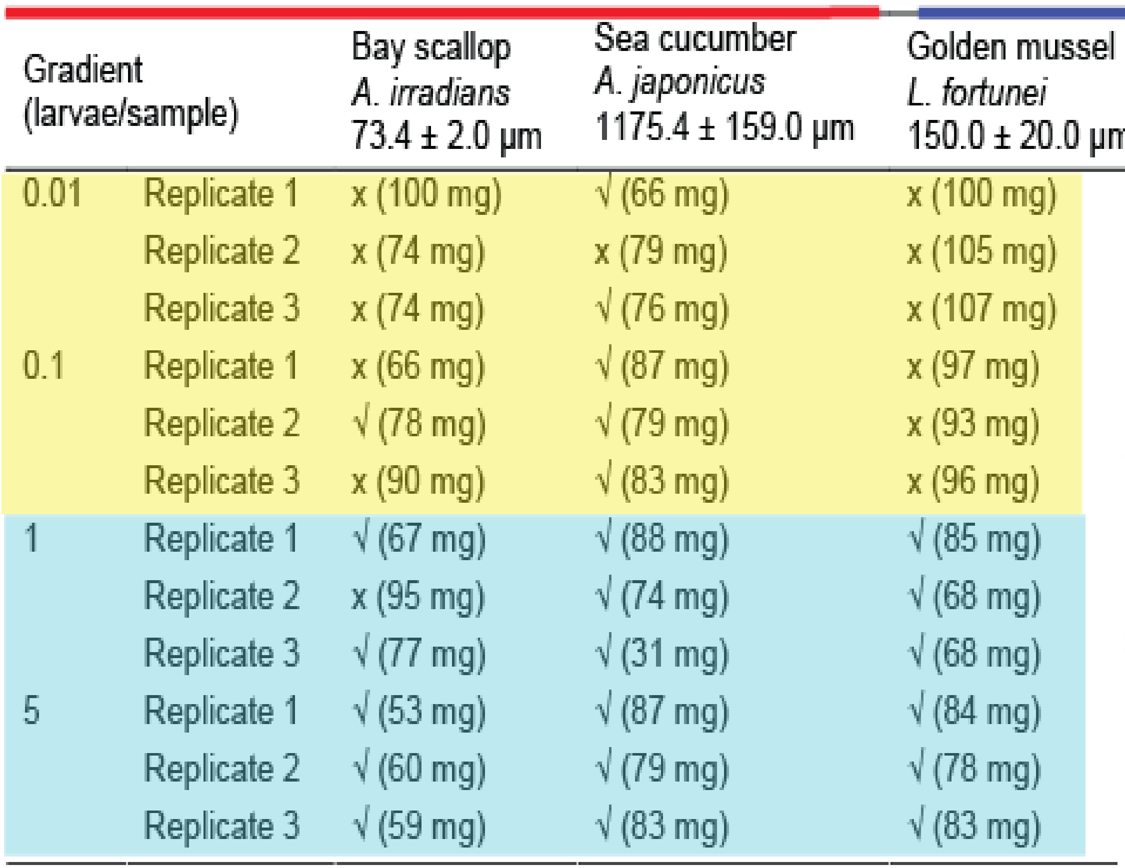
↓ 幼虫（野外）

海水浮游群落

1. 检测限问题

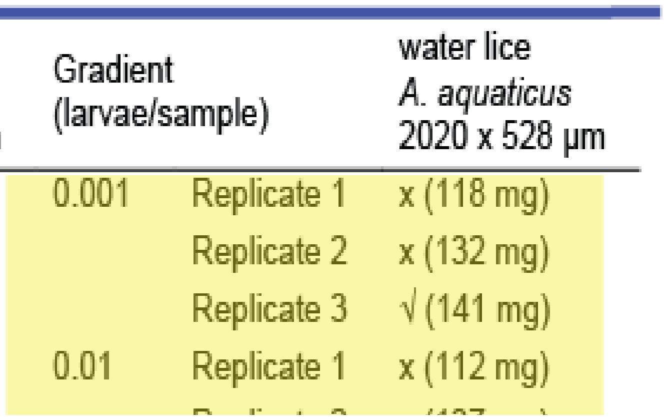
检测结果

海洋生物

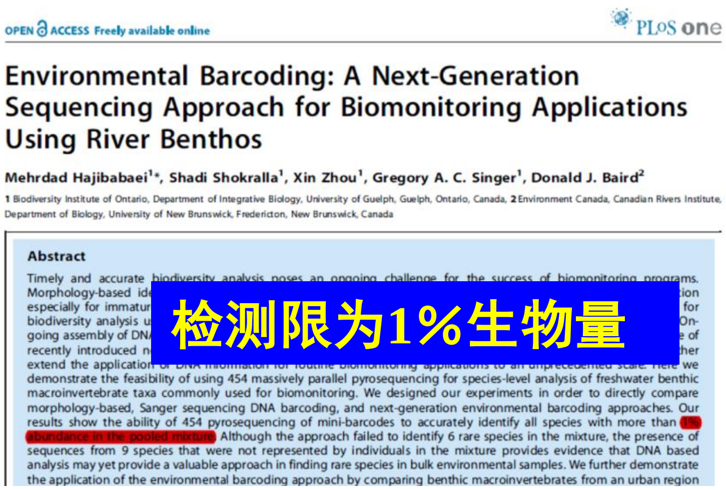


Gradient (larvae/sample)		Bay scallop *A. irradians* 73.4 ± 2.0 μm	Sea cucumber *A. japonicus* 1175.4 ± 159.0 μm	Golden mussel *L. fortunei* 150.0 ± 20.0 μm
0.01	Replicate 1	x (100 mg)	√ (66 mg)	x (100 mg)
	Replicate 2	x (74 mg)	x (79 mg)	x (105 mg)
	Replicate 3	x (74 mg)	√ (76 mg)	x (107 mg)
0.1	Replicate 1	x (66 mg)	√ (87 mg)	x (97 mg)
	Replicate 2	√ (78 mg)	√ (79 mg)	x (93 mg)
	Replicate 3	x (90 mg)	√ (83 mg)	x (96 mg)
1	Replicate 1	√ (67 mg)	√ (88 mg)	√ (85 mg)
	Replicate 2	x (95 mg)	√ (74 mg)	√ (68 mg)
	Replicate 3	√ (77 mg)	√ (31 mg)	√ (68 mg)
5	Replicate 1	√ (53 mg)	√ (87 mg)	√ (84 mg)
	Replicate 2	√ (60 mg)	√ (79 mg)	√ (78 mg)
	Replicate 3	√ (59 mg)	√ (83 mg)	√ (83 mg)

淡水生物



Gradient (larvae/sample)		water lice *A. aquaticus* 2020 x 528 μm
0.001	Replicate 1	x (118 mg)
	Replicate 2	x (132 mg)
	Replicate 3	√ (141 mg)
0.01	Replicate 1	x (112 mg)



OPEN ACCESS Freely available online

PLoS one

Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos

Mehrdad Hajibabaei[1]*, Shadi Shokralla[1], Xin Zhou[1], Gregory A. C. Singer[1], Donald J. Baird[2]

[1] Biodiversity Institute of Ontario, Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada, [2] Environment Canada, Canadian Rivers Institute, Department of Biology, University of New Brunswick, Fredericton, New Brunswick, Canada

Abstract

Timely and accurate biodiversity analysis poses an ongoing challenge for the success of biomonitoring programs.
Morphology-based ide … tion
especially for immatur … for
biodiversity analysis u … On-
going assembly of DN … e of
recently introduced n … ther
extend the application … here we
demonstrate the feasibility of using 454 massively parallel pyrosequencing for species-level analysis of freshwater benthic macroinvertebrate taxa commonly used for biomonitoring. We designed our experiments in order to directly compare morphology-based, Sanger sequencing DNA barcoding, and next-generation environmental barcoding approaches. Our results show the ability of 454 pyrosequencing of mini-barcodes to accurately identify all species with more than 1% abundance in the pooled mixture. Although the approach failed to identify 6 rare species in the mixture, the presence of sequences from 9 species that were not represented by individuals in the mixture provides evidence that DNA based analysis may yet provide a valuable approach in finding rare species in bulk environmental samples. We further demonstrate the application of the environmental barcoding approach by comparing benthic macroinvertebrates from an urban region

检测限为1%生物量

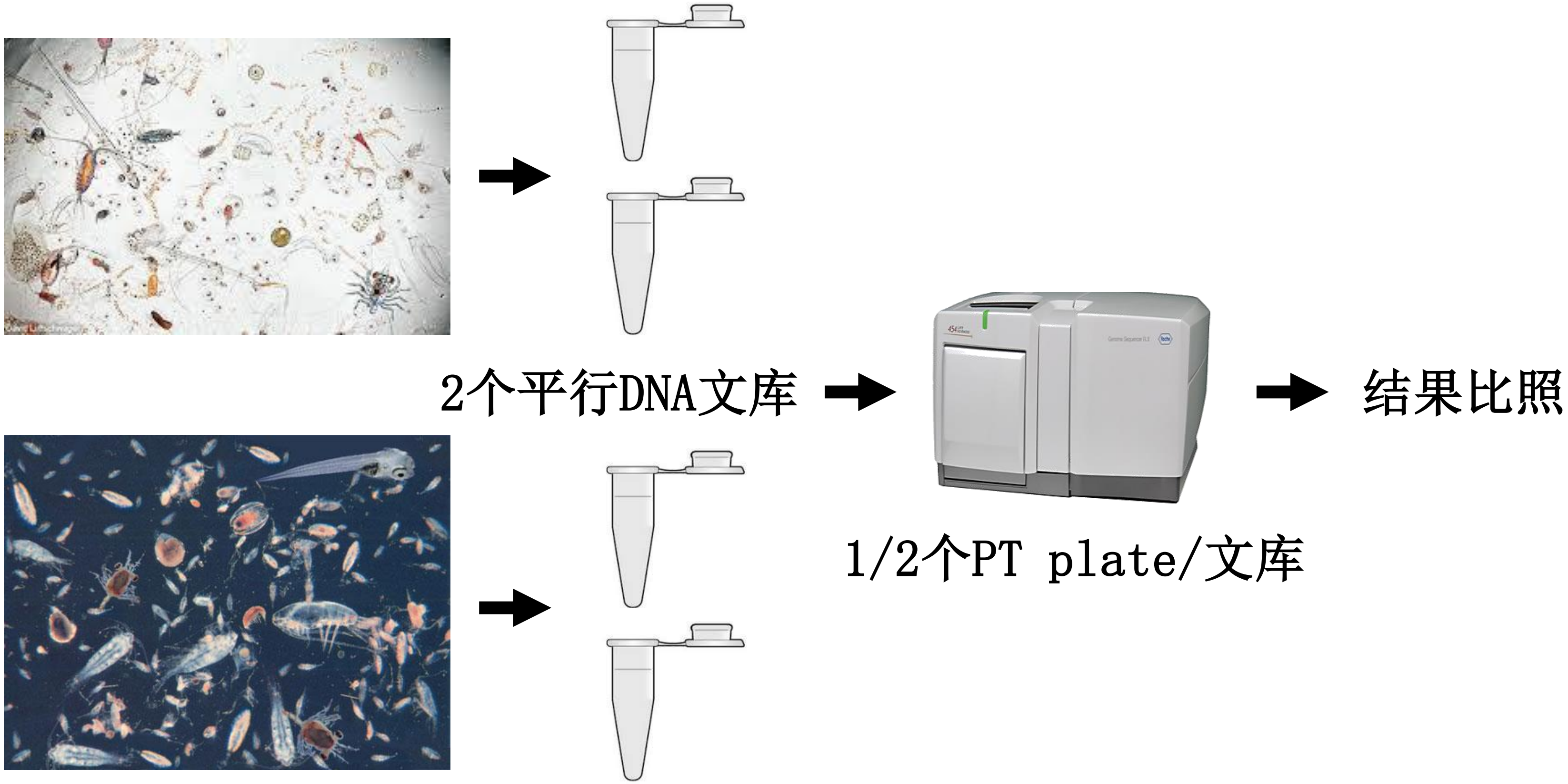
>1幼虫，即可检出
灵敏度~ 2.3 x 10^{-5}% (生物量)

5个数量级

← **1% (生物量)**

2. 检测重复率

实验设计



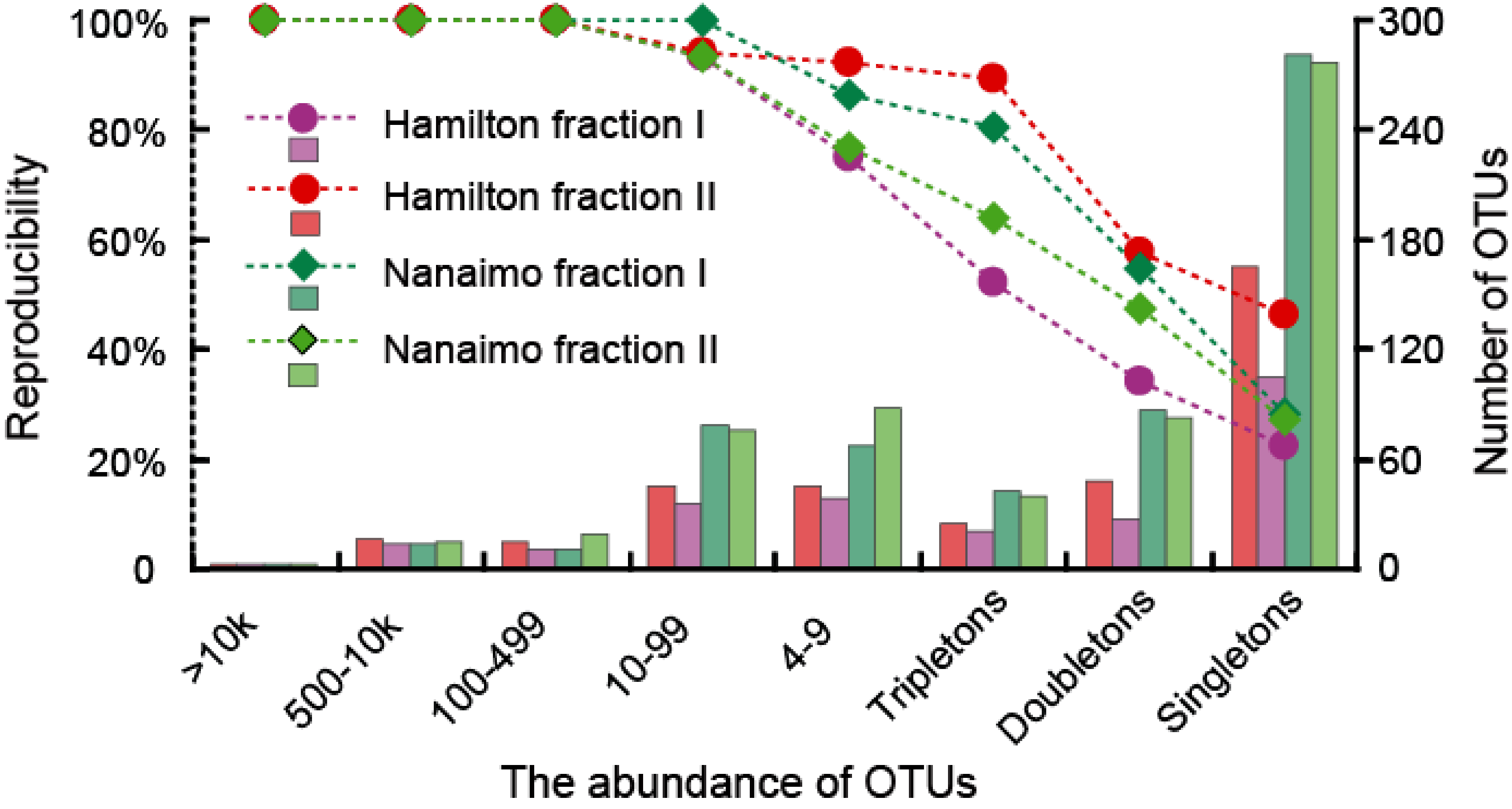
2个平行DNA文库

1/2个PT plate/文库

结果比照

2. 检测重复率

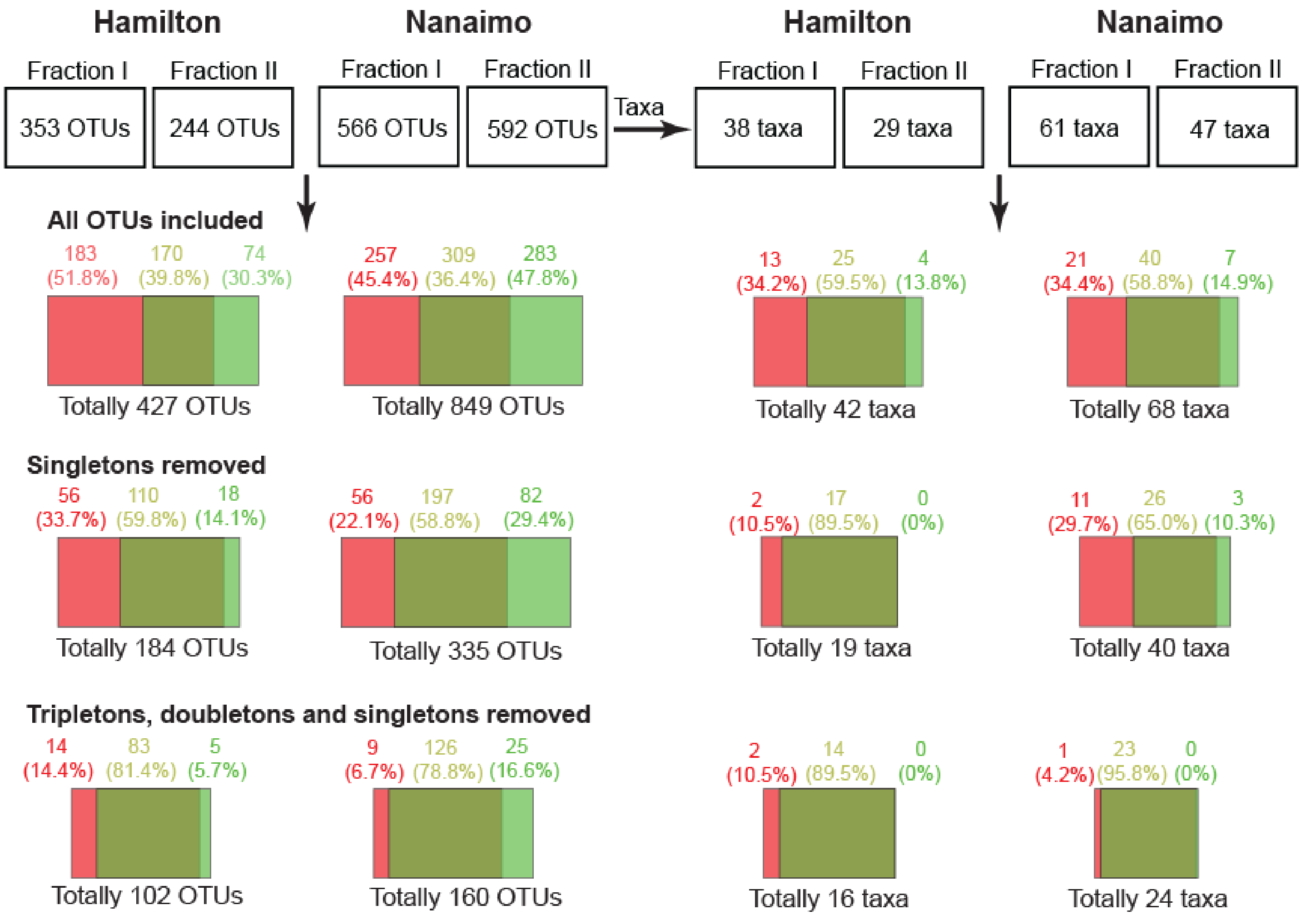
OTU重复率



Zhan et al. (2014) Methods in Ecology and Evolution

2. 检测重复率

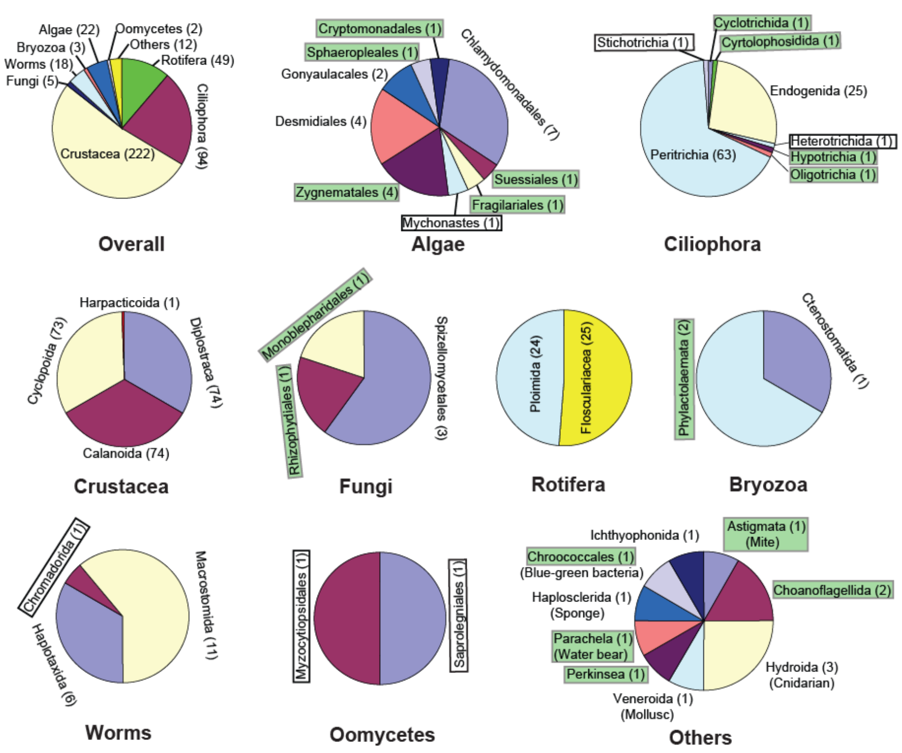
OTU重复率与其丰度关系



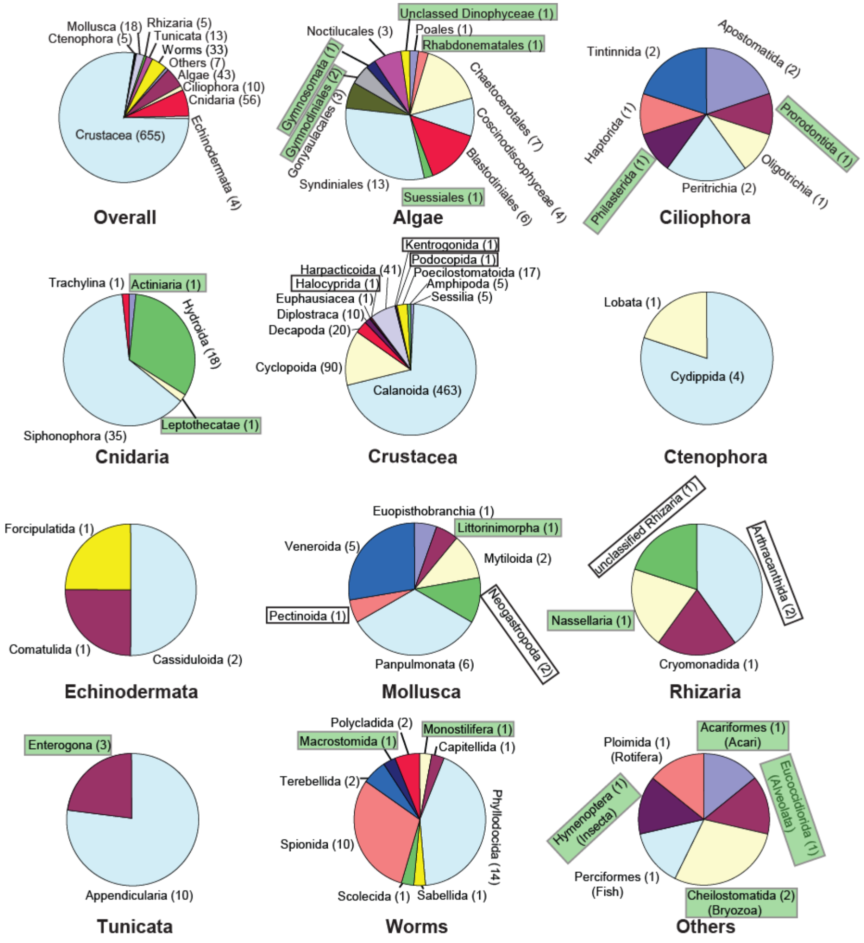
Zhan et al. (2014) Methods in Ecology and Evolution

2. 检测重复率

类群水平重复率



Hamilton

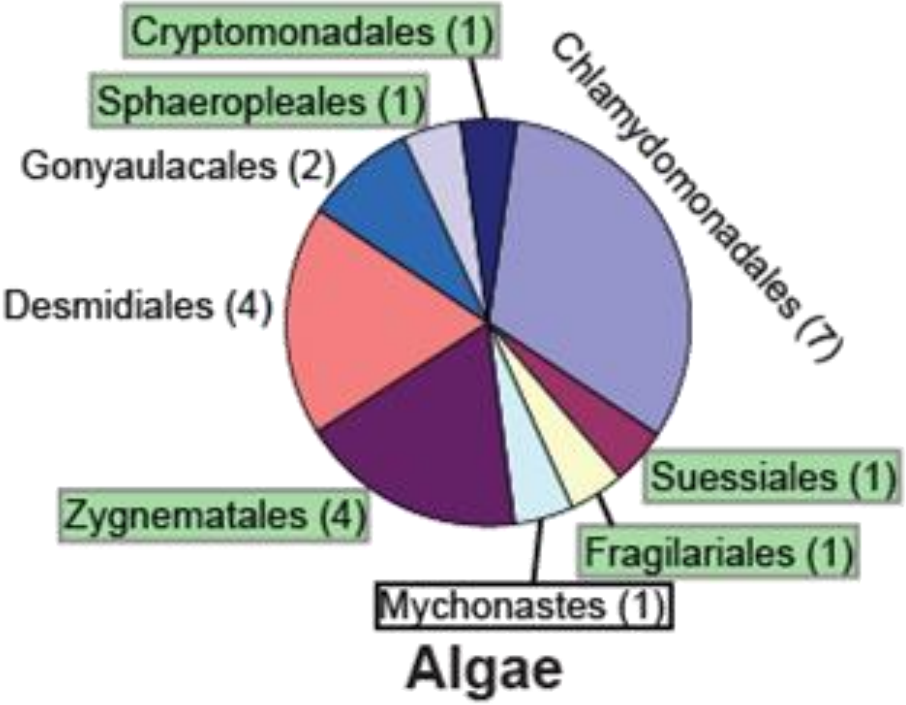


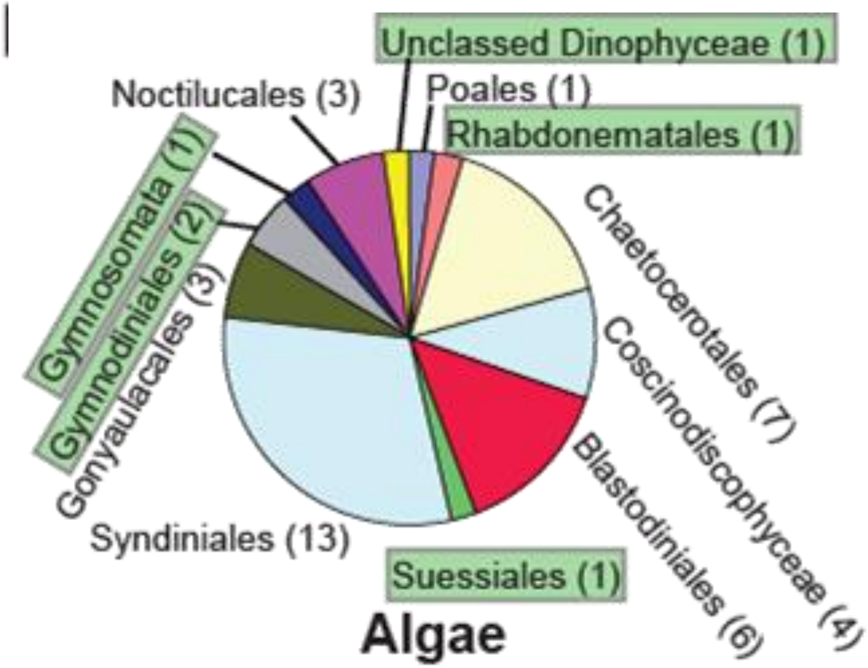
Nanaimo

Zhan et al. (2014) Methods in Ecology and Evolution

2. 检测重复率

类群水平重复率





1、单体不是由于测序误差等原因引起

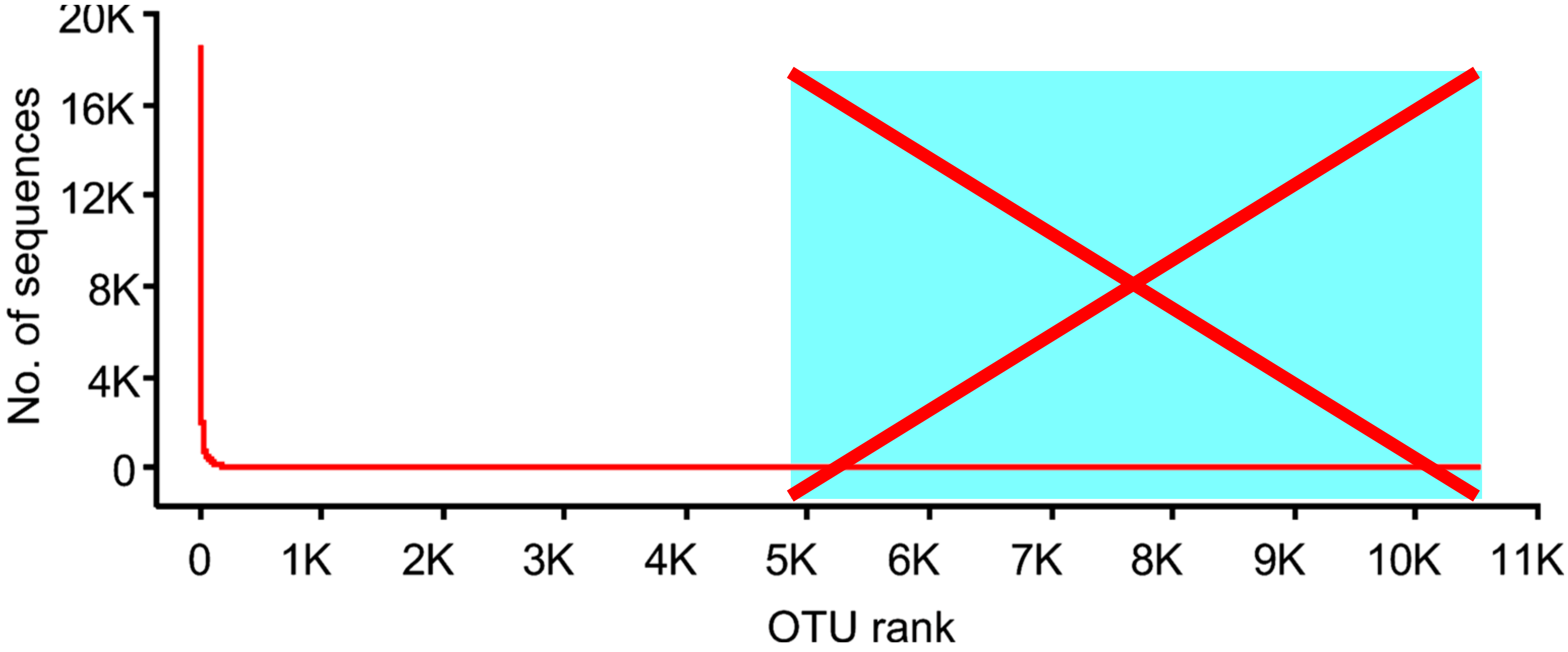
2、低重复率是由取样偏差引起

3、简单删除单体极大降低生物多样性，尤其低丰度生物

4、如此多的低丰度生物的生物及生态学效应值得深入研究

Zhan et al. (2014) Methods in Ecology and Evolution

3. 滤错的影响



1、直接删除（显然是错误的）

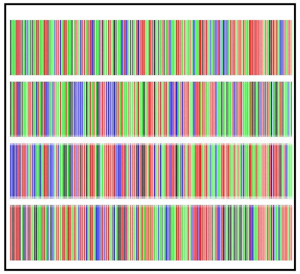
2、Phred Score (Q)滤错

-滤错对稀有物种是否有影响？

3. 滤错的影响

实验设计

方法1：自然群落



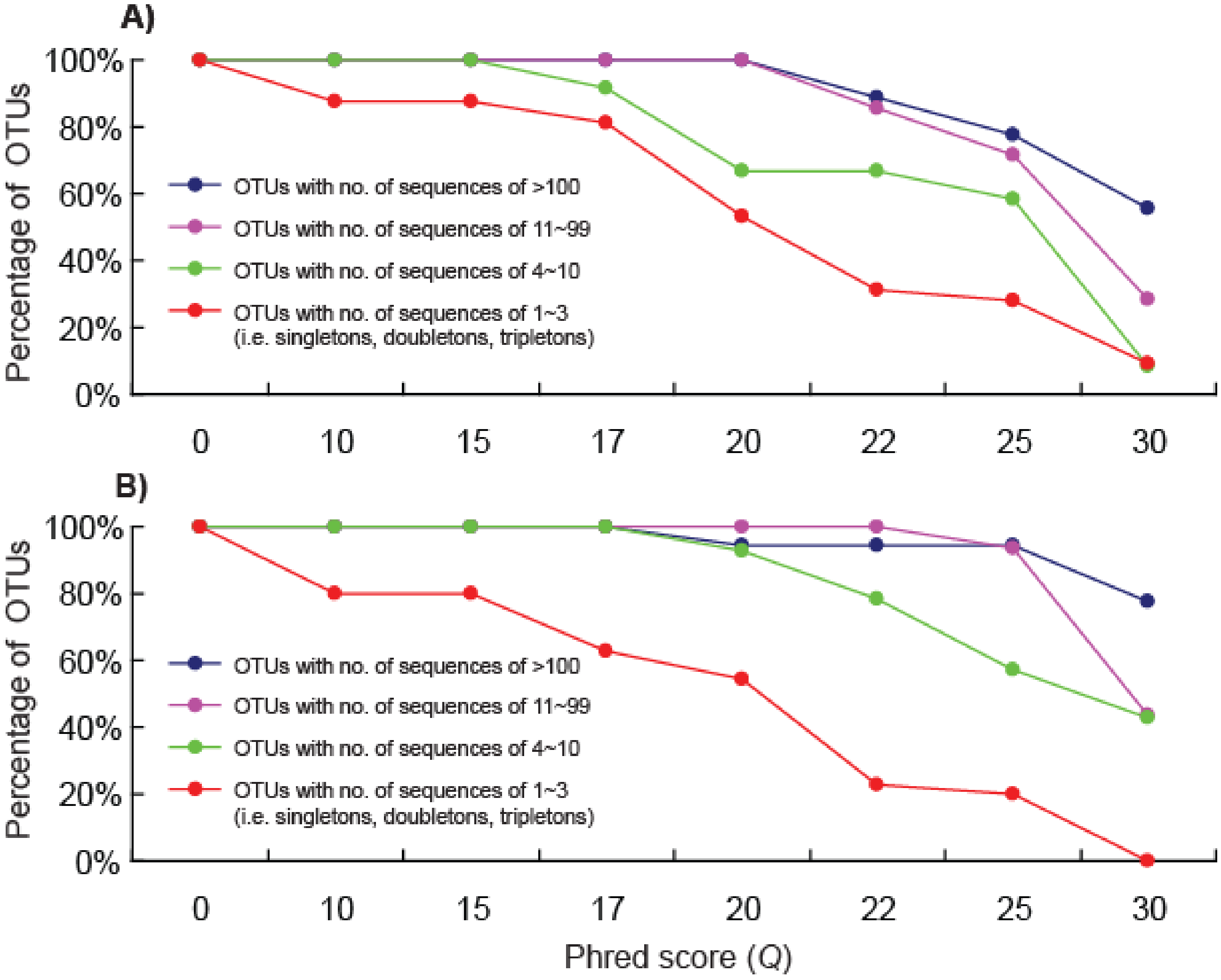
→ S>99%,
C>99% → Q filtering

方法2：人工群落

→ 48 communities → Q filtering

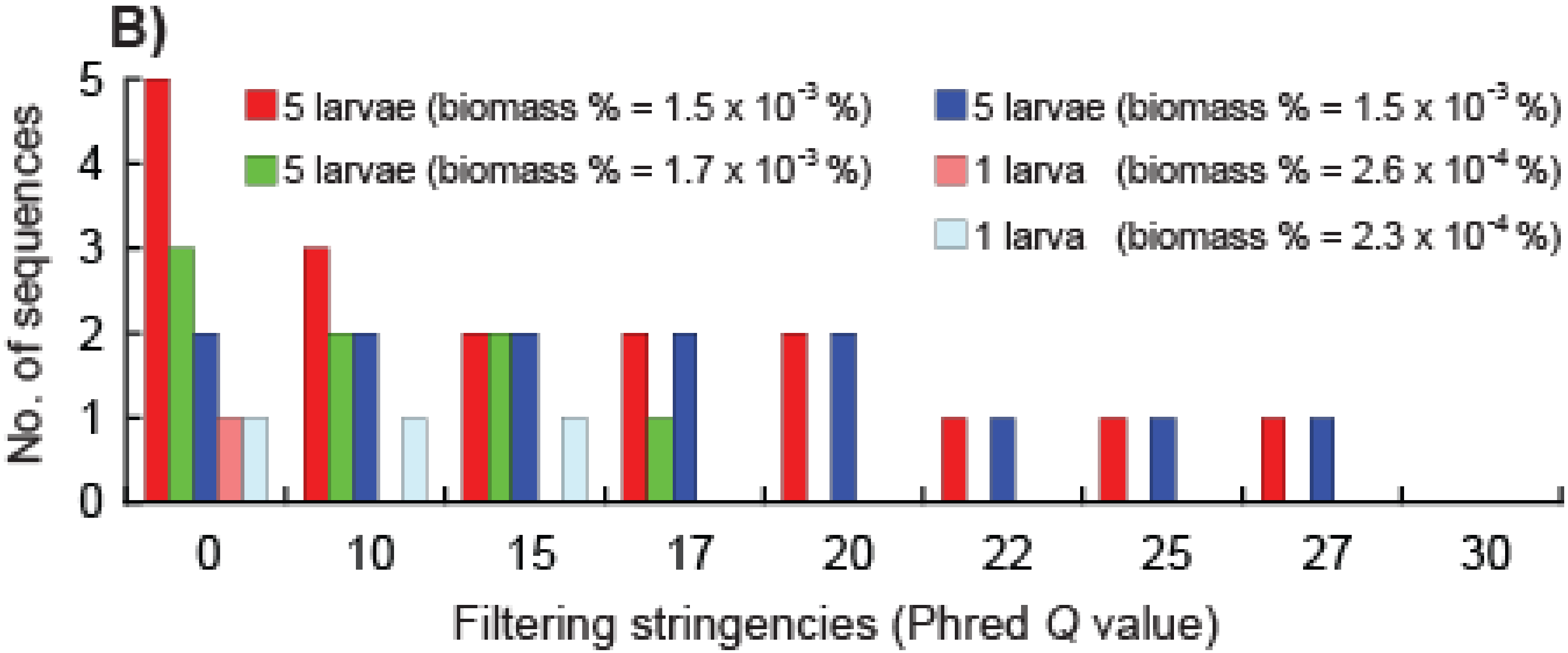
3. 滤错的影响

自然群落


A)
100%
80%
60%
40%
20%
0%
Percentage of OTUs
OTUs with no. of sequences of >100
OTUs with no. of sequences of 11~99
OTUs with no. of sequences of 4~10
OTUs with no. of sequences of 1~3
(i.e. singletons, doubletons, tripletons)
0
10
15
17
20
22
25
30
B)
100%
80%
60%
40%
20%
0%
Percentage of OTUs
OTUs with no. of sequences of >100
OTUs with no. of sequences of 11~99
OTUs with no. of sequences of 4~10
OTUs with no. of sequences of 1~3
(i.e. singletons, doubletons, tripletons)
0
10
15
17
20
22
25
30
Phred score (Q)

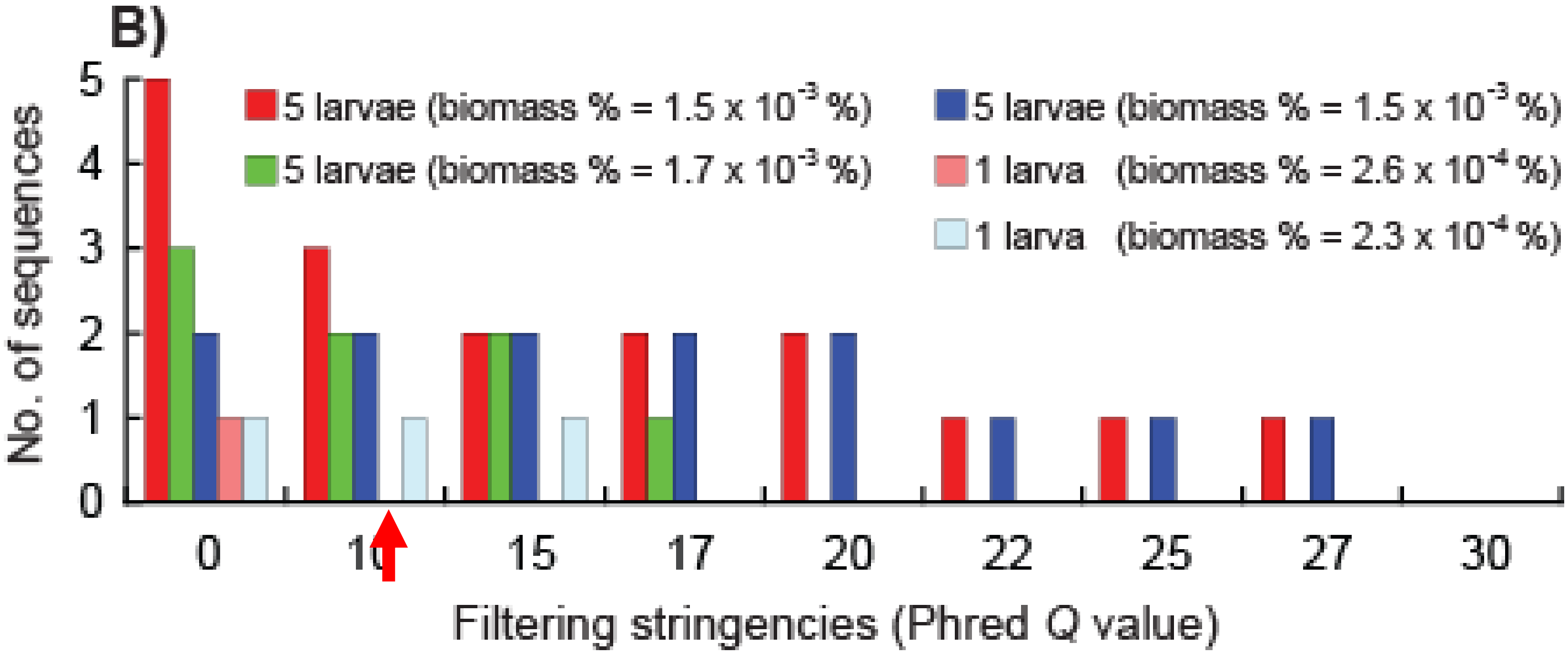
3. 滤错的影响

人工群落


B)
5
4
3
2
1
0
No. of sequences
🟥 5 larvae (biomass % = 1.5 x 10-3 %)
🟩 5 larvae (biomass % = 1.7 x 10-3 %)
🟦 5 larvae (biomass % = 1.5 x 10-3 %)
🟥 1 larva (biomass % = 2.6 x 10-4 %)
☐ 1 larva (biomass % = 2.3 x 10-4 %)
0
10
15
17
20
22
25
27
30
Filtering stringencies (Phred Q value)

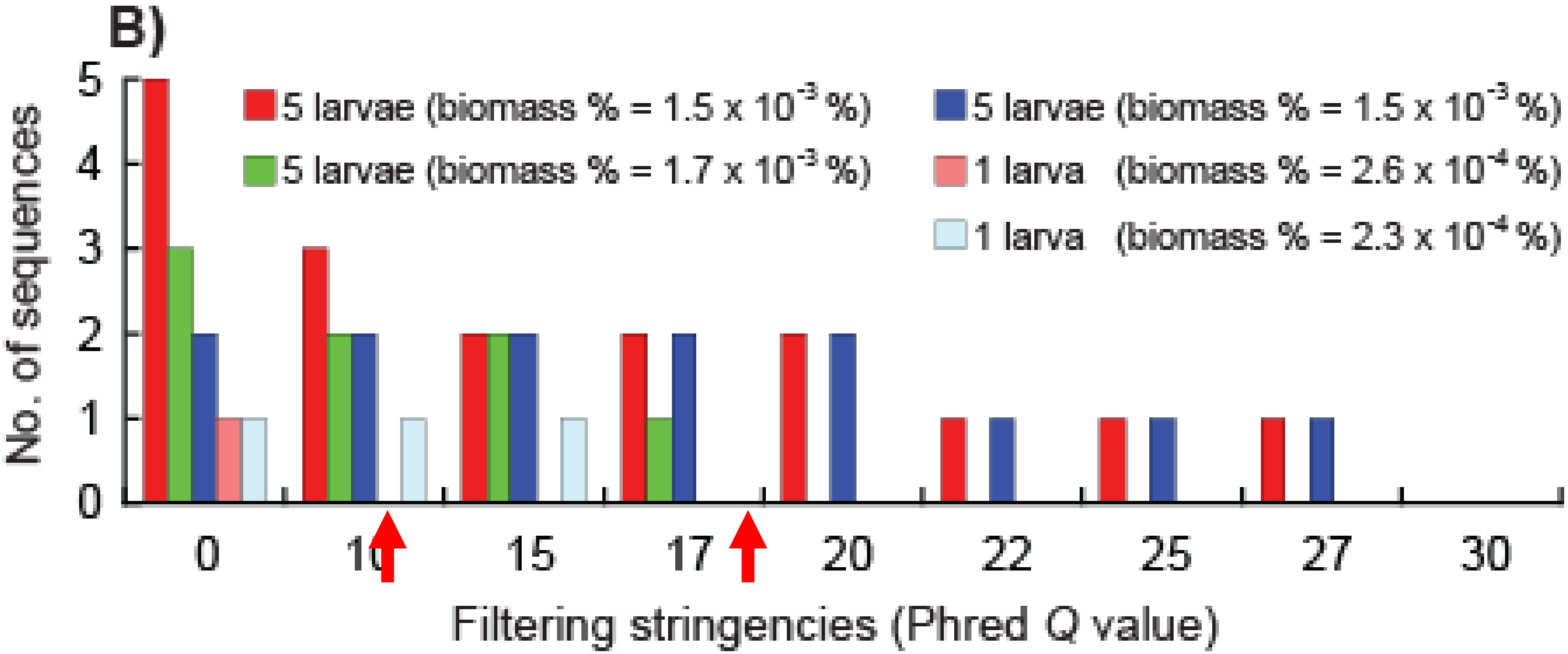
3. 滤错的影响

人工群落


B)
5
4
3
2
1
0
No. of sequences
5 larvae (biomass % = 1.5 x 10⁻³ %)
5 larvae (biomass % = 1.7 x 10⁻³ %)
5 larvae (biomass % = 1.5 x 10⁻³ %)
1 larva (biomass % = 2.6 x 10⁻⁴ %)
1 larva (biomass % = 2.3 x 10⁻⁴ %)
0
10
15
17
20
22
25
27
30
Filtering stringencies (Phred Q value)

3. 滤错的影响

人工群落



B)
5 larvae (biomass % = 1.5 x 10^{-3} %)
5 larvae (biomass % = 1.7 x 10^{-3} %)
5 larvae (biomass % = 1.5 x 10^{-3} %)
1 larva (biomass % = 2.6 x 10^{-4} %)
1 larva (biomass % = 2.3 x 10^{-4} %)
No. of sequences
0
1
2
3
4
5
0
10
15
17
20
22
25
27
30
Filtering stringencies (Phred Q value)

4. 序列丰度与生物量关系

方法一：人工群落构建

海洋生物

海湾扇贝

仿刺参

幼虫（实验室）

淡水浮游群落

淡水生物

沼蛤

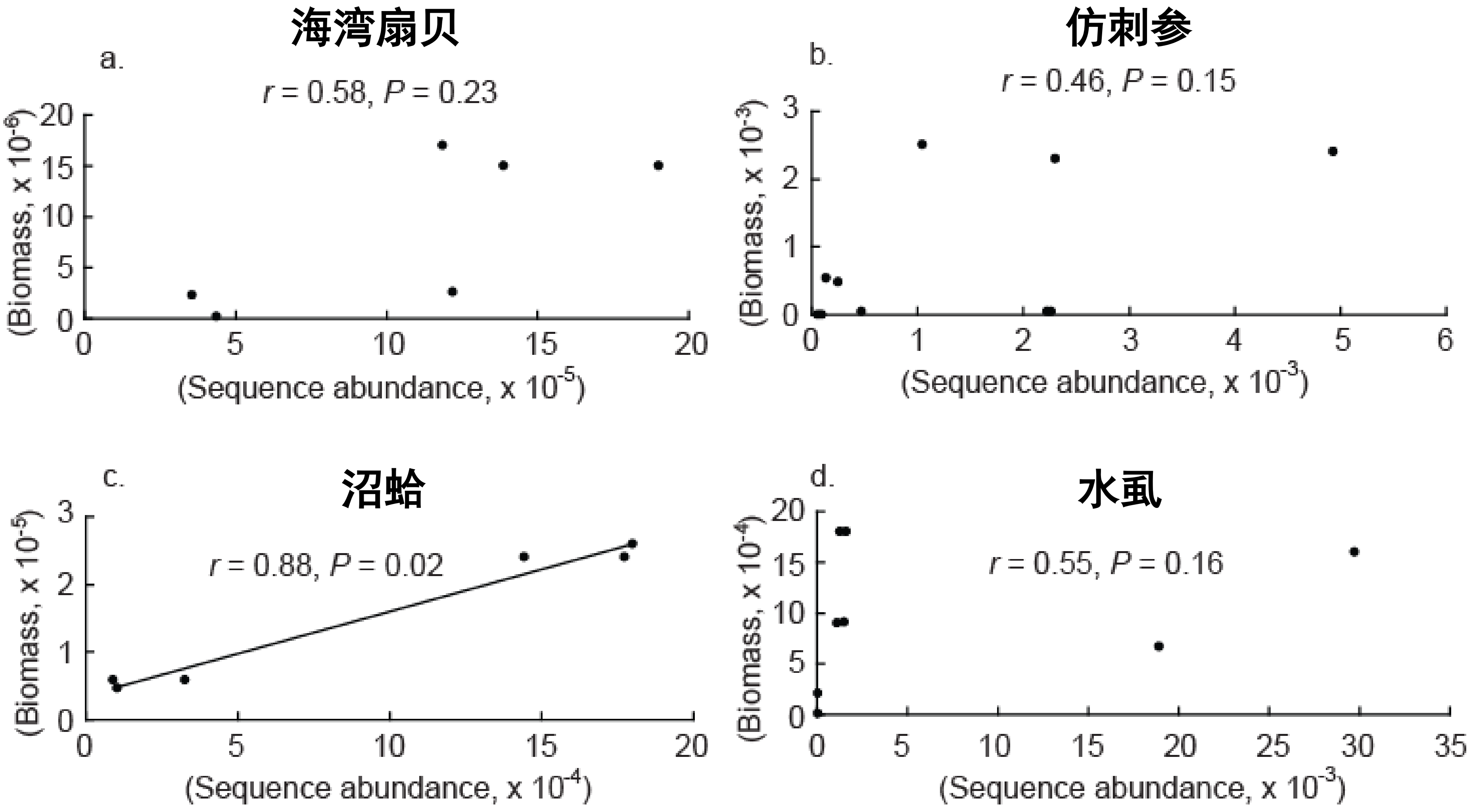
水虱

幼虫（野外）

海水浮游群落

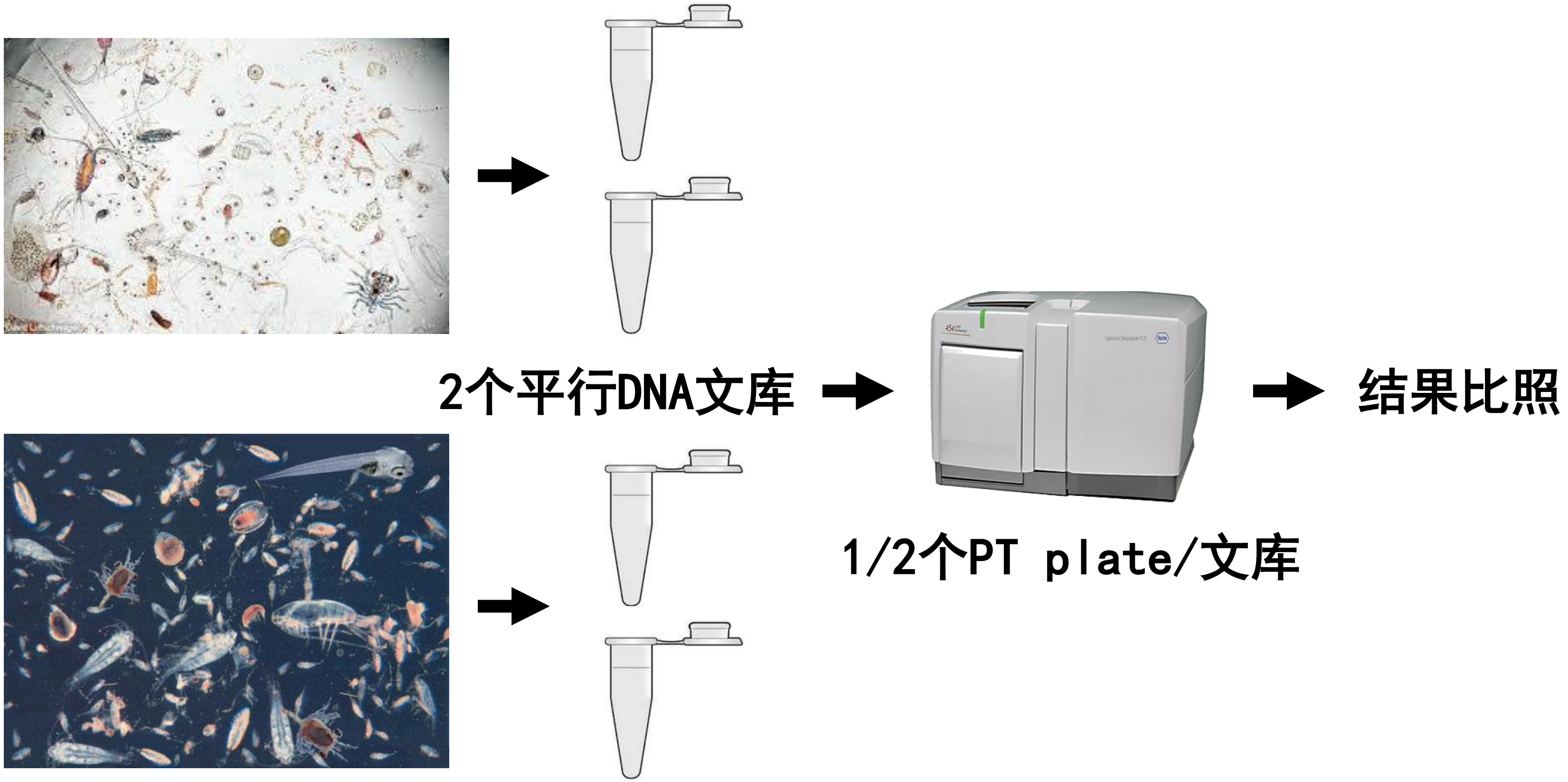
4. 序列丰度与生物量关系

方法一：人工群落构建结果



4. 序列丰度与生物量关系

方法二：设置重复



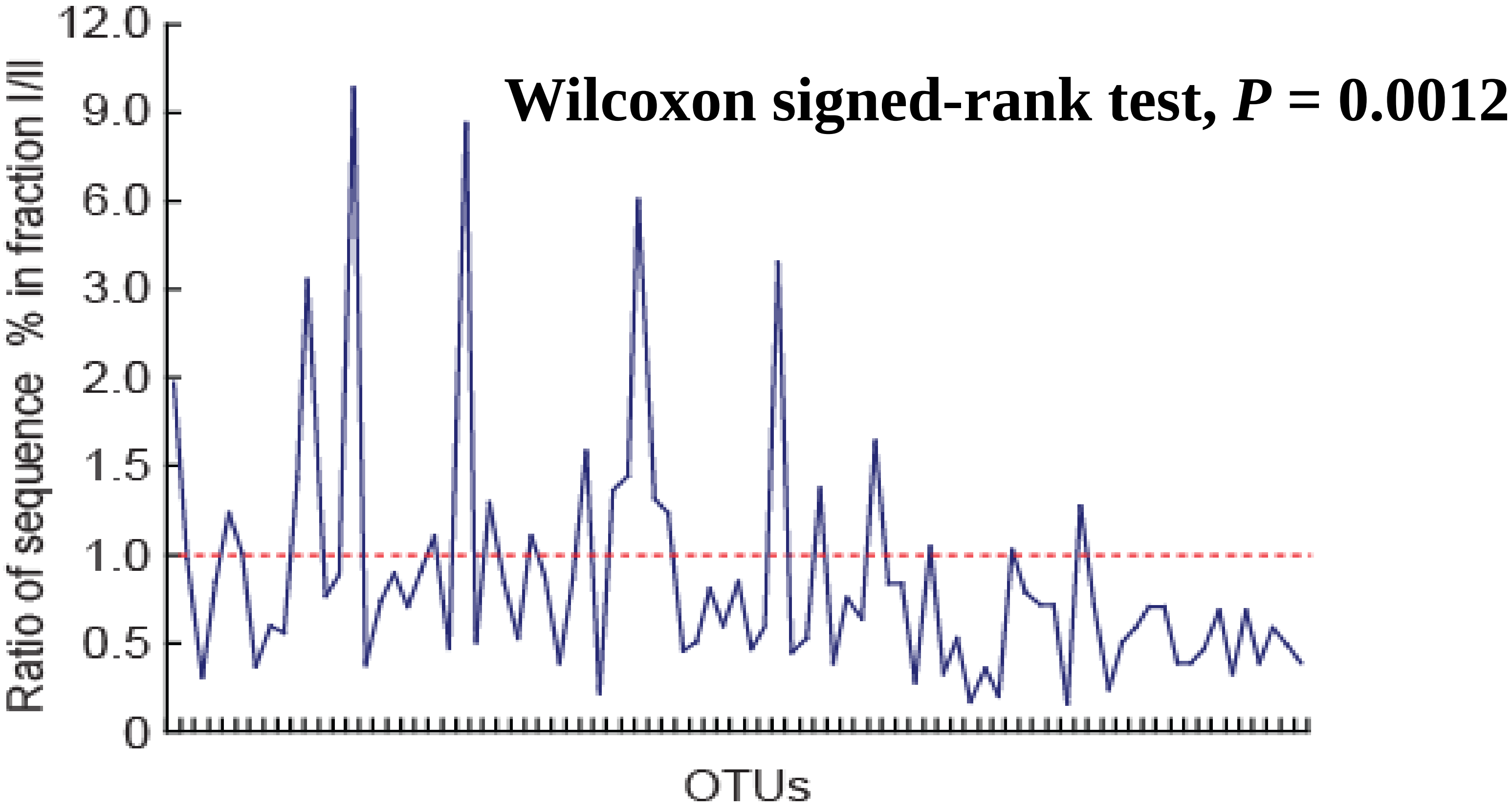
2个平行DNA文库

1/2个PT plate/文库

结果比照

4. 序列丰度与生物量关系

方法二：设置重复的结果


Wilcoxon signed-rank test, *P* = 0.0012

结论：1. 序列丰度不能定量量化生物量

2. 但序列丰度基本可定性反应生物量

致谢及相关文献

NSFC

中国科学院 CHINESE ACADEMY OF SCIENCES

国家海洋局 STATE OCEANIC ADMINISTRATION

1. Zhan A. et al. 2013. Methods Ecol Evol, 4: 558-565

2. Zhan A. et al. 2014. Mol Ecol Resour, 14: 1049-1059

3. Zhan A. et al. 2014. Methods Ecol Evol, 5: 881-890

4. Zhan A. et al. 2014. PLoS ONE, 9: e96928.

5. Sun C…. Zhan A. et al. 2015. Aquat Biol, Accepted