**Introduction**

DNA barcoding is a molecular approach to species identification that uses a short region of DNA, such as Cytochrome Oxidase I (COI), to differentiate species. Next-generation sequencing (NGS) is a high-throughput technology that can be used with DNA barcoding to identify species within bulk insect samples. This process is termed ‘metabarcoding’ and can significantly upscale vector surveillance. Metabarcoding is performed by extracting nucleic acid (Step 1), amplifying the COI region (Step 2), adding illumina adapters (Step 3), Normalizing and pooling libraries (Step 4), conducting quality checks (Step 5) and Sequencing (Step 6).

1. **Non-destructive DNA extraction protocol for single/multiple insects.**
2. Heat heating block at 56°C.
3. Put single insect/multiple insects in a 1.5mL/2.0mL Eppendorf tube (possibly with secure cap, to avoid it popping open during the night).
4. For single insects: add 180μL of buffer ATL and 20μL of proteinase K.

For multiple insects: add the same ratio ATL/protK in an amount that can cover the insects (e.g. 360μLATL/40μLprotK)

1. Incubate overnight at 56°C.
2. Quick spin to remove liquid from the lid.
3. Add 200μL (or same volume of ATL/protK) of buffer AL and 200μL (or same volume of ATL/protK) of 100% ethanol to each sample.
4. Vortex
5. Transfer 600μL of liquid into a QIAGEN kit column (filter inserted in the collection tube) and centrifuge at 6000g for 1 minute. Discard lower collection tube. Repeat with another 600μL if you are working with multiple insects.
6. Insert filter column in a new collection tube.
7. Add 500μL of Buffer AW1 and centrifuge at 6000g for 1 minute. Discard the collection tube.
8. Add 500μL of Buffer AW2 and centrifuge at maximum speed (usually 13000) for 3 minutes. Discard collection tube.
9. Transfer filter column in a new 1.5mL Eppendorf tube. Add 100μL of AE buffer on the filter (without touching it with the tip!).
10. Incubate at room temperature for 1 minute.
11. Centrifuge at 6000g for 1 minute.
12. Repeat a second elution if you want to obtain more DNA (consider if adding the second elution in the same Eppendorf or keep them separate).
13. Discard filter column and retain the Eppendorf with the DNA.
14. Store DNA at -20°C

**This protocol uses the DNEasy Blood and Tissue kit protocol slightly modified from Bahder *et al.* 2015. The “single insect version” is published in Martoni *et al.* 2019.**

1. **Initial PCR to amplify COI barcode**

An initial PCR is performed to create the COI amplicons for sequencing. We are using the fwhF2-fwhR2n primers of Vamos et al 2017, modified with partial Illumina sequencing adapter tails, underlined in the diagram below

5'– ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGDACWGGWTGAACWGTWTAYCCHCC-3'

5’– GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTRATWGCHCCDGCTARWACWGG-3'

The optimal number of PCR cycles used during this step can change depending on community size, taxonomic composition, and amount of DNA from the extraction step. As a rule, if the amplicon is run on a gel and a band is visible, it is suitable for library prep. The protocol below suggests 30 cycles, which has been shown to be reliable for successful library prep, however if strong bands are present, they may require dilution prior to indexing PCR

Note: it is important to use high fidelity polymerase, such the MyFi DNA polymerase listed in this protocol to avoid the formation of primer dimers caused by long the adapter-tailed primers.

Equipment/materials needed:

* Thermal cycler
* Vortexer
* Repeater pipette
* 5 mL microcentrifuge tubes
* 96 well PCR plates
* 1X BSA
* MyFi DNA Polymerase
* 5X MyFi Reaction Buffer
* 10 µM forward primer with adapter tail:
* 10 µM reverse primer with adapter tail:

|  |  |  |
| --- | --- | --- |
| **Reagent** | **x 1 reaction (µl)** | **x 100**  **reactions (µl)** |
| 1X BSA | 14.7 | 1470 |
| 5X MyFi Reaction Buffer | 5 | 500 |
| Forward primer | 1 | 100 |
| Reverse primer | 1 | 100 |
| MyFiDNA Polymerase | 0.8 | 80 |
| DNA template | 2.5 | 2.5 |
| **Total volume** | **25** | **4500** |

94˚C for 2 min

94˚C for 30 sec

49˚C for 45 sec x30 cycles

72˚C for 45 sec

72˚C for 1 min

15˚C indefinitely

1. Once the PCR is finished, store the plate at 4c

**1.2 Gel electrophoresis**

Running the amplicons on a gel is a good quality control step to ensure the PCR has worked. It is particularly important if you have lowered the number of PCR cycles used in Step 1, because amplicons not visible on a gel can sometimes lead to failed libraries.

Equipment/materials needed:

* Gel loading plate, dock and tank
* Plastic beaker
* Agarose
* TE buffer
* SYBR safe
* 6X DNA gel loading dye
* 1 Kb DNA ladder

Method:

1. Add the appropriate number of well combs to the gel plate, secure it and ensure it is evenly placed in the dock by using a spirit level.
2. Weigh out 2 g of agarose in the plastic beaker.
3. Add 100 mL of TE buffer and swirl to mix.
4. Microwave until the mixture is boiling and there are no visible small bubbles.
5. Add 10 µl of SYBR safe to the agarose and pour it into the plate. Ensure there are no bubbles or dust in the gel.
6. Allow the gel to dry for at least 30 min, then place into a tank with the wells closest to the black electrodes.
7. Mix 5 µl of DNA ladder with 2 µl of gel loading dye and add to the first well.
8. Mix 5 µl of PCR product with 2 µl of gel loading dye and add it to another well.
9. Run the gel at 100 V for approximately 40 min, or until the bands are close to the bottom.
10. Drain the plate, place in gel dock and photograph the gel. All samples that have a visible band can be sequenced, those that do not should be re-amplified with more PCR cycles or DNA template used (with proportionally adjusted BSA volume).

**Optional step: DNA Dilution 1/10**

A 1:10 DNA dilution may be required following the initial PCR, depending on the concentration of the amplicons

1. **Second (Indexing) PCR**

A second PCR is performed to add the remainder of the illumina sequencing adapters, as well as the indexes. The [i5] and [i7] component represent a variable 8bp index sequence, while the underlined part anneals onto the primer tails from the previous step

3’-A\*ATGATACGGCGACCACCGAGATCTACAC[i5]ACACTCTTTCCCTACACGAC-5’

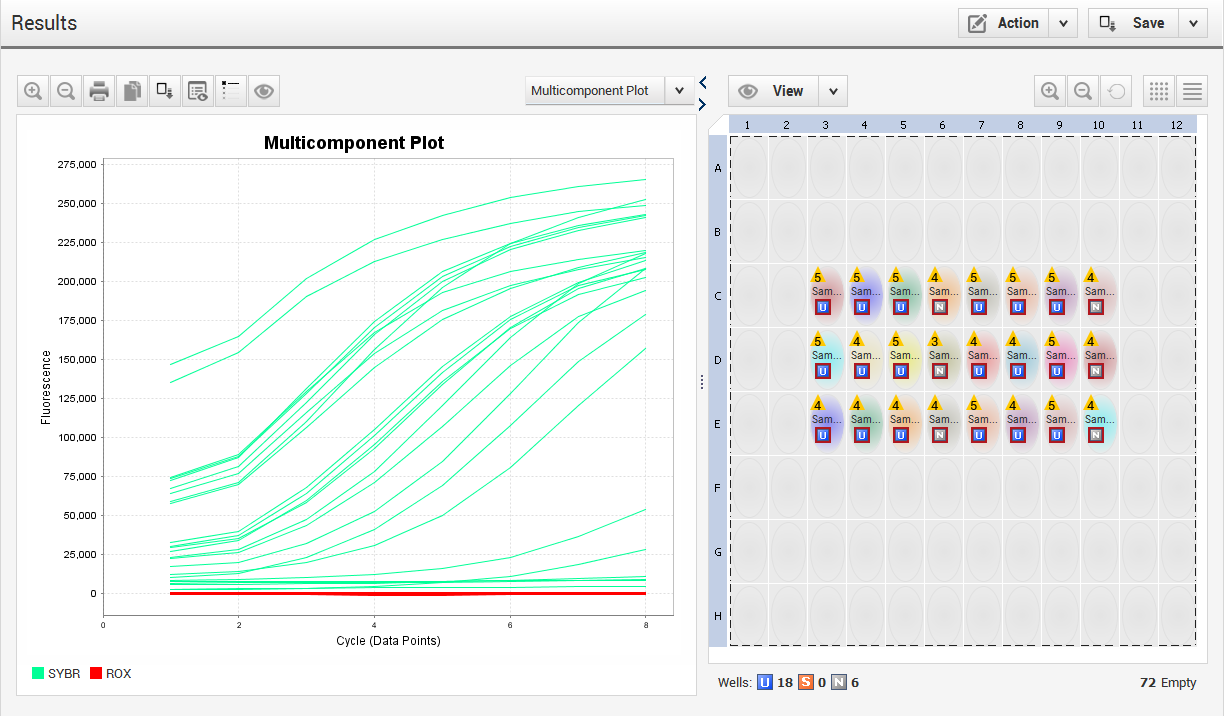
3’-C\*AAGCAGAAGACGGCATACGAGA[i7]GTGACTGGAGTTCAGACGTG-5’

Equipment/materials needed:

* 10mM dNTP Mix
* 10X SYBR Green I Mix
* 5X Phusion HF Buffer
* Phusion DNA polymerase
* 2.5mM barcode primer mix
* Molecular grade water
* Real-time thermal cycler
* Vortexer
* Plate spinner
* Plate sealing film
* 96-well plate (if using the Bio-Rad CFX qPCR machine in L2Q1, the special Bio-Rad qPCR plate is required)
  + Dilute PCR product as recommended by dilution table
  + Prepare PCR master mix on ice. Briefly vortex and spin down reagents before use. Ensure optimal mixing by pipetting up and down.

|  |  |  |
| --- | --- | --- |
| **Reagent** | **x 1 reaction (µl)** | **x 100 reactions (µl)** |
| Water | 15.75 | 1575 |
| 5X Phusion HF Buffer | 5 | 500 |
| 10mM dNTP Mix | 0.5 | 50 |
| 1/1000 SYBR Green I Mix\* | 0.5 | 50 |
| Phusion DNA polymerase | 0.25 | 25 |
| **Total volume** | **23** | **2300** |

* + Add 45 µl of the master mix into each of the qPCR plate wells.
  + Add 2 µL barcode primer mix (2.5 µM) to each well from the 96 well barcode primer plates.
  + Add 1ul of the diluted PCR product from the initial PCR into the qPCR plate
  + Seal tightly with BioRad tape strip, **be careful not to leave marks on top of plastic seal tape strip**
  + Vortex to mix, then pulse centrifuge. Tap plate on desk to remove bubbles then vortex again.
  + Amplify on qPCR machine using the following profile
    - 1. 98˚C for 30 sec
      2. 98˚C for 10 sec
      3. 65˚C for 30 sec x30 cycles
      4. 72˚C for 30 sec + plate read
  + Create a new “plate”, mark samples as Unknown, mark blank as NTC and select Fluorophore used (SYBR in this case)
  + Stop amplification in the **72c extension phase** when leading plots reach 90% CT or when the recommended cycle number is reached– To identify this:
    - Once the run has started, select “Raw Data Plots” tab
    - Select, all, right click anywhere inside the plate and ensure “Consolidate plots enabled” is ticked. Then, right click again and select show plots. This will allow you to enlarge the plots and visualise them all simultaneously. Y-axis display range can be adjusted on the right for ease of viewing.
    - 90% CT is identified at the first cycle after the exponential phase has started where dR increases by less than the previous cycle.
    - Important: If there is no amplification on the screen, the reaction may still be progressing but the SYBR is not imaging- This could be due to bubbles. **If the progress of the reaction is not visible, stop the PCR around 5 cycles**



**4: Normalisation and pooling**

Equipment/materials needed:

* SequalPrep™ Normalization Plate Kit, 96-well (ThermoFisher)
* 96-well plates

**SequalPrep normalization Protocol**

Binding step:

1. Transfer PCR product (5-25 ul; at least 250 ng amplicon/well) from PCR plate to SequalPrep Normalization plate
2. Add equivalent volume of Binding Buffer (25 ul).
3. Seal plate and mix by vortexing, followed by a brief spin down.
4. Incubate at room temp for 1 hour.

Washing step:

1. Aspirate liquid from wells. **do not scrape the side of wells**
2. OPTIONAL: store amplicon/binding buffer liquid (<30 days)
3. Add 50 ul Wash Buffer to wells. Mix by pipetting gently up and down twice. Completely aspirate the buffer and discard. **do not scrape the side of wells**

Elution step:

1. Add 15 ul Elution buffer to each well.
2. Seal plate and mix by vortexing again.
3. Incubate at room temp for 5 min.
4. Transfer to clean tube and pool all purified DNA.
5. Store at 4 C for short-term storage, or –20 C for long term.

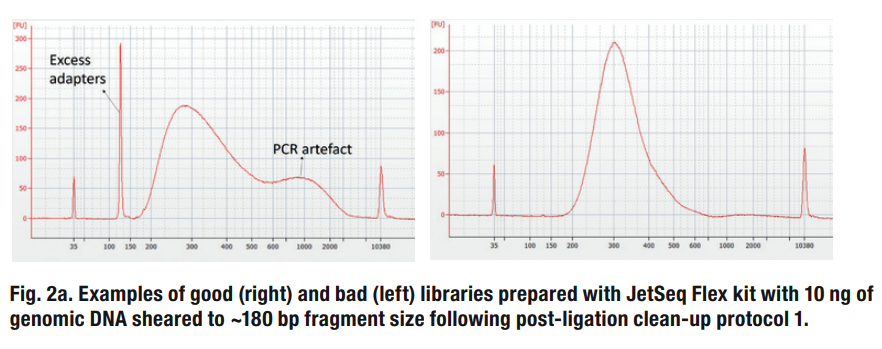
Expect 1-2 ng/ul per well or ~25 ng/well. I pool immediately and both times the pool has been 1nM.

**5: Library quality check**

Equipment/materials needed:

* D1000 screentape
* D1000 screentape reagents

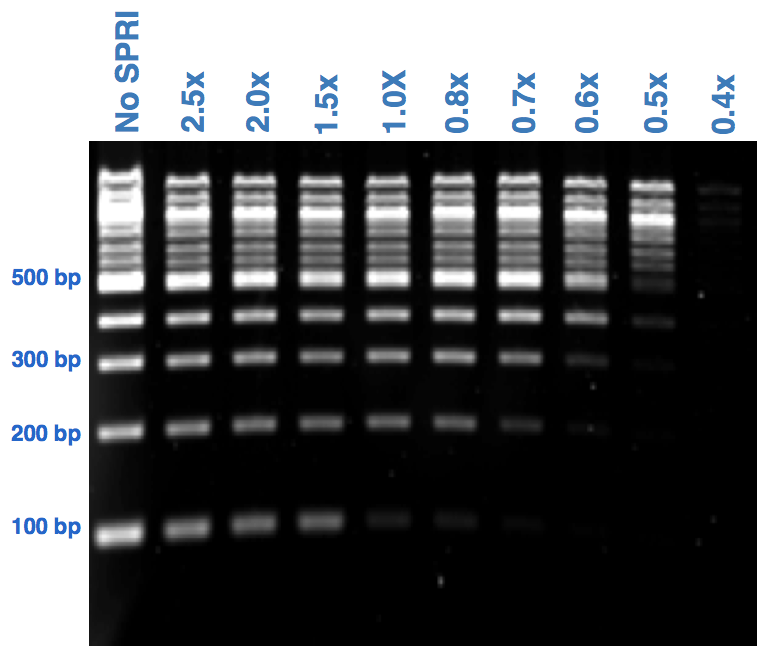
Run a D1000 tape to quantify and measure the fragment size of the final pooled library, and also check for presence of adapter dimers (fig a), if there are any peaks smaller than the desired product, these will need to be size selected out with a 0.8x or 0.7x cleanup as they will be preferentially sequenced. If there are no smaller peaks, a 1.0x cleanup can be conducted to purify the DNA from leftover PCR reagents



**Optional: Additional library cleanup**

Equipment/materials needed:

* Ampure XP beads
* Fresh 70% ethanol
* Bead magnet for 1.5mL Eppendorf tubes
* 10mM tris-HCl pH 8.0 (or ddH20)
  1. Let the AMPure beads come to room temperature for at least 30 minutes prior to using
  2. Mix the AMPure beads well by vortexing so they appear homogenous and consistent in colour
  3. Add the appropriate amount AMPure beads to the pooled library sample, then vortex to mix and spin briefly to collect the liquid without pelleting the beads
  4. Incubate at room temperature for 5-10 mins
  5. Put the tube on a magnetic stand at room temperature and wait for the solution to clear (approx. 2-5 mins)
  6. While keeping on the magnet, carefully remove and discard the supernatant from each sample, do not touch the beads while removing the solution
  7. Continue to leave the plate or tube on the magnetic stand and add enough 70% ethanol to cover the beads
  8. Wait for 1 min to allow any disturbed beads to settle and then remove the ethanol and discard
  9. Repeat steps 7 and 8 once for a total of 2 washes
  10. Briefly spin down and return to the magnet. Use 10ul tips to make sure all ethanol has been removed
  11. While on the magnet, air dry the samples for 2-5 mins, making sure not to over-dry – when the beads are dry they should lose their sheen and become a matte colour
  12. Add 52ul 10mM tris-HCl pH 8.0 (or ddH20) to the bead pellet (reduce volume if library concentration is low)
  13. Seal plate or cap tubes, then vortex and briefly spin to collect the liquid
  14. Incubate for 2 mins at room temperature
  15. Return plate or tube to the magnetic stand and leave for 2 mins or until solution clears
  16. Remove 50ul clear supernatant and transfer to a new tube and discard the beads.
  17. **Run a D1000 tape**
  18. **If adapters or primer dimers are still present in sample, run the library through an additional 0.8x or 0.7x bead clean up and elute in 32ul**



**6: MiSeq Sequencing (Running the machine)**

The initial part of this step is based on what is found in the Illumina ‘Preparing Libraries for Sequencing on the MiSeq’ document. Prepare a sample sheet for the MiSeq before starting this step. Email the completed sample sheet to an email address that can be accessed from the MiSeq (gmail, hotmail etc. **Not** a work email address).

Equipment/materials needed:

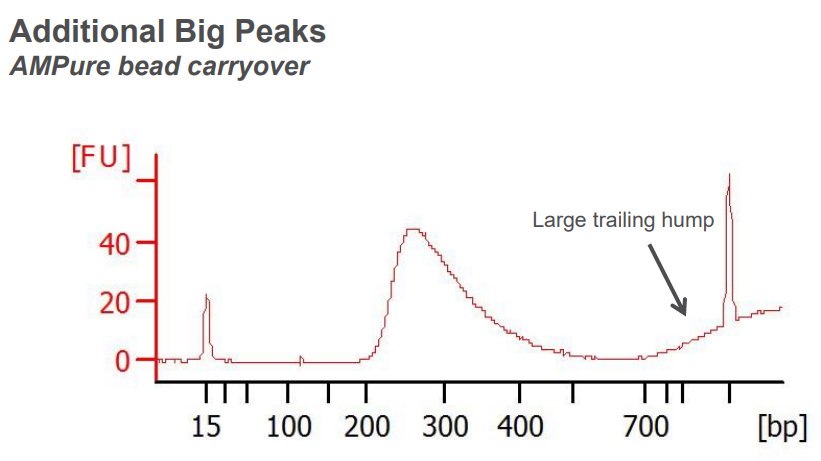
* Molecular grade water
* Freshly diluted 0.2 N NaOH – **Prepare at least 1ml to prevent pipetting errors** (20uL 10N NaOH, 980uL H20)
* MiSeq Reagent Kit (v3, 600 cycles)
  + The cartridge and Ht1 are stored in the freezer – defrost overnight before the run (if not then the cartridge can be defrosted in warm water for 1-2 hours)
  + The flow cell and PR2 bottle are stored in the fridge
* 20pM PhiX control (in sequencing room)
* MilliQ water
* Tween (10% and 0.5%, in sequencing room)
* Bleach (in sequencing room)
* Microcentrifuge
* Vortexer
* 1.5 mL microcentrifuge tubes

**Method:**

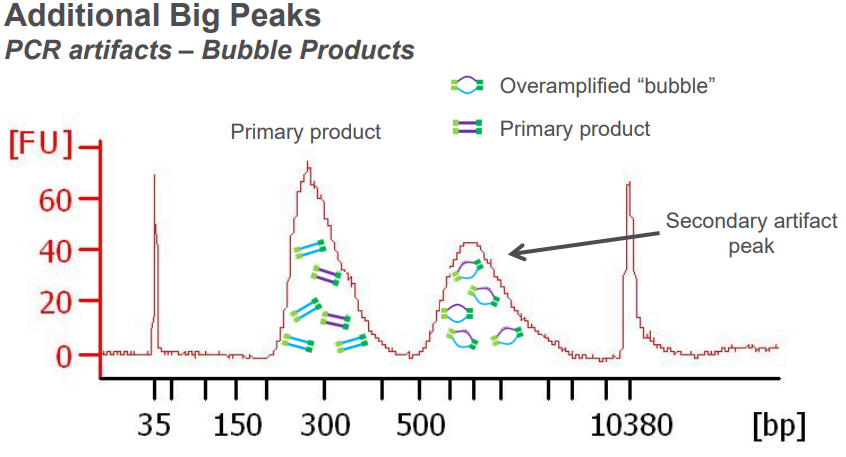
1. Use ng/ul to nM calculator to calculate molarity and dilute down to 2nM – This is an appropriate concentration for MiSeq sequencing
2. Transfer 5 µl of the 2nM library to a clean 1.5 mL tube.
3. Add 5 µl of freshly diluted NaOH, vortex and spin down.
4. Incubate for 5 min at room temperature to denature the DNA.
5. Add 990 µl of chilled HT1 – this dilutes the library to 10pM.
6. Transfer 420 µl of the 10pM library to a clean 1.5 mL tube.
7. Add 180 µl of chilled HT1 – this dilutes the library to 7pM.
8. Invert several times and spin down.
9. Transfer 570 µl of the 7pM library to a clean 1.5 mL tube.
10. Add 30 µl of 20pM PhiX, vortex, spin down and place on ice.
11. **Optional: additional heat denaturation**
    * Incubate the diluted library at 96c for 2 minutes using a heat block
    * After heat incubation, invert the tube 1-2x to mix
    * Quickly move library to an ice water bath for 5 minutes
    * Proceed immediately to cluster generation

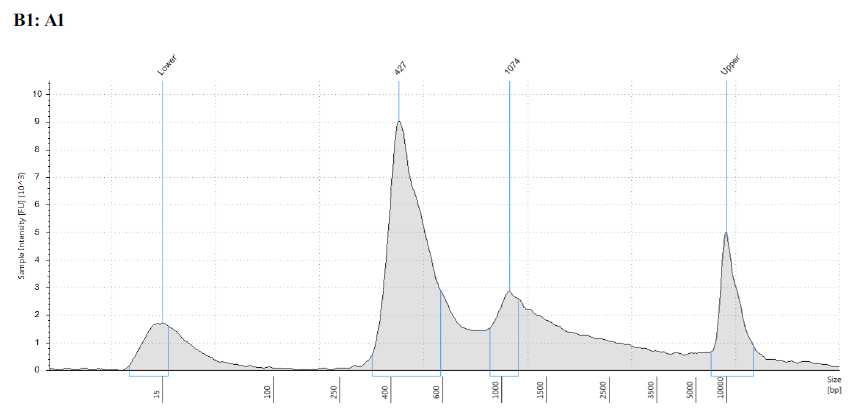
1. Check that the MiSeq booked for this run has been washed and is ready to run. Start the MiSeq process and follow the instructions on the screen. If necessary, login to BaseSpace.
2. Download the sample sheet from the email account it was sent to, and then change the name of it so it matches the number of the cartridge that will be used for the run. Check the read length is correct.
3. Write down the flow cell number, and then wash the flow cell with water. Ensure no fibres or dust are present on the flow cell. If necessary, use a Kimwipe with ethanol to get rid of any fibre or dust. Ensure the ethanol does not touch the black line on the flow cell. Insert the flow cell into the MiSeq.
4. Open the door, lift the lever and empty the old wash bottle. Put the new PC2 buffer inside and lower the lever. Make sure the waste bottle is empty.
5. Invert the cartridge 10 times to ensure there are no ice particles. Tap the cartridge on a benchtop to ensure there are no air bubbles. Take a 1000 µl pipette tip and pierce the foil covering the sample hole. Add the 600 µl of library.
6. Remove the dummy cartridge from the MiSeq and insert the loaded cartridge. Close the door.
7. Perform the check and start the run. Record the run in the MiSeq log book.
8. After the run is finished:
   * Empty the waste bottle into the cytotoxic liquid waste bin.
   * Put the old cartridge into the purple bin.
   * Tip the leftover wash buffer down the sink and throw the bottle into a yellow bin.
   * Add 500 mL of MilliQ and 25 mL of 10% tween to the MiSeq’s wash bottle, and put it into the MiSeq.
   * Squirt 0.5% tween into the dummy cartridge. Add a tube of bleach into position I7. Place the cartridge into the MiSeq.
   * Run a post-run wash.
   * Transfer over the read files to BASC if immediate data analysis is required.
   * Record the run stats in the MiSeq log book.

**Troubleshooting**



Ampure beads were carried over from cleanup step. Place library back on magnet for 5 minutes then re extract library. It is recommended to take off 1 or 2 ul less solution that what was put in to avoid bead carryover





This can be caused by too much input DNA and too many cycles. The bubble products have apparent high molecular weight, as the ends are duplex while the middle comprises two non-complementary strands. These bubble products can appear as daisy chains of sizes double, or triple the desired amplicon if they are very entangled products. To diagnose bubble products, re-heat samples to 95c for 5 minutes, and allow to cool slowly to room temperature in heatblock. Re-run on tape station and look for a size shift in product. These libraries may still sequence fine, however quantification will be challenging. Quantify these libraries extra carefully using KAPA qPCR. Alternatively, index PCR should be repeated with less cycles or diluted DNA input.