16S rRNA probe design for HCR-FISH V.2

Kayley Hake[1]
[1]University of California, Berkeley

Version 2
Apr 19, 2021
1 Works for me
This protocol is published without a DOI.
King Lab

Kayley Hake
University of California, Berkeley

ABSTRACT

This protocol outlines how to design HCR-FISH probes targeting 16S rRNA sequences. It covers downloading and installing software (ARB, MacPorts, XQuartz), importing SILVA 16S rRNA database, adding new 16S rRNA sequences, aligning 16S rRNA sequences, and generating custom HCR-FISH probes. Many of these ideas are adapted from Amy Apprill's ARB Tutorials (Arb-Home Documentation Page), and Philip Hugenholtz in Methods in Molecular Biology

PROTOCOL CITATION

Kayley Hake 2021. 16S rRNA probe design for HCR-FISH . **protocols.io** https://protocols.io/view/16s-rrna-probe-design-for-hcr-fish-wdffa3n

KEYWORDS

HCR-FISH, 16S rRNA, ARB, SILVA, choanoflagellate, bacteria

CREATED

Dec 12, 2018

LAST MODIFIED

Apr 19, 2021

PROTOCOL INTEGER ID

18567

BEFORE STARTING

- Additional information regarding ARB and probe design can be found at the ARB-SILVA website.
- Published 16S rRNA probes can be found at ProbeBase.
- Many of these steps were adapted from Amy Apprill's Tutorials which can be found at the bottom of the Arb-Home Documentation Page, and Hugenholtz et al. from Methods in Molecular Biology

Installing ARB, MacPorts, and XQuartz Software

1

1. Download zip file of ARB software for your operating system from the ARB website.
2. Download the MacPorts software through the MacPorts website.
3. Install MacPorts. (Requires Xcode for Mac which can be downloaded in the App store)
4. Once MacPorts is installed, open a new Terminal.
5. In the terminal window write the following command and let it run.

```
sudo port install arb
```

6. Before ARB is operational, you have to set ARBHOME and add ARB to your PATH. To do this, open a new terminal window and type the following command. This sets the necessary environment variables automatically.

```
arb_macsetup
```

7. Download XQuartz software to run ARB from the XQuartz website.
8. To test if everything is operational, open a new terminal window and type the following to open ARB. XQuartz should automatically launch and run the ARB software.

```
arb
```

Importing SILVA 16S rRNA databases

2

1. Create a new folder in the Documents section of your computer called ARB where you will keep all your Arb-related files.
2. Go to the SILVA website, under the Download tab at the top, and select ARB Files.
3. Download the .gz file titled **Ref NR 99** which is the recommended 16S/18S reference database for ARB (Also linked here Ref NR 99).
4. Move the downloaded file to your ARB Folder in the Documents section and double click to open the fine uncompress it.
5. You should now have a file called SILVA_132_SSURef_NR99_13_12_17_opt.arb in your ARB Folder.
6. Open ARB Software by typing ' arb ' in a new terminal window.
7. Naviagate to your ARB folder in your Documents section through the software and open the .arb file to open the database in arb.
8. Re-save it as a new database that you will be modifying in future steps.

Importing new 16S sequences to the database

3

1. To generate probes to novel bacterial sequences, it is important to add the specific 16S sequence you are targeting. To do this, first generate a FASTA file using the program Text Edit which is found on your Mac. For a PC, you will have to look up otherways to make a FASTA file like using EMACS.
2. Look up the FASTA sequence on NCBI (For example: *A. machipongonensis* 16S rRNA is NR_121685.1).
3. Copy the FASTA sequence into a new Text Edit document. Make sure it has the proper FASTA format with the first line including ' >name of sequence '.
4. Under the Format Tab, click change to Plain Text.
5. Save File to your ARB Folder. At the end of the file name add ' .fasta ' and un-check the box that says ' If no extension is provided, use ".txt" '. This will save the file as a FASTA file.
6. In ARB, with the new database opened, click > File, > Import, > Import from external format.
7. Once ARB Import pops-up, navigate to your ARB Folder and select your FASTA sequence.
8. Under Import selected format, select 'fasta.ift' and than 'Go'.

9. A question box will open, select 'Generate unique species IDs'.
10. The box SEARCH and QUERY should open once it's imported with the information for the sequence. Record the species ID ('acc') for your record. This is how the sequence will appear later in the tree.
11. Click "Write to Fields of Listed"
12. Scroll to the field name 'author'
13. In the 'Enter new field value:' enter your intials and click 'Append' . This is a nice feature that allows you to quickly search for imported sequences by searching your initials.
14. Save the database as a new database and give it a name. Everytime you add a sequence you will want to save.

Aligning new 16S rRNA sequence in the database

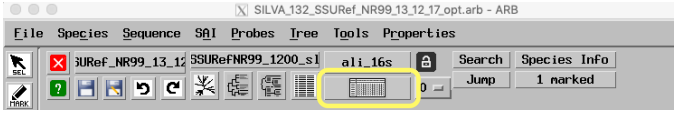
4 1. In the SEARCH and QUERY box, select your sequence in the HITLIST and select "Mark Listed Unmark Rest'. This highlights your sequence and no other sequences in the database. The ability to "Mark" a sequence is an important feature that you will use throughout ARB to let the software know what sequences to focus on. You can also manualy "mark" sequences using the buttons along the left-hand side of the main window, scrolling through the tree, and clicking on sequences of interest. A change in text color illustrates sequences are marked vs. unmarked.
2. Close the SEARCH and QUERY, and under the Probes tab, select >PT_SERVER Admin. This ADMIN allows you to build a server to align sequences against. Anytime you add a sequence, change an alignment, or add a sequence to a tree, you need to build or update the server.
3. In the PT_SERVER Admin, select a PT_SERVER Template like user1.arb, and remember which one you selected.
4. Next select 'Build server'.This can take several hours if you are building it from scratch. To update an existing server when you've made a change, select your current server and select 'Build server'. This will take less time compared to building from scratch. You can follow the progress in the Terminal. Close PT_SERVER ADMIN when it's done.
5. To align your sequence of interest against the server you need to open the aligner. In the main ARB program, select the button under 'ali_16S'. See image below.
6. 

Image depicting the button to press to open the aligner.
7. The ARB_EDIT4 *1* [ali_16s] window should open after pressing the button above. In the center of this window, select the button that says 'Aligner' to open 'Integrated Aligners'.
8. Select the following settings in the Integrated Aligner window:
9. ***Aligner*** > Fast aligner
10. ***Align what?*** > Marked Species
11. ***Refernce*** > Auto search by pt_server. In the box, switch 'probe-server.arb' to the **server you built** (For this example, select user1.arb)
12. Change the ***Number of relatives to use*** from 1 to 10.
13. ***Range*** > Whole sequence
14. Everything else should be default settings.
15. Select ' GO ' and let the aligner run. A status box should open to let you know when it's done.

Adding new alignment to the database and tree

5 1. While the species is still marked in the ARB software, select the tab >Tree in the main window followed by >Add Species to Existing Tree.
2. Select >ARB Parsimony (Quick add marked)
3. The SET PARSIMONY OPTIONS should open with the tree already selected and the alignment ali_16S highlighted. Select ' GO '.
4. A status box will open saying ' loading tree ' and close when it's done.
5. To visualize where your sequence was added to the tree, select the tab >Tree, >Collapse/Expand Tree, >Group all except marked. This will collapse the tree containing un-marked species leaving the marked species and close relatives expanded.
6. Update your server by selecting >Probes, >PT_SERVER Admin.
7. Select your previous server, and select 'Build server' to replace it with the updated database with the new aligned sequences. (Took about 30 min. to build over the old server)

Designing probes for specific 16S rRNA sequences

6 1. To design a probe, first "mark" all the sequences in the tree you want to design a probe against. You can switch the mouse between zoom, select, and mark by clicking the buttons on the left hand side. (Note 1: your newly added sequence should still be the only one marked unless you clicked something else. For this example, keep your added *A. machipongonensis* sequence, and select two sequences related to it that are the same species. (1) AAXU02000001 and (2) EU636233, Algorphagus sp. PR1 (2008))(Note 2: In ARB, you can mark as many or as little species as you want in this step. This may be advantageous for you future experiements to design probes against entire genera, families, or orders.)
2. With your three sequences "marked", under >Probes, select >Design Probes.
3. Change the PT-Server to your updated server. In this case it should user1.arb.
4. The design parameters allow you to adjust probe length and annealing temperature. By default, the probe length is 18nt which is optimal for 16S rRNA FISH. The length of output is how many potential probes it will return. The Max. non group hits allows you to determine how specific your probe is. Having it set to zero means it will not generate probes that recognize sequences you do not have "marked." If you are not able to generate probes with zero non-group hits, you can increase this value and identify which non-marked species your probe is recognizing in a later step to decide how it may affect your experiement. The Min group hits (%) allows you to adjust how many of your selected sequences are being recognized by potential probes.
5. Change the Min group hits (%) to 100, and press ' GO '. If you get no probes with your own sequences, adjust the Min group hits (%) and the Max. non group hits.
6. A status box will open showing the progress of the probe design.
7. A PD RESULT window will pop-up when the probe design is finished. At the top is a summary of the paramteters and results followed by a list of probes in order of quality. The left column is the target sequence followed by the length. Next is an alphabetic letter which helps organize the probes by grouping probes under the same letter that are within a few nucelotides of the same target region. A + and - following the letter shows change in placement either upstream or downstream of the first lettered probe that does not have a + or -. The 'ecol' is the position of the target sequence in respect to the 16S rRNA from *E. coli*. I often reference the *E. coli* position because probe placement on the *E. coli* rRNA can affect probe acessibility. Although it is not completely compatable between bacterial species, there are some clear regions that are not ideal for probe placement. See attached paper for more information and to reference their accessibility maps.
   Behrens_et_al-2003-Applied_and_Environmental_Microbiology.pdf
8. In the probe sequence column for the *A. machipongonensis* example, the first probe is **UACUGGCCAGUACAAUGC**. If you select the probe, and click the button 'MATCH' at the top, it opens a PROBE MATCH window with the respective target sequence already written in the Target String box.
9. Click 'MATCH' in the Probe Match window to blast the target region across the database. This will help confirm the probe specificity and what other non-marked sequences it may be recognizing. A list of target species will load, and for the *A. machipongonensis* example, only the top three should be a perfect match. The following target species, have atleast one mismatch and increase in mismatches as you go down the list. Although HCR-FISH is a robust method, you want probes with as little mismatches towards other sequences (at a minimum, your probe should

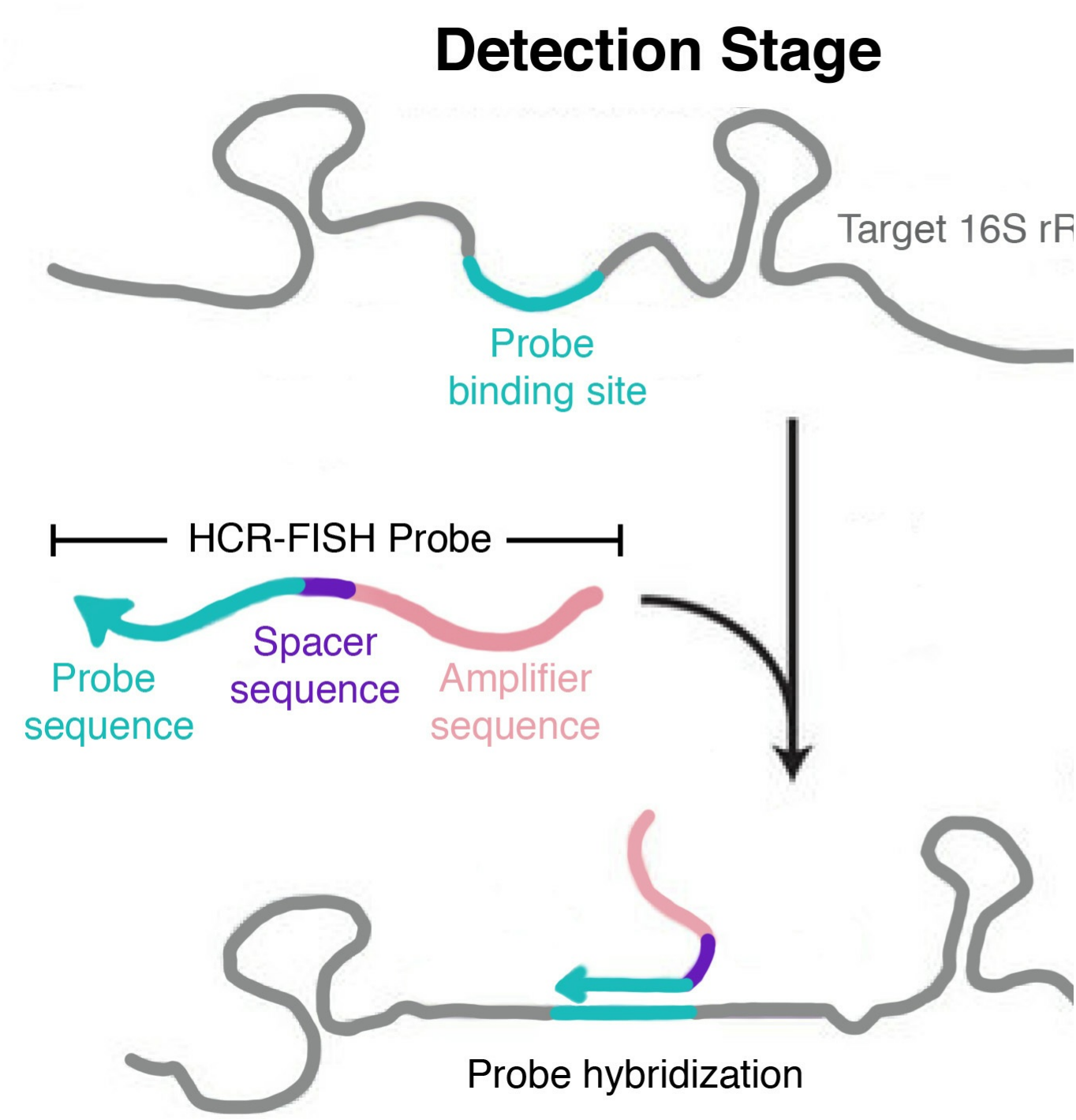
have at least 3 mismatches with non-target species). (If you have a mismatch of 1 with a non-target species, you can look into making a competitor probe which is an unlabeled probe with the exact sequence of the mismatch non-target. You can read more about competitor probes here.)

Selecting the right HCR hairpin and spacer for a new probe

7 1. The complete HCR probes is: **[Probe Seq. 18nt]+[Spacer Seq. 5nt]--[Hairpin Amplifier Sequence 36nt].** Until this point, you have only determined the Probe sequence using ARB. Now you need to determine which hairpin corresponding fluorophore you want to use with this probe and add on the respective spacer sequence and hairpin amplifier sequence. This is important because the hairpins that contain the fluorophore recognize the amplifier sequence on your probe and assemble off of it resulting in the fluorescent signal during the HCR-FISH Protocol. See image below for a diagram of how the hairpins assemble off of the amplifier sequence in the probe.
2. In the King Lab we have three sets of hairpins you can use with your probes with three different fluorophores (B1 - 488; B2 - 594; B3 - 647)(**Table 1** below).
3. There are two things to consider. (1) What color fluorophore do you want your probe to be? and (2) Will the sequence upstream of the probe target binding site anneal to the spacer sequence? If the immediate nucleotides (especially the first 1-2) upstream of the target region binds to the spacer, this may affect the ability of the hairpins to bind to the amplifier sequence and properly assemble.
4. To determine the sequence upstream of your probe, return to the PROBE MATCH window where the probe was tested against the database.
5. In the list on the left you can see the alignment of the target sequence, and the neighboring sequences. The sequence immediate to the left of the target is the sequence that has potential to bind to the spacer sequence. In the case of the *A. machipongonensis* example, the next three nucleotides upstream are GAA. Because the first nucleotide is a G, any of the spacers (all the spacers begin with either A or T, so a G or C is fine). If the next nucelotide had been a T, hairpin amplifier B3 would be the better choice because the spacer also begins with a T compared to B2 which begins with an A and would result in additional binding beyond the probe sequence.
6. Probes are ordered through IDT with a final length of 59 nucleotides in the correct 5'-3' orientation of Probe Sequence followed by the respective spacer and amplifier sequence.
7. Finishing with the *A. machipongonensis* example, you first identified a probe you liked (UACUGGCCAGUACAAUGC). Next, convert the U's to T's (TACTGGCCAGTACAATGC). Finaly, to make this probe compatable with the B2 amplifiers, add on the respective sequences in the table resulting in the final probe oriented 5'-3': TACTGGCCAGTACAATGCAAAAAAGCTCAGTCCATCCTCGTAAATCCTCATCAATCATC.

Hairpin Name	Fluorophore	PROBE SEQ + SPACER -- AMPLIFIER SEQUENCE
B1	488	PROBE SEQ + ATATA -- GCATTCTTTCTTGAGGAGGGCAGCAAACGGGAAGAG
B2	594	PROBE SEQ + AAAAA -- AGCTCAGTCCATCCTCGTAAATCCTCATCAATCATC
B3	647	PROBE SEQ + TAAAA -- AAAGTCTAATCCGTCCCTGCCTCTATATCTCCACTC

Table 1: The King Lab Hairpins available with their fluorophores and respecitive spacers and amplifier sequences.



Probe layout and overview of HCR-FISH targeting 16S rRNA. Image modified by Kayley Hake from Choi et al. (2010) Nature Biotechnology.