

Protocol for generating double-stranded DNA sequencing libraries for Illumina using NEBNext Ultra II DNA library Prep Kits followed by enrichment of specific targets (eg. antimicrobial resistance genes - AMR baits) using Arbor Biosciences custom myBaits® technology (<https://arborbiosci.com/products/targeted-sequencing-kits/>). This protocol has been tested on clinical genomes from *Escherichia coli*, *Klebsiella pneumoniae*, *Staphylococcus aureus*, *Pseudomonas aeruginosa* and metagenomic DNA isolated from a human stool sample.

This protocol is highlighted in Guiton AK, Raphenya AR, Klunk J, Kuch M, Alcock BP, Surette MG, McArthur AG, Poinar HN, Wright GD. 2019. Capturing the Resistome: A targeted capture method to reveal antibiotic resistance determinants in metagenomes. Antimicrobial Agents and Chemotherapy, in press.

Steps in the workflow:

- Day 1: DNA library preparation (assuming DNA is already prepared and passes quality control)
- Day 2: Library quality control (½ day), begin hybridization (¼ day) - 24 hours
- Day 3: Finish enrichment and library amplification
- Day 4: Library quality control and begin sequencing
- Data Analysis

Materials and Reagents:

- Pure genomic or metagenomic DNA from sample of interest (user supplied) need at least 100 ng to 1 µg of DNA
- Elution buffer (10 mM Tris-HCl, pH 8.0) or 0.1X TE (1 mM Tris-HCl, pH 8.0, 0.1 mM EDTA)
- 10 mM Tris-Cl, 0.05% TWEEN-20 solution (pH 8.0-8.5)
- Nuclease-free molecular biology grade water
- 80% (v/v) Ethanol (freshly prepared with nuclease-free water)
- qPCR reagents (eg. KAPA SYBR Fast qPCR Mix, appropriate primers and appropriate standard - Illumina PhiX control library or other).

Equipment:

- Nanodrop
- Qubit fluorometer (or equivalent)
- Covaris sonicator and MicroTUBEs (or another method of sonication of DNA)

- Low-bind/maximum recovery 1.5 mL tubes, DNase-free
- 200 µL low-bind/maximum recovery tubes, DNase-free
- 50 mL nuclease-free tubes
- Magnetic rack
- Vortex
- Centrifuge, mini centrifuge with adapters for 1.5-1.8 mL and 0.2 mL tubes/strips is
- Recommended:
 - Multichannel pipettor capable of 20 µL and up to 200 µL volume
- Water bath or incubation oven capable of 65°C
- Heat block capable of 65°C
- Thermal cycler
- qPCR machine
- Agilent Bioanalyzer (or equivalent)
- Filter pipette tips

Specific kits/materials required:

- NEBNext Ultra II DNA Library Prep Kit for Illumina with sample purification beads (NEB #E7103 - <https://international.neb.com/products/e7645-nebnext-ultra-ii-dna-library-prep-kit-for-illumina>)
- NEBNext Multiplex Oligos for Illumina (Dual Index Primers, NEB #E7600)
- PCR reagents for post-capture amplification (KAPA HiFi HotStart ReadyMix, Kapa Biosystems) KK2610 - 10 rxns, KK2611 - 50 rxns, KK2612 - 250 rxns
- PCR primers for amplifying sequencing libraries after capture ('reamp' primers described in Meyer and Kircher 2010 (doi: 10.1101:pdb.prot5448) or 10X Library Amplification Primer Mix from Kapa (KK2623))
- PCR purification system (eg. KAPA Pure Beads - KK8000 5mL, or AMPure beads or silica columns - MinElute)
- Using 1 of the kits available from Arbor Biosciences (eg. myBaits Custom, designs with 20-40K probes, 48 rxns #300248) = 5.5 µL per reaction - Use nanogram amounts based on a Nanodrop value or the concentration of the baits provided by the manufacturer. The cost of the custom probes per reaction can be significantly reduced by optimizing the input amount into hybridization.
- Reagents ordered separately: myBaits Reagents (48 rxns) - Arbor biosciences - 300048 (~\$590.00 USD)

Protocol:

A. DNA Library Preparation - NEBNext Ultra II DNA Library prep

<https://international.neb.com/-/media/nebus/files/manuals/manuale7645.pdf>

1. Measure the quantity of DNA via absorbance and fluorescence. Quality measured via absorbance on Nanodrop (Ideal A260/A280 ratio 1.80 and A260/A230 ratio >2.00)
 - a. Ideally would like at least 1000 ng in 50 µL for libraries but a lower input is still feasible to work with (~100 ng).
2. **Optional:** Run sonicated/provided DNA samples on the Bioanalyzer to determine average size of DNA fragments.
3. **Optional:** If DNA is fragmented but of high enough quantity, consider performing a size selection to remove smaller DNA fragments.
4. Libraries are to be prepared in PCR clean hoods that have been UV-irradiated and with the use of bleached/UV-treated equipment.
5. Sonicate 100 ng to 1 µg of DNA to 600 bp using a Covaris MicroTUBE (50 µL total volume). DNA should be diluted in Nuclease-free water or Elution buffer (10 mM Tris-HCl, pH 8.0).
6. Prepare an Illumina Sequencing DNA library using the NEBNext Ultra II library preparation protocol
 - a. **Optional (but highly recommended):** Include a negative/blank DNA extraction alongside library preparations.
7. Follow the protocol as described in the manual with necessary changes such as adapter dilution and appropriate size selection or clean-up.
8. Depending on the input amount into the library, follow the recommended # of cycles for indexing PCR (Table 4.2 in the manual).
 - a. Using NEBNext Multiplex Oligos for Illumina (Dual Index Primers, NEB #E7600)
 - b. Depending on which indexes are used specific Blockers must be used in Part B - Enrichment. These are provided with the myBaits® Reagents. Arbor Biosciences should be made aware of which libraries/indexes are being used so the proper Block A can be sent.
9. Clean-up and elute DNA as described in the manual.
10. Perform library quality control (Bioanalyzer or BA and qPCR)

- a. If exact inputs of adaptor-ligated library fragments going into enrichment should need to be known, the Agilent Bioanalyzer or TapeStation analysis and qPCR are recommended quality control checks.
 - b. If an estimated input amount is needed, the concentration values from the Agilent Bioanalyzer results can be used.
11. Convert pM concentrations into ng/μL concentrations using average size distribution to determine the total nanogram amount of each library available for enrichment.

B. Enrichment with myBaits® protocol V4

- Using 1 of the kits available from Arbor Biosciences (eg. myBaits® Custom, designs with 20-40K probes, 48 rxns #300248)
- Reagents ordered separately: myBaits® Reagents (48 rxns) - 300048 (~\$590.00 USD)
- Protocol taken from the myBaits-Manual-v4 <https://arborbiosci.com/wp-content/uploads/2018/04/myBaits-Manual-v4.pdf>

Considerations for enrichments:

1. With 7 μL for library input, the lowest input tested was 50 ng library from a complex metagenome.
2. The manual states that baits are shipped at 44 μL / 8 rxns = 5.5 μL per reaction however we have been using nanogram amounts of bait rather than a supplied volume.
3. Based on experience with the AMR Baits, enrichments perform well with the following input/ratios (Table 1). Enrichment may perform differently with other sample types and it is recommended that control experiments are performed to assess optimal enrichment conditions. Also, if more library/DNA is available for enrichment the input amounts could be increased - whether the amount of probes need to be increased is not known as trials with 200 ng library and 400 ng baits did not perform better than lower input of baits (200 and 100 ng).

Table 1: Input amounts tested with AMR baits. These ratios have worked best or equally best. Varying the bait ratio did not always cause global improvement of recovery of genes across three replicates. *No significant difference between the two input bait amounts.

Sample type	Input library (ng)	Input baits (ng)
Pure bacterial genome	100	100
Stool metagenome	200	100/200*
Stool metagenome	100	50/100*
Stool metagenome	50	25/50*

4. With the AMR baits only the following hybridization conditions have been tested: used 65°C for 16 or 24 hours. Other temperatures and times may work as well.
5. Other important considerations are described in the myBaits® manual V4.
6. Only 1 round of enrichment has been tested as the results after this were promising and additional rounds of enrichment were not required. In some cases, 2 rounds may be beneficial.

Enrichment is performed following the myBaits® Manual v4.0 with the following considerations and changes:

Part 1: Hybridization Mix Setup considerations

1. Most mixes/reagents are prepared in excess. Added below are considerations for diluting the baits in order to achieve the desired final amount in the hybridization reaction.
2. This calculation is done by using the final amount (ie. 100 ng) and accounting for the excess volumes in each step prior to step 1.4 in the manual.
 - a. Baits are diluted to ~20.2 ng/μL in a volume that is suitable for the number of reactions to be processed.
 - b. 5.5 μL of this is used in the Hybridization Mix Setup (step 1.2) in 20 μL = 111.1 ng
 - c. 18 μL of the hybridization mix is transferred to the library mix in step 1.4 = ~100 ng in final step
3. Blocker mix - ensure you have the required amount of library DNA in 7 μL

Part 2: Bind and Wash ("Cleanup")

1. Start by ensuring the water bath is turned on and heating to 65°C

2. Cleanup can be performed in 1.7 mL or 0.2 mL tubes
3. Only make enough Wash Buffer X for the reactions being processed at one time.
4. 2.3 - Bead preparation
 - a. Recommended to prepare beads in 7 (or fewer) reactions in 1.7 mL tubes = 210 μ L of beads (ensure they are vortexed completely), pellet then wash 3 times in 1.4 mL of Binding Buffer
 - b. Consider preparing one reaction in excess to account for loss throughout washing.
 - c. At the final step, resuspend in 490 μ L then aliquot 70 μ L to individual low-bind tubes.

Part 3: Library Resuspension and Amplification

1. With the KAPA HiFi Hotstart polymerase, the enriched library can remain “on-bead” for amplification. The set to incubate at 95°C for 5 minutes and use off-bead libraries for amplification has not been tested.
2. For amplification - use the 10X Library Amplification Primer Mix from KAPA (KK2623) - 2.5 μ L with 7.5 μ L NF water and 25 μ L 2X KAPA HiFi HotStart Ready Mix and 15 μ L enriched library.
3. For library amplification cycles, follow the program from the myBaits® Manual V4 with step 4 (72°C for 1 min) and steps 2 - 4 repeated for 14 cycles. Depending on the input amount, the number of cycles can be reduced. With the stool metagenomes and clinical strains, the starting material ranging from 50 - 200 ng so 14 cycles were used.
4. After amplification, perform a PCR clean-up (0.8X to 1.0X beads). Pellet the beads that have carried over in the PCR reaction before moving the supernatant to a tube containing the purification beads. The protocol suggests this is only necessary for purification with silica columns, but this has not been tested with the purification beads. Elute in 35 μ L (or less) 10 mM Tris-HCl, pH 8.0 and store at 4°C.

C. Post-enrichment quantification and sequencing

1. With enriched, amplified, and purified libraries it is sometimes difficult to get informative results from the Nanodrop - values are usually <5 ng/ μ L, therefore the Nanodrop is not accurate.

2. The Qubit usually provides values less than 1 ng/μL and this should be sufficient for sequencing.
3. Perform qPCR and use the pre-enrichment average sizes to calculate/estimate the concentration of enriched libraries. If the concentration is high enough based on Qubit, the Bioanalyzer can be used to provide more details on quality, size, and concentration prior to qPCR, pooling and sequencing.
4. Prior to sequencing but after qPCR, the samples are pooled equimolar and sent for QC - qPCR and Bioanalyzer as a pooled sample.
5. Depending on the average insert size of the libraries and available sequencing chemistry – 2 x 250 bp reads are preferred. If starting DNA is on-average shorter, then shorter reads would be more appropriate.
6. Depending on the sample/application, usually a sequencing depth of between 100,000 to 300,000 clusters per enriched library is sufficient. Though deeper sequencing has not been attempted and is predicted to provide additional information (based on rarefaction).

D. Enriched sequencing analysis

- There are various approaches to analyzing metagenomic sequencing data, let alone targeted sequencing data. Suggested approaches are described in Guiton *et al.* 2019.
- RGI also offers a module for analyzing metagenomic reads (<https://github.com/arpcard/rqi>).
- Reads should be trimmed, filtered, and duplicate removal should be considered.
- Reads can potentially be merged prior to mapping or assembly.
- In general, reads are mapped to the reference of the probes as well as the reference that the probes were designed from.
- Various normalization approaches should be considered – based on total reads obtained or through subsampling.
- If blanks/negative controls are carried through all experiments and sequenced, their contents should be reported and considered in the final analysis and reporting.