Fungal screening and phylogenetic tree construction

Jianyang Zhou1,

1Institute of Marine Biology, Ocean College, Zhejiang University, No. 1 Zheda Rd., Zhoushan 316021, China.

Abstract. The second-generation sequencing technology has achieved rapid development in recent years. It has the advantages of fast speed, high accuracy and low cost compared to the former sequence, so it has a widely application. Marine fungi are important sources of active Marine natural products. Due to the differences in the thermodynamic properties of land and sea and the concentration of salt oxygen, it has a great potential to produce new bioactive substances. Nowadays, scientists isolate and identify a large number of secondary metabolites with novel structures which is from marine fungi, and these compounds show excellent biological activities. For example, some are potential drugs to treat cancer. In order to search for new fungi, this article used a series of biological methods to isolate and study the bacteria isolated from huimin wharf and determine their specific species. And this paper also uses 18S rDNA to analyse the gene sequence and construe the phylogenetic tree.

1 Introduction

1.1 High-throughput sequencing

High-throughput sequencing technology includes several mainstream technologies including Roche/ 454 pyrosequencing (2005), Illumina/Solexa polymerase synthesis sequencing (2006), ABI/SOLiD ligase sequencing (2007) and Thermo Sciences/Ion Torrent semiconductor chip sequencing (2010)1,2. It uses parallel sequencing technology, and a single run can complete millions or even tens of millions of sequencing experiments, with a large amount of sequence data (depending on the specific application, the sequencing flux is between 1gbp-1tbp), and compared with the former sequence, the average sequencing cost is extremely low3.

One of the important applications of high throughput sequencing in microbial research is the metagenome research. In 1998, a scientist called Handelsma first proposed the concept of Metagenome, and believed that the study of Metagenomics should focus on the total genome of bacteria and fungi in environmental samples. Metagenomics considers the genetic information of all microorganisms in the environment as a whole and studies the relationship between microorganisms and the natural environment or organisms from top to bottom. It not only overcomes the difficulty of microbial culture, but also opens up a new way to fully understand the ecological characteristics and functions of microorganisms from the level of community structure by combining bioinformatics methods4. Macro genome sequencing is using all microbial genome sequencing technology to test some environmental samples, and to analyse the genetic composition and function of the microbial community, then interpret the diversity and richness of microscopic objects groups, so as to explore the relationship between microorganisms and the environment and host, explore and research the new specific work, etc5. Compared with other technologies, high-throughput sequencing technology has obvious advantages in macro-genome research6. It has high accuracy, sensitivity and comprehensiveness in the detection and identification of microorganisms in the community. It can accurately quantify the biological quantity of a certain microorganism and discover new microorganisms.

1.2 16S rRNA Gene sequence analysis technique

16S rRNA is a ribosomal RNA of prokaryotes system. At present, the bacteria taxonomy study of the most useful and the most commonly used molecular clock is rRNA. rRNA is less kinds and content of large (about 80% of total bacteria RNA), molecular size than normal RNA7]. And in the long evolution process, its genetic sequence change very slowly, so could be used for marking some biological evolutionary distance or affinity. What’s more, rRNA’s structure and function are highly conservative.

16S rRNA has many characteristics8. First of all, it has Multiple copies. Every bacterium has five to ten copies of 16S rRNA, which can be detected more sensitively. Secondly, it can contain more information. In general, the 16S rRNA have the variable and conserved region. The conserved areas were extremely common to all bacteria, and the variable areas were different from each other to some extent. Therefore, scientist can design some common primer to a large number of bacteria according to the important conserved region. Furthermore, researcher can also design some particular makers or primers to the particular bacteria because of the variable region. Thirdly, generally speaking, the 16S rRNA’s
length is 1,500 bp or so long and contains about 50 functional domains.

16s rRNA gene sequence analysis technology is extracted from microbial samples. Researchers can clone and then sequence the 16s rRNA genes. When they accomplish this, they can also get much information about this gene through enzyme digestion and probes hybridization. Then compared with 16s rRNA sequences in the database or other data comparison, scientists are able to know what’s kind of the microbe is.

2 Isolation and identification

2.1 Experiment Material and apparatus

2.1.1 Material

Marine fungi intertidal sediments samples from zhoushan city, zhejiang province, in 2018, collected from huimin bridge pier (30 ° 00' N, 122 ° 10' E).

2.1.2 Reagents

Yeast genome extraction kit (Omega DNA kit, 50 times), 5× TBE electrophoresis buffer agarose, 2× Mix Taq enzyme.

2.1.3 Medium and Antibiotics

YPX medium, YPD medium, ampicillin solution and kanamycin solution.

2.1.4 Main Instruments and Equipment

Centrifuge 5424, Eppendorf; PCR Amplifier Nexus ×2, Eppendorf; electronic balance BSA124S, Sartorius; Vertical pressure steam sterilizer, Systec V-40, Systec; constant temperature incubator, DHG-9123A, jinghong company; ultra-clean bench, BBS-DDC, BIOBASE; Shaker Infors Ag Ch-4103 Bottmingen, INFORS HT.

2.2 Method

2.2.1 Strain isolation and purification

1. Take 100 mL to 200 mL of the sediment to conical flasks and cultured at room temperature for 3 days.
2. draw 1 mL of the culture with a pipette and dilute it with normal saline at 10 times gradient (10⁻¹, 10⁻², 10⁻³).
3. prepare three YPX and YPD solid plates, Cover with rubber plug high-pressure steam sterilizer and sterilize at 121 °C for 20 min,then take 0.1 ml product to each plate.
4. Observe the fungus growing on the plate and record its morphology.
5. Select Single colonies of yeast-like fungi and strew them on the plate to purify the single colonies. The single colonies were cultured in a 25-degree incubator for 3 days.
6. A small number of cells were picked up and observed with an optical microscope to depict the cell morphology.

2.2.2 Extraction from Marine fungi

1. Select the single colony and incubate it at 25 in 25 mL liquid YPD for 3 days.
2. Use the Omega DNA extraction kit to extract genomic DNA.
3. Put the yeast into YPD medium until the OD600 is between 0.8 and 1.0, then centrifuge 3 mL this mixed or so at 6500rpm about 10 mins. (Use centrifuge 5425, Eppendorf, the same below)
4. Discard the supernate and add 480μL SE Buffer, 10μL 2-mercaptoethanol and 20 μL lyricize solution to make cells resuspend. Then place it in 30°C for 40 minutes or so and centrifuge it about 5 minutes at 6500rpm.
5. Add 200 μL YL Buffer to the tube and gently blow and beat the cells. After doing that, add 50 mg Glass Beads into the tube.
6. Take a new EP tube and use pipette to draw the supernatant into it. Then put 25 μL Proteinase K Solution into the tube, too. Gently blow and beat the mixed then incubate at 65°C.
7. Use pipette to take 5 μL RNase A and mix them together then let them react for about 10 min at room temperature.
8. Use pipette to take 220 μL YDL Buffer and absolute ethanol to the sample and also mix them. Follow the induction and extract the target DNA.
9. Store DNA at -20°C.
10. Extract the target DNA by 1% agarose gel electrophoresis: 50 mL 0.8% agarose gel with TAE (including DNA-dye) in a 200 ml flask, boil it in the microwave oven until the liquid is completely transparent without any lumps, cool, and pour it into an agarose gel mould with a narrow well comb.
11. After gel solidification (about 20-30 min), transfer gel into gel tank. Mix 5 μL (50-500 ng) DNA sample with 2 μL DNA loading buffer on a piece of clean plastic paper using a 10 μL pipette, and transfer it into a well by 10 μL pipette. Add 5 μL DNA ladder DL5000 into another well.
12. Run gel for 20-30 min at about 200V. Remove gel from gel tank, image and save photo. Keep the remaining genomic DNA sample at -20°C-80°C until analysis. Frozen DNA samples are stable for more than one year without frequent freezing and thawing.

2.2.3 Marine fungi were identified by 18srDNA sequencing

1. Use PCR with 2×Mix Taq enzyme to amplify 18sr DNA, Primer:
   18S1:ATCTGGTTTGATCCTGCCAGT,
   18S2:GATCCTTCCGCAGGTTCACC
Table 1. PCR amplification procedure.

| Purpose          | Temperature | Time  | Cycle |
|------------------|-------------|-------|-------|
| Pre-denaturation | 95 °C       | 5 min |       |
| Denature         | 95 °C       | 30 sec|       |
| Primer annealing | 52°C        | 30 sec| 30 cycles |
| Extension        | 72°C        | 2 min |       |
| Final extension  | 72°C        | 10 min| add A base in the end of product |

1) According to the introduction of PCR Cleanup kit, add a volume of buffer PCR-A that is 3 CR reaction volume to the mixture. If the PCR mixture are in several tubes, collect them and add buffett PCR-A.

2) Set a Spin Column into the tube, and take mixture (S700 BL once in one Spin Column) from Step I to the Spin Column, then centrifuge at 10,000 rpm about 1 mins.

3) To improve recovery rate of DNA, transfer the liquid to Spin Column again and centrifuge again. If necessary, repeat this step 2-3 times.

4) If the mixture are in several tubes, combine the solution of these tubes into one Spin Column by repeating step 2-3.

5) Add 700 mL of buffer W2 into the Spin Column. Then centrifuge at 10,000 rpm for 1 min or so. Discard the flow-through. Repeat this procedure again and then centrifuge at 10,000 rpm for about 1 min. Abandon the flow-through.

6) Centrifuge the empty Spin Column and Collection Tube at 10,000 rpm for 2 mins so as to get rid of residual liquid.

7) Put the Spin Column into another clean EP tube. Add 30 over-heated ddwater (or Eluent) to the spin column membrane. Put it at atmospheric temperature at least 5 min.

8) Centrifuge at 10,000 rpm for 2 min in order to elute the DNA. Take 2-5 purified DNA to run gel to check if you obtain the target DNA. Keep the purified DNA at -20 or -80 oc until further use.

2. Run 1% agarose gel electrophoresis to observe PCR amplification products and send them to the company for sequencing.

3. Blast was used to identify the microbial species.

4. Use Mega 7 to build up phylogenetic trees.

3 Result

![Fig. 1. Gel electrophoresis](image1)

This figure is the result of DNA extraction. After the gel electrophoresis, we can clearly see that the extracted DNA is very pure, and the base pair is between 1000 and 2000. As a result, we get the target DNA sequence and take it to the company for next gene sequencing.

We can see that although the correct nucleotide cannot be detected at the beginning. It is because that the ddNTP labeled fluorescently was excessive during sequence, and after the end of the reaction, they would precipitate and interfered with the signal of a few nucleotides behind the primer. The type of nucleotide can be clearly determined later. However, after 800bp it would also mixed. So we should sequence them twice and then joint them together, at last we can get the true data. This sequence conclude 1028bp.

When we get the data, change one of the sequence into Inverse complement sequence and then joint it and another sequence together. Then we blast the detected genome in NCBI and found that this species is a known candida, which has also been found in the south China sea.

![Fig. 2. The result of gene sequence](image2)
4 Conclusion

The team collected Marine sediment samples at a fixed point and used dilution coating method to isolate the microorganisms, and then enriched and cultured them, so as to build a bacterial strain resource pool of different Marine environmental sources. As part of the establishment of strain pool, this paper isolated Marine sediment samples from huimin bridge wharf, and selected a strain of yeast for analysis and identification. The 16s rRNA was extracted for PCR amplification and sequencing, and the approximate sequence was obtained. It was then compared in BLAST gene bank and found to be most similar to KY118177.1 Candida parapsilosis, a strain belongs to Debaryomycetaceae. Of course, this paper also has some shortcomings. For example, due to time, we cannot expand the culture to extract its active metabolites. And beacuse of the shortage of time, this paper didnot get the sequence of TA clone. It is widely acknowledged that the TA clone for target DNA is able to get more brilliant sequence.

Reference

1 Johnston, P. R., Park, D. & Smissen, R. D. Comparing diversity of fungi from living leaves using culturing and high-throughput environmental sequencing. *Mycologia* 109, 643-654, doi:10.1080/00275514.2017.1384712 (2017).

2 Kircher, M. & Kelso, J. High-throughput DNA sequencing--concepts and limitations. *Bioessays* 32, 524-536, doi:10.1002/bies.200900181 (2010).

3 Huss, M. Introduction into the analysis of high-throughput-sequencing based epigenome data. *Brief Bioinform* 11, 524-536, doi:10.1093/bib/bq014 (2010).

4 Ni, T., Wu, H., Song, S., Jelley, M. & Zhu, J. Selective gene amplification for high-throughput sequencing. *Recent Pat DNA Gene Seq* 3, 29-38, doi:10.2174/187221509787236183 (2009).

5 Budowle, B. *et al.* Validation of high throughput sequencing and microbial forensics applications. *Investig Genet* 5, 9, doi:10.1186/2041-2223-5-9 (2014).

6 Di Bella, J. M., Bao, Y., Gloor, G. B., Burton, J. P. & Reid, G. High throughput sequencing methods and analysis for microbiome research. *J Microbiol Methods* 95, 401-414, doi:10.1016/j.mimet.2013.08.011 (2013).

7 Matsumoto, T. & Sugano, M. [16S rRNA gene sequence analysis for bacterial identification in the clinical laboratory]. *Rinsho Byori* 61, 1107-1115 (2013).

8 Deutscher, M. P. in *Progress in Molecular Biology and Translational Science* Vol. 85 369-391 (Academic Press, 2009).