HHS Public Access
Author manuscript
Cold Spring Harb Protoc. Author manuscript; available in PMC 2023 October 20.

Published in final edited form as:
Cold Spring Harb Protoc. ; 2023(1): 22–34. doi:10.1101/pdb.prot108067.

Isolation and Sequencing of Novel *Vibrio* Species

Andrew Camilli[1]
Department of Molecular Biology & Microbiology, Tufts University, School of Medicine, Boston, Massachusetts 02111, USA

Abstract

Whole-genome sequencing of bacteria facilitates their genotypic analysis and expands our understanding of the tree of life on Earth. The *Vibrio* genus comprises many halophilic species of bacteria, including some that are pathogenic to humans and other animals. Here I describe methods for isolating and sequencing both known and novel species of *Vibrio* from saltwater environments. The first section outlines methods of isolating and phenotypically characterizing strains, followed by purification of their total DNA. The second and third sections outline methods of whole-genome sequencing and assembly, annotation, and phylogenetic analysis.

MATERIALS

It is essential that you consult the appropriate Material Safety Data Sheets and your institution's Environmental Health and Safety Office for proper handling of equipment and hazardous materials used in this protocol.

RECIPES: Please see the end of this protocol for recipes indicated by <R>. Additional recipes can be found online at http://cshprotocols.cshlp.org/site/recipes.

Reagents

Molten agarose <R>

Use 0.8% agarose for Step 27 and 1.5% agarose for Step 41.

Reagents required for isolation of culturable bacteria of the *Vibrio* genus procedures only (Steps 1–9)

CHROMagar *Vibrio* (Chromagar VB912)

Make the plates with CHROMagar Vibrio and Petri dishes (150-mm-diameter) according to the manufacturer's instructions. Let the plates dry for 2–7 d at room temperature, depending on ambient humidity, so that liquid suspensions spread on the surface can dry into the agar plate. Alternatives to CHROMagar Vibrio for isolating Vibrio species exist, such as mTCBS (Meier-Kolthoff et al. 2022).

Glycerol (Fisher Scientific G33)

Reagents required for purification of total DNA procedures only (Steps 10–27)

[1]Correspondence: andrew.camilli@tufts.edu.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Disodium EDTA (Fisher Scientific BP120)

DNA size markers (New England Biolabs N3200)

DNeasy Blood & Tissue Kit (QIAGEN 69504)

Alternatives to the DNeasy Blood & Tissue Kit for isolating bacterial total DNA exist, such as the Quick-DNA Fungal/Bacterial Miniprep Kit (Zymo Research D6005).

Ethyl alcohol (190 proof; Sigma-Aldrich 493511)

Gel loading buffer (10×) <R>

Isopropanol–dry-ice bath (freshly prepared)

Immediately before use, make a small isopropanol–dry-ice bath in a pipette tip box lid by adding enough dry ice to cover the bottom, and then pouring in enough isopropanol to achieve a depth of 1 cm. If no dry ice is available, then an isopropanol −70°C bath can be substituted (see Step 17).

Lysozyme (5%; Millipore Sigma-Aldrich L6876) in AE buffer (QIAGEN)

Warm to 37°C and vortex to dissolve; store at −20°C in 1-mL aliquots. Aliquots are stable for up to 5 yr at–20°C. Aliquots can be freeze–thawed up to four times before becoming inactive.

RNase A (20 µg/mL; Fisher EO0491)

Sea salts powder (Millipore Sigma-Aldrich S9883)

Dissolve 1 g of sea salts in 100 mL H_2O to make a 1% solution. Filter sterilize using a 0.22-µm vacuum filter unit. Store at room temperature; it will be stable for up to 1 yr at room temperature.

SYBR Safe DNA stain (Invitrogen S33102)

TAE <R>

Dilute to 1×.

Reagents required for whole-genome-sequencing procedures only (Steps 28–58)

Gel Green DNA Stain (Fisher Scientific 41004)

Mineral oil (Millipore Sigma-Aldrich M8662)

Nextera XT DNA Library Preparation Kit (Illumina FC-131–1024)

Nextera XT Index Kit v2 Set A (Illumina FC-131-2001)

QIAquick PCR Purification Kit (QIAGEN 28104)

Alternatives to the QIAquick PCR Purification Kit for purifying PCR products exist, such as the DNA Clean & Concentrator Kit (Zymo Research D4029).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Equipment

Agarose gel casting tray and comb

Bacterial cell spreader (Fisher Scientific 08-769-2A)

Block heater for tubes (56°C and 85°C)

CLC Genomics Workbench Software (QIAGEN 832021)

Cryovials (Nunc 377267)

Computer

Freezer at −70°C

Alternatively, a frost-free or non-frost-free −20°C freezer could be used for shorter time periods (see note below Step 9.iii).

Gel electrophoresis chamber (horizontal) and power supply

Glass microscope slides (25-mm × 75-mm; Fisher Scientific 48366–067)

Glass microscope slide coverslips (22-mm × 22-mm; VWR 48366–067)

H_2O bath preset to 56°C

Ice

Illumina NextSeq 550

Image-analysis software (optional; see Step 43)

Incubator at 30°C

Microcentrifuge

Microcentrifuge tubes (1.5-mL; USA Scientific 1615–5500)

PCR tubes (0.2-mL; USA Scientific 1402–4700)

Phase-contrast microscope with 40× objective, 100× immersion oil objective, and immersion oil Spectrophotometers (UV and visible)

Stereomicroscope with a low-magnification objective or a magnifying glass (optional; see Step 7.ii)

Wooden sticks (Fisher Scientific 23400102)

Sterilize wooden sticks by autoclaving them in a plastic or glass beaker with an aluminum foil cover. The sticks can be reused many times by placing used sticks in a labeled beaker and sterilizing again by autoclaving.

Thermocycler

UV or blue light box

Vortexer

Waste container

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

METHOD

Four procedures are described here: isolation of culturable bacteria of the Vibrio genus (Steps 1–9), purification of total DNA (Steps 10–27), whole-genome sequencing (Steps 28–58), and genome analysis (Steps 59–72). The first procedure has been optimized for isolation of bacteria from the Vibrio genus, whereas the rest have been selected for sequencing and analysis of any bacterial species.

Isolation of Culturable Bacteria of the *Vibrio* Genus

Isolate Bacteria

Sample processing usually takes 1 h.

1. Collect several H_2O samples from the ocean or an estuary in microcentrifuge tubes from surface H_2O, bottom H_2O, and H_2O plus a small amount of sediment to enhance the diversity of species collected. Fill each tube. Store for up to 1 wk at room temperature.

2. Concentrate bacteria by centrifugation at 10,000g for 2 min at room temperature in a microcentrifuge. Immediately pour off all but ~100 μL of the supernatant into a waste container.

3. Resuspend the pellet in the residual supernatant (~100 μL); if necessary, add 1% sea salts to bring the volume up to 100 μL. Vortex samples vigorously to disperse biofilms and clumps of bacteria.

4. Make three consecutive 10-fold serial dilutions of each sample in microcentrifuge tubes using 1% sea salts as diluent.

5. Spread 50 μL of undiluted samples and each dilution onto CHROMagar Vibrio plates using a bacterial cell spreader.

   Use sufficiently dry plates such that liquid spread on the surface soaks into the agar. If the liquid does not soak in by 1 min after spreading, then dry the plates with lids off in a biosafety cabinet or, if not available, in a fume hood until the surface liquid is gone.

6. Incubate the plates for 1–2 d at 30°C, or for 2–3 d at room temperature.

   The plates can be stored for up to 1 wk at room temperature.

Purify and Differentiate Bacterial Colonies

Differentiation by microscopy and cryopreservation typically takes 1 h.

7. Colony-purify diverse colony types.

   i. Record the size, color, and morphology of different colony types.

   ii. (Optional) Use a stereomicroscope with a low-magnification objective or a magnifying glass to better differentiate colony types.

   iii. To colony-purify, pick a tiny portion of the colony with a sterile wood stick and streak for single colonies on a fresh CHROMagar Vibrio plate.

iv. Incubate plates as described in Step 6.

8. Further differentiate bacterial isolates by microscopy.

i. Place 10 μL of 1% sea salts on a microscope slide.

ii. Pick a tiny portion of a purified colony and resuspend it in the 10 μL of sea salts. Place a coverslip over the suspension.

iii. Examine the bacteria by phase-contrast microscopy under medium magnification (40× objective), followed by high magnification (100× immersion oil objective with immersion oil).

iv. Note cell size, shape, clumping pattern, motility pattern, presence of vacuoles, and any other distinguishing characteristics.

If the cells appear lysed or have lost their cytoplasmic contents (ghost cells), they are probably not a Vibrio species and were under lethal stress on the CHROMagar Vibrio plate, which is a selective and differential medium for Vibrio species. To increase the chance of sequencing a novel species of Vibrio, proceed with as many different types of bacteria as possible. The methods described below (Steps 9–58) are designed to sequence up to a total of 96 different bacterial isolates, although typically one will obtain six to 12 different colony types using CHROMagar Vibrio plates.

9. Cryopreserve bacterial isolates.

i. Restreak a pure colony from each type of bacteria from Step 8 onto a fresh CHROMagar *Vibrio* plate and incubate as described in Step 6.

This will result in a pure culture plate of each bacterial type. Use these culture plates for cryostorage of the strain (Step 9.ii) and for DNA purification (Step 10).

ii. Using a sterile wood stick, scrape up ~3 mm^3 of cells from the pure plate culture and resuspend the cells in 1 mL of sea salts in a cryovial.

iii. Add glycerol to 15% (vol/vol). Store the cryovial at −70°C.

If a −70°C freezer is not available, the cryovials can be stored in a −20°C freezer. However, the viability of the bacteria could wane after several months. Note that, when storing any biological samples or reagents in a frost-free −20°C freezer, they should be stored in a styrofoam box, or less ideally but still sufficient, a cardboard box, not in contact with the interior walls. This will insulate the samples/reagents from temperature extremes due to the freezer going through periodic cycles of heating the interior walls to eliminate ice buildup.

Purification of Total DNA

The procedure below uses reagents and materials in the DNeasy Blood & Tissue Kit (QIAGEN). All centrifugations are performed at room temperature in a microcentrifuge.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Lyse Cells (2 h)

10. Scrape up ~3 mm^3 of cells from the pure plate culture, being careful not to scrape up the underlying agar. Fully resuspend the cells in 1 mL of 1% sea salts in a microcentrifuge tube.

11. Measure the optical density at 600 nm of the resuspended cells using a spectrophotometer. Move a volume of the cell suspension equal to 1 OD_{600} unit to a new microcentrifuge tube. Bring up the volume to 1 mL with 1% sea salts.

12. Pellet the cells by centrifugation at 15,000g for 2 min in a microcentrifuge. Remove the supernatant using a 1000-μL pipette tip. Note the size of the cell pellet for each isolate.

- If the cell pellet is small, proceed immediately to Step 14, and then Step 16.
- If the cell pellet is large, proceed immediately to Step 13.

  Large fluffy cell pellets are likely a result of large amounts of exopolysaccharide, which can interfere with DNA purification.

  The cell pellets can be stored for up to 2 mo at −20°C if necessary.

13. Suspend the cell pellet in 1 mL of deionized or distilled H_2O. Heat for 10 min at 85°C in a block heater with occasional vortexing.

In some cases, this will dissolve some or all of the exopolysaccharide.

14. Pellet the cells as described in Step 12, and remove the supernatant using a 1000-μL pipette tip.

If the cell pellet got noticeably smaller, then you may have eliminated some or all of the exopolysaccharide.

15. If necessary to remove even more exopolysaccharide, repeat Steps 13 and 14 one or two more times.

If the cell pellet gets smaller after each repetition of this step, then you are likely removing exopolysaccha-rides. Alternatively, some species of bacteria will lyse upon resuspension in deionized or distilled H_2O. You can use microscopy to test this. If you find that the cells are lysing, then carry out Steps 13–15 using 1% sea salts instead of deionized or distilled H_2O.

16. Remove the supernatant from Step 14 or 15 (if not already done) using a 1000-μL pipette tip. Fully resuspend the cell pellet in 40 μL of AE buffer (provided in the QIAGEN kit) plus 10 μL of 5% lysozyme in AE buffer. Incubate for 30 min at 37°C.

AE buffer contains EDTA, which will chelate Mg^{2+} and thus inhibit DNases. This and most other reagents and materials in Steps 17–19, including proteinase K, are part of the DNeasy Blood & Tissue Kit (QIAGEN). RNase A is not included in the kit.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

17. Add 130 μL of ATL solution (provided in the QIAGEN kit) and 20 μL of proteinase K (provided in the QIAGEN kit). Close the lid and mix by inverting the tube eight times, rotating the tube after each inversion such that all sides are wetted. Incubate for 1 h at 56°C in the block heater.

    - If the cell suspension clears within the hour, proceed to Step 18.

      Most Vibrio species and other Gram-negative bacteria will immediately lyse upon addition of ATL buffer, indicated by a clearing of the suspension. Some Gram-negative and most Gram-positive bacteria will take longer to lyse.

    - If, by the end of the 1-h incubation, the cell suspension has not cleared, then subject the tube to two cycles of rapid freeze–thawing using an isopropanol–dry ice bath and a 56°C H_2O bath, incubating for 5 min in each. If no dry ice is available, then an isopropanol −70°C bath can be substituted for the isopropanol–dry ice bath. After the second freeze–thaw cycle, incubate the tube again for 30 min at 56°C to digest proteins, and then proceed to Step 18.

      Immediately before use, make a small isopropanol–dry-ice bath in a pipette tip box lid by adding enough dry ice to cover the bottom, and then pouring in enough isopropanol to achieve a depth of 1 cm. Insert tubes between dry ice pieces to be held upright or held in a floating foam or plastic rack.

Digest RNA and Purify DNA (45 min)

18. Cool the tube from Step 17 to room temperature, and then briefly centrifuge (hold until the microcentrifuge reaches 3000g, and then stop the rotor) to move the contents to the bottom of the tube.

19. Add 2 μL of 20 μg/mL RNase A. Mix by inverting the tube eight times, wetting all sides. Briefly centrifuge to move the contents to the bottom of the tube as described in Step 18. Incubate for 10 min at room temperature.

20. Add 200 μL of AL solution (provided with the QIAGEN kit), close the lid, and vortex vigorously for 3 sec to mix. Turn the tube upside down and vortex again for 3 sec to ensure complete mixing. Briefly centrifuge to move the contents to the bottom of the tube as described in Step 18.

    If there are a lot of white solid clumps within the mixture, pipette up and down repeatedly with a 1000-μL pipette tip to break up the clumps into smaller sized clumps.

21. Add 200 μL of 190 proof ethyl alcohol, and then vortex vigorously for 3 sec to mix. Turn the tube upside down and vortex again for 3 sec to ensure complete mixing. Briefly centrifuge to move the contents to the bottom of the tube as described in Step 18.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

- If there are still large and numerous white clumps within the mixture, pellet the clumps by centrifugation at 2000g for 2 min in a microcentrifuge. Use the supernatant in Step 22.
- If there are clumps, but not large or numerous enough to clog the column below, proceed immediately to Step 22.

22. Use a 1000-μL pipette to move the mixture from Step 21 onto a DNeasy Mini spin column sitting in a 2-mL collection tube (provided with the QIAGEN kit). Centrifuge at 6000g for 1 min. Discard flowthrough and collection tube, and move the spin column to a new 2-mL collection tube.

23. Add 0.5 mL wash buffer AW1 (provided with the QIAGEN kit). Centrifuge at 6000g for 1 min. Discard flowthrough and collection tube, and move the spin column to a new 2-mL collection tube.

24. Add 0.5 mL of wash buffer AW2 (provided with the QIAGEN kit). Centrifuge at 20,000g for 3 min. Discard flowthrough and collection tube. Move the spin column to a microcentrifuge tube.

25. Add 100 μL AE buffer (provided with the QIAGEN kit) onto the center of the spin column. Wait 1 min, and then centrifuge at 6000g for 1 min. Run the flowthrough through the column a second time for greater recovery of DNA.

After elution, keep DNA on ice when in use, and store permanently at −20°C to limit degradation. DNA is stable for up to 1 yr at $-20^{\circ C.}$

26. Determine the concentration of the DNA using a UV spectrophotometer.

27. Confirm the concentration of DNA determined by spectrophotometry and assess the size range of the DNA using agarose gel electrophoresis.

i. Pour a 0.8% agarose gel in a casting tray. Immediately add 0.0001 volume SYBR Safe DNA Stain using a pipette, and mix throughout the molten agarose by stirring with a 1000-μL pipette tip. Place a comb in the gel, and then let the agarose solidify and cool.

ii. ii. Place the gel in a horizontal electrophoresis chamber with running buffer (1× TAE).

iii. Mix 9 μL of DNA with 1 μL of 10× gel loading buffer. Load 2 and 8 μL in separate wells. Load a known amount of DNA ladder in a third well.

iv. Perform electrophoresis at 5 V/cm until the orange dye front reaches the bottom of the gel.

v. Photograph the gel with a blue or UV light transilluminator.

Chromosomal DNA purified using the above procedure should migrate as a diffuse band at 10 kbp in size. The concentration of DNA can be estimated by comparing the intensity of the band with the high-molecular-weight band of the DNA ladder. Plasmids, if present, can migrate at almost any size depending on their size, but usually appear

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

as a sharp band having a different intensity than the chromosomal DNA band. Multiple plasmid bands may be present. The presence of DNases may result in partial degradation of the DNA. As long as the DNA is migrating at a size greater than ~1 kbp, then the molecules are large enough to proceed with the sequencing steps below.

Whole-Genome Sequencing

All centrifugations are performed at room temperature in a microcentrifuge.

Generate a Sequencing-Ready Sample (2 h)—The procedure below uses reagents in the Nextera XT DNA Library Preparation Kit and Nextera XT Index Kit v2 Set A (Illumina). Kits from other manufacturers are available for generating samples suitable for sequencing on the Illumina Massively Parallel Sequencing (MPS) platform, such as the NEBNext Ultra II DNA Library Prep with Sample Purification Beads Kit (New England Biolabs E7103).

28. Thaw the following reagents from the Nextera XT DNA Library Preparation Kit: Tagment DNA Buffer (TD) and Nextera PCR Master Mix (NPM). Thaw index primers 1 and 2 from the Nextera XT Index Kit v2 Set A.

29. After thawing, briefly centrifuge reagents to move liquid to the bottom of tubes as described in Step 18.

    Keep TD, genomic DNA, and index primers on ice, and NPM at room temperature. Reagent Neutralize Tagment Buffer should be kept permanently at room temperature. Reagent Amplicon Tagment Mix should be kept at −20°C until use. Note that other Nextera XT Index kits can be used instead of v2 Set A.

30. Make 10 μL of a 0.5-ng/μL working stock of total DNA using pure, autoclaved deionized or distilled H_2O as diluent and keep it on ice.

31. Add 2.5 μL of TD, 1.25 μL of 0.5-ng/μL total DNA, and 1.25 μL of Amplicon Tagment Mix to a PCR tube.

32. Mix by pipetting or gently vortexing, and then briefly centrifuge to move the contents to the bottom of the tube as described in Step 18.

    Note that the volumes used are one-quarter of what is recommended in the kit manual.

33. Incubate for 5 min at 55°C in a thermocycler with heated lid on.

34. Cool to 10°C either by changing the temperature of the thermocycler or by placing the tube in ice for 30 sec.

35. Stop the tagmentation reaction by adding 1.25 μL of Neutralize Tagment Buffer, mix, briefly centrifuge as described in Step 18, and then incubate for 5 min at room temperature.

36. Add 3.75 μL of NPM.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

37. Add 1.25 µL of index 1 and index 2 primers. Mix, and briefly centrifuge as described in Step 18.

Use of indexed primers allows pooling of samples for sequencing (multiplexing). Each sample to be multiplexed must have a unique index 1 and/or index 2 primer.

38. Add 5 µL of mineral oil.

The mineral oil will form a thin layer across the top of the aqueous mixture and prevent evaporation during the PCR.

39. PCR-amplify using the following parameters:

Cycling step	Temperature	Time	No. of cycles
Initial denaturation	72°C	3 min	1×
Denaturation	95°C	30 sec	12×
Annealing	55°C	30 sec	
Elongation	72 °C	30 sec	
Final elongation	72°C	5 min	1×
Hold	10°C		

40. After the PCR is complete, place the PCR tubes on ice until Steps 41–43 are completed.

Store the PCR tubes for extended periods at −20°C.

41. Estimate the concentration and size range of the PCR products by agarose gel electrophoresis.

i. Pour a 1.5% agarose gel in a casting tray. Immediately add 0.0001 volume Gel Green DNA Stain using a pipette, and mix throughout the molten agarose by stirring with a 1000-µL pipette tip. Place a comb in the gel, and then let the agarose solidify and cool.

ii. ii. Place the gel in a horizontal electrophoresis chamber with running buffer (1× TAE).

iii. Insert a pipette tip until it touches the bottom of the PCR tube containing PCR products. Slowly release the plunger to take up 4.5 µL of PCR products. Save the remaining PCR products on ice.

This helps to avoid taking up mineral oil into the pipette tip.

iv. Mix 4.5 µL of PCR products with 0.5 µL of 10× gel loading buffer. Load the mixture into the agarose gel. Load 5 µL of DNA ladder in another well.

v. Perform electrophoresis at 5 V/cm until the orange dye front reaches the bottom of the gel.

vi. Photograph the gel with a blue or UV light transilluminator.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The PCR products should form a broad smear from ~0.2 to 2 kbp in size. The 900–1200 bp molecules will be sequenced; therefore, it is important that products are visible in this size range.

42. Ensure that the size range is correct.

- If the size range of products is too small, then repeat Steps 28–41 using a working stock of DNA with a concentration >0.5 ng/μL (e.g., 1 or 2 ng/μL).
- If the size range of products is too large, then repeat Steps 28–41 using a working stock of DNA with a concentration <0.5 ng/μL (e.g., 0.1 or 0.2 ng/μL).

  These modifications alter the ratio of DNA to transposomes in the tagmentation reaction, thus shifting the size range of products.

43. Estimate the volume of each sample needed for multiplexing based on visual examination of the agarose gel result to achieve approximately equivalent amounts of 900- to 1200-bp products for each sample.

Although not necessary, one can use image-analysis software such as ImageJ (Schneider et al. 2012) to more precisely calculate the volumes to be multiplexed.

44. Withdraw the aqueous-phase PCR products, not the mineral oil, and place them in a clean 1.5-mL microcentrifuge tube. Place the sample tube on ice.

- If more than one isolate is being sequenced, then pool (multiplex) anywhere from 1 to 6.75 μL of PCR products, achieving an approximately equivalent amount of products in the 900- to 1200-bp range for each sample.
- If only sequencing one isolate, then transfer all 6.75 μL of PCR products to a 1.5-mL microcentrifuge tube.

  Save the remainder of each PCR at −20°C in case more sequence data are needed at a later date. The multiplexed sample can be stored for up to 2 wk at −20°C.

45. Add five volumes buffer PB (provided with the QIAquick PCR Purification Kit) to the multiplexed DNA sample. Vortex vigorously for 3 sec to mix.

There is no need to remove mineral oil before this step.

46. Apply the sample to a QIAquick spin column in a 2-mL collection tube (provided in the QIAGEN kit). Centrifuge at 17,900g for 1 min in a microcentrifuge.

47. Discard flowthrough, and then place the QIAquick column back into the same collection tube.

48. (Optional) Remove any remaining liquid on the rim of the collection tube by touching to a clean paper towel.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

49. Add 0.75 mL of buffer PE (provided in the QIAGEN kit) to the QIAquick column and centrifuge at 17,900g for 1 min.

50. Discard flowthrough and place the QIAquick column back into the same collection tube.

51. (Optional) Remove any remaining liquid on the rim of the collection tube by touching to a clean paper towel.

52. Centrifuge for an additional 1 min at 17,900g.

53. Place the QIAquick column in a clean microcentrifuge tube.

54. In another clean microcentrifuge tube, dilute 30 µL of buffer EB (provided with the QIAGEN kit) in 70 µL of deionized or distilled H_2O to make a 0.3× working stock.

The purpose of diluting buffer EB is to reduce the concentration of the buffer, so that it is less likely to interfere in subsequent pH- or salt-sensitive reactions or manipulations.

55. Add 40 µL of 0.3× buffer EB to the center of the QIAquick membrane. Incubate for 1 min at room temperature.

56. Centrifuge the column at 17,900g for 1 min.

57. Run the flowthrough through the column a second time for greater recovery of DNA.

The DNA can be stored for up to 6 mo at −20°C.

58. Measure the concentration of the purified DNA by UV spectrophotometry.

- If the concentration is <10 ng/µL, then repeat Steps 28–57, but do 16 cycles of PCR in Step 39 instead of 12 cycles.
- If the concentration is in the range of 10–100 ng/µL, proceed to Step 58.

MPS (1–2 wk)

59. Submit the multiplexed sample from Step 56 to a sequencing facility for MPS on the Illumina NextSeq 550 or similar Illumina instrument, for example, the Tufts University Core Facility (www.tucf.com).

i. Request purification of samples in the size range of 900–1200 bp using the Pippin Prep purification system (Sage Science).

If the sequencing facility lacks the Pippin Prep purification system, you may perform gel purification of the pooled sample in the size range of 900- to 1200-bp (see Protocol: **Recovery of DNA from Low-Melting-Temperature Agarose Gels: Organic Extraction** [Green and Sambrook, 2020]). Instruct the sequencing facility to use a high amount

of DNA for cluster generation, as this is necessary when sequencing large-sized DNA molecules on the Illumina MPS platform.

ii. Request mid-output, 150-cycle, paired-end sequencing.

iii. Provide the sequencing facility with the index 1 and index 2 primer sequences used in Step 37.

They will use this information to demultiplex the reads and bin them for each bacterial isolate.

The primer index sequences are listed in the Nextera XT Index Kit v2 Set A manual (Illumina). For example, if one were to multiplex and sequence a total of eight bacterial isolates using the following index 1 and index 2 primers, the following table would need to be provided to the sequencing facility:

Isolate	Index 1 primer	Index	Index 2 primer	Index
1	N701	TAAGGCGA	S/N501	GCGATCTA
2	N702	CGTACTAG	S/N502	ATAGAGAG
3	N703	AGGCAGAA	S/N503	AGAGGATA
4	N704	TCCTGAGC	S/N504	TCTACTCT
5	N705	GGACTCCT	S/N505	CTCCTTAC
6	N706	TAGGCATG	S/N506	TATGCAGT
7	N707	CTCTCTAC	S/N507	TACTCCTT
8	N708	CAGAGAGG	S/N508	AGGCTTAG

Genome Analysis

Genome Assembly (2 h)

60. Open CLC Genomics Workbench Software (QIAGEN) and import the “R1” and “R2” pair of sequence data files (one for each end of the paired-end data) for each bacterial isolate.

These files will have the extension “.fastq.gz”.

61. Import as Illumina data. Under “Paired read information”, choose “Minimum distance = 413” and “Maximum distance = 813”.

The Illumina index primer sequences appended to each end of the DNA molecules total 137 bp. Thus, for PCR products ranging from 900 to 1200 bp, the minimum and maximum distance of genomic DNA between 150- bp paired-end reads is 463 and 763 bp, respectively. However, there may be some error in the Pippin Prep size selection process, on the order of ±50 bp. Therefore, the minimum and maximum distance between paired-end reads can be set to 413 and 813 bp, respectively.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

62. Trim the reads to remove poor quality sequence and Illumina tagmentation and index primer 1 and index primer 2 sequences.

i. Paste Table 1 into Microsoft Excel, save, and then import into CLC Genomics workbench and rename it "Nextera XT adapter trim list".

ii. Highlight all pairs of imported paired-end sequence files and choose the "Trim Reads" function under the "Prepare Sequencing Data" menu.

iii. Click the "batch" checkbox on the first option screen when doing more than one input file as in this case.

iv. Choose the following Trim Reads parameters: (1) trim using quality scores (limit, 0.01), (2) trim ambiguous nucleotides (maximum number of ambiguities, 4), (3) automatic read-through adapter trimming (trim adapter list = Nextera trim adapter list), (4) filter on length (discard reads <20 bp; discard reads >151 bp).

63. Run the "De Novo Assembly" function within CLC Genomics Workbench.

i. Highlight all pairs of trimmed read files and choose the "De Novo Assembly" function under the menu "De Novo Sequencing."

ii. Click the "batch" checkbox on the first option screen.

iii. Choose the following De Novo Assembly function parameters: (1) automatic word size; (2) automatic bubble size; (3) minimum contig length of 300; (4) auto-detect paired distances; (5) perform scaffolding; (6) map reads back to contigs (slow) (mismatch cost, 2; insertion cost, 3; deletion cost, 3; length fraction, 0.5; similarity fraction, 0.95); (7) global alignment; (8) update contigs.

The resulting output is a set of contigs, which together comprise a draft genome of the strain. You should obtain ~20–200 contigs that together comprise ~95%–99% of the genome, depending on the size and complexity of the genome as well as the depth of sequencing. The complexity of the genome refers to the number and size of repeated sequences. Assuming good depth of sequencing (>50-fold coverage), most of the missing genome sequence is composed of repeated sequences longer than ~900 bp (e.g., transposons, insertion sequences, and ribosomal gene operons). However, each repeat sequence should be present at least once in the contigs, typically at the 5′ or 3′ end of contigs. Extrachromosomal elements such as plasmids or excised prophages should form single, complete contigs.

64. Compare the average coverage in the assembly report to calculate the copy number of any extrachromosomal elements.

65. Use the average coverage of contigs comprising repeated sequences such as transposons to calculate the copy number in the genome.

Small contigs with low average copy number may be due to contaminating DNA.

66. Perform sequence homology analysis using BLAST to identify contaminating DNA contigs and remove them from the data set.

Annotate Draft Genome—This typically takes 2 h, aside from registering at the RAST annotation server and RASTtk annotation.

67. Open the de novo assembly contig file in CLC Genomics Workbench.

68. Export the contig sequences in FASTA format.

69. Register at the RAST annotation server (rast.nmpdr.org).

    It may take a few hours or overnight to receive notification of registration.

70. Log in to RAST. Choose “Upload New Job” under the “Your Jobs” menu, and upload your FASTA file.

    Alternatives to RAST exist, such as DFAST (dfast.nig.ac.jp/dfc/).

71. In RAST, leave the string fields for Taxonomy ID and Taxonomy blank. Choose “Domain = Bacteria.” For genus and species, enter “Unknown” and “sp.”, respectively. Leave the strain field blank. Choose “Genetic Code = 11” (archaea, most bacteria, most virii, and some mitochondria). Finally, choose “Use this data and go to step 3” at the bottom.

    i. On the Complete Upload page, choose the RASTtk annotation scheme. Check “Customize RASTtk pipeline”. Next, enable all options in the table. At the bottom of the page, check the first three options: “Automatically fix errors,” “Fix frameshifts,” and “Build metabolic model.” Finally, choose “Finish the upload”.

       Annotation can take several hours. You will receive an email notifying you when your job is finished.

    ii. When your job is complete, log in and go to “status page.” Choose “view details.” Choose “Download” in GenBank format.

    iii. Open the annotation file in CLC Genomics Workbench. Under “Annotation Types” in the sequence list settings to the right, check the RNA box. Next, do a “Find in annotation” for “small ribosomal subunit”. Search until you find the full-length small ribosomal subunit 16S rRNA gene (~1.5 kb in size).

       Bacteria have multiple identical or nearly identical copies of the rRNA genes. However, unless mate-pair sequencing was done, de novo assembly software, including CLC Genomics Workbench, is incapable of placing more than one copy of large repeated sequences such as this into contigs when analyzing short-read MPS data. Therefore, you should find only one complete copy of the 16S rRNA gene in your data.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Taxonomic Classification (4 h)

67. Copy and paste the 16S rRNA gene sequence into NCBI's Targeted Loci Nucleotide BLAST page (blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=TargLociBlast). Choose "16s ribosomal RNA sequences (Bacteria and Archaea)" as the database. Also choose "Taxonomy reports" and "Distance tree of results" under "Other reports:" to see more results.

You may right-click hits and nodes to expand them.

Taxonomic classification using the 16S rRNA gene sequence is capable of classifying bacteria to the family or genus level, but may be unable to discern the species, as some closely related species have identical 16S rRNA sequences, and some strains of the same species have nonidentical 16S rRNA sequences.

68. Perform a deeper taxonomic classification.

i. Upload the five largest contigs to the NCBI Genome Blast page (blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=MicrobialGenomes), and perform the search. Record the top ten hits for each contig and the corresponding genome GenBank accession numbers.

If your five contigs are part of the same genome, then the top hits should overlap substantially.

ii. Perform BLAST analysis on all remaining contigs to make sure that all of them belong to the same species or group of related species.

If some contigs are not homologous to your top hits, then your genomic DNA sample may have been contaminated, your strain may contain one or more plasmids, and/or your strain may contain large genomic islands (e.g., prophages) that are absent from the top hits. Plasmids and prophages can be identified by performing BLASTx analysis of these particular contigs against the non-redundant protein sequence database, and then looking at the bioinformation of the homologous proteins.

iii. Perform in silico DNA–DNA hybridization using the program dDDH (ggdc.dsmz.de/ggdc. php#). Upload all of your contigs as the query genome, except those that you believe are due to contamination, or are plasmids or prophages. As reference accessions, enter the GenBank accession numbers of the top 10–20 hits from your NCBI Genome BLAST search in Step 72.i.

dDDH uses three different algorithms. Hybridization of >70% from any of the algorithms implies that your strain is the same species, whereas hybridization of <70% supports the hypothesis that your strain is a different species.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RECIPES

Gel Loading Buffer (10×)

Reagent	Quantity for 10 mL
Orange G (Fisher Scientific AC416550100)	100 mg
Tris-EDTA solution (10×, pH 7.4; Fisher Scientific BP2477100)	5 mL
Glycerol (80% v/v)	5 mL

1. Add Orange G to 1× Tris-EDTA solution (pH 7.4).
2. Vortex to dissolve.
3. Add 80% glycerol.
4. Vortex to mix.
5. Store at room temperature.

The buffer is stable for up to 5 yr at room temperature.

Molten Agarose

69. Add either 0.8 or 1.5 g of agarose (Fisher Scientific BP160) to 100 mL of 1× TAE <R> electrophoresis running buffer in a 250-mL glass bottle to make a 0.8% or 1.5% solution, respectively.
70. Microwave until boiling starts, and then carefully swirl the bottle.
71. Continue microwaving at reduced power with occasional swirling until the agarose is fully dissolved.
72. Cool to 60°C before use.

    The molten 0.8% agarose can be stored for up to 1 wk at 60°C, but the 1.5% agarose should be used fresh. Single-use aliquots of either can be stored for up to 6 mo at 4°C, and then microwaved to remelt.

TAE

Prepare a 50× stock solution in 1 L of H_2O:

242 g of Tris base

57.1 mL of acetic acid (glacial)

100 mL of 0.5 M EDTA (pH 8.0)

The 1× working solution is 40 mM Tris-acetate/1 mM EDTA.

ACKNOWLEDGMENTS

I thank Cecilia Alejandra Silva Valenzuela, Lauren Shull, Miriam Ramliden, Jacob Bourgeios, and David Lazinski for expert assistance and modifications of these procedures.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

REFERENCES

Green MR, Sambrook J. 2020. Recovery of DNA from low-melting-temperature agarose gels: organic extraction. Cold Spring Harb Protoc doi:10.1101/pdb.prot100461

Meier-Kolthoff JP, Sardà Carbasse J, Peinado-Olarte RL, Göker M. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. Nucleic Acid Res 50: D801–D807. doi:10.1093/nar/gkab902 [PubMed: 34634793]

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Meth 9: 671–675. doi:10.1038/nmeth.2089

blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_-SPEC=TargLociBlast [accessed July 24, 2022].

blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=MicrobialGenomes [accessed July 24, 2022].

dfast.nig.ac.jp/dfc/ [accessed July 24, 2022].

ggdc.dsmz.de/ggdc.php# [accessed July 24, 2022].

rast.nmpdr.org [accessed July 24, 2022].

www.tucf.com [accessed July 24, 2022].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

TABLE 1.

Nextera XT adapter trim list

Name	Sequence	Reads	When an adapter is found	For reads without adapters
P5	AATGATACGGCGACCACCGAGATCTACACNNNNNNNN	All	Trim 3′ end	Keep the read
P7	CAAGCAGAAGACGGCATACGAGATNNNNNNNN	All	Trim 3′ end	Keep the read
Tagmentation read 1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG	All	Trim 3′ end	Keep the read
Tagmentation read 2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG	All	Trim 3′ end	Keep the read