Genomics Training Academy (GTAC)

NHS England

Virtual reality resource supporting material: Massively parallel sequencing

Model answers document

V2 16/10/24

About this document

This document has been designed to support the massively parallel sequencing virtual reality resource in training sessions. The questions and model answers below provide a reference point for discussions that will have happened during the session, and can be kept by learners for future reference, or used by trainers to aid in session planning.

Model answers

Bookmark 1: Samples and reagents (05:12)

1. **Some toxic reagents are used in this procedure – how do laboratories assess the safety of these and control their use?**

All reagents will have a COSHH (Control of Substances Hazardous to Health) risk assessment in place. A COSHH risk assessment identifies toxic reagents and defines the safest possible approach to their use and disposal.

2. **Why is it important to accurately measure DNA concentration at the start of this procedure?**

Variations in DNA concentrations would lead to inconsistent library preparation between samples. By measuring the DNA concentration of each sample we can then bring each sample to a standard concentration – this is called normalisation.

3. **What could be the consequence of an error during one of the sample transfers?**

Ultimately, the reporting of an incorrect result for a patient, which could significantly affect patient management. Therefore, a witness check (i.e. a colleague witnessing each transfer and double-checking labels) is necessary at each step in the procedure where a patient sample is moved across to a new location.

4. **What is meant by a safe stopping point? Why is it important to have these in the SOP?**

This is a point where the staff member is able to take a break from the procedure for a specified amount of time. Safe stopping points allow work to be divided into manageable stages and allow some flexibility for staff undertaking the procedure. Some reagents are very sensitive to exposure to certain conditions, so safe stopping points are worked around these crucial steps.

Bookmark 2: Fragmentation, end-repair and A-tailing, and adaptor ligation (12:47)

5. **What is the purpose of the fragmentation step and how is this achieved?**

To create smaller DNA fragment sizes appropriate for downstream sequencing. This is most commonly done using enzymes but can also be done using other methods such as mechanically (for example, by using high-frequency sound waves) or physically (for example, shearing by force).

6. **What is the purpose of using the microplate shaker or vortex and centrifuge?**

The microplate shaker and vortex are used to ensure rapid and efficient mixing. Centrifuging returns the liquid to the bottom of the tube, providing better access for pipetting. Centrifuging the sample also helps to prevent contamination through the release of aerosols and splash contamination from the tube when opened.

7. **What do the adaptors do?**

They contain sequences that allow the library to bind to barcoding indexes and the flow cell for sequencing.

8. **How do we isolate and purify the DNA fragments?**

We use magnetic beads that bind the DNA to them only under certain conditions, for example, the introduction of a particular binding buffer. By placing the tube in a magnetic field, the beads with the DNA bound to them are kept at the bottom of the tube, enabling easy removal of the liquid containing any impurities without removing the DNA. DNA can be washed while attached to the magnetic beads and then released from the beads and eluted back into a liquid. The purified DNA in the liquid can then be transferred to a new tube while the magnetic beads remain behind in the magnetic field.

Bookmark 3: Sample indexing and library amplification (17:52)

9. **What do the indexes do?**

Sample indexes are short, unique DNA sequences added to each DNA sample during library preparation. They allow multiple DNA samples to be pooled together and sequenced in a single run.

10. **What happens at each stage of the PCR reaction? What temperature is required for each?**

- Denaturation (94°C–98°C): To separate the double-stranded DNA into single strands. This high temperature breaks the hydrogen bonds between the DNA strands, providing single-stranded templates for the next step.

- Annealing (50°C–65°C): To allow primers to bind (anneal) to their complementary sequences on the single-stranded DNA. The temperature is optimised to ensure specific binding of the primers to the target sequence without non-specific binding.
- Extension/elongation (72°C): To enable the DNA polymerase enzyme to synthesize a new DNA strand by adding nucleotides to the annealed primers. The temperature is optimal for the activity of Taq polymerase, a commonly used enzyme in PCR.

11. **Thermocyclers are used a lot in this procedure. How do we ensure that the temperatures they hold are accurate?**

Regular calibration and temperature validation using control samples.

12. **What is the purpose of adding ethanol?**

Ethanol precipitation is used to purify DNA by removing salts, enzymes and other contaminants that may be present after various steps in library preparation, such as after fragmentation or adapter ligation. The process helps ensure that only high-quality DNA is included in the final library.

Bookmark 4: Library pooling and target enrichment (23:05)

13. **Why are we able to pool the samples at this stage without risk of cross-contamination?**

Each fragment has had a sample index added to it to ensure it can be traced back to the correct patient during the demultiplexing stage of the bioinformatics pipeline.

14. **What method of target enrichment is being used here and why is this done? Can you think of another method that is used by other NGS workflows?**

Hybrid capture – this ensures that only the regions of interest will be sequenced. Amplicon-based is another method which uses region-specific primers only.

15. **Why do the libraries of lower concentration need to be added separately?**

Massively parallel sequencing platforms, such as Illumina, rely on a certain concentration of DNA to achieve optimal cluster density on the sequencing flow cell. Libraries that are too concentrated can lead to overlapping clusters, which can cause problems with data quality and result in reduced accuracy. Conversely, libraries that are too dilute may not produce enough clusters to generate sufficient data. Separate handling of lower concentration libraries ensures that each library is prepared to an optimal concentration for sequencing.

Bookmark 5: Bead washes, amplifying and purifying libraries (28:07)

16. **Why do the different buffers need to be held at different temperatures?**

Some buffers contain components that are sensitive to temperature. For example, salts and cofactors may precipitate or degrade if the buffer is not kept at the appropriate temperature. Proper temperature control ensures the stability of these components and the overall effectiveness of the buffer.

17. **What are the benefits of using a multi-channel pipette in this procedure?**

Increased throughput and consistency, and also a reduction in the number of transfers which reduces risks of pipetting error or contamination. Fewer pipetting actions for staff reduces risk of RSI.

Bookmark 6: Library quantification, denaturation and sequencing (33:08)

18. **What instrument is used to measure the fragment sizes prior to sequencing and how does this work?**

An electrophoresis instrument (in this case, a TapeStation) which measures fragment sizes by electrophoresis.

19. **What also needs to be checked before loading the sample on to the sequencer? Why is this important?**

The molarity. This check is important to ensure optimal cluster density – too high would result in overcrowded clusters, and too low would result in insufficient clusters.

20. **What is the purpose of denaturing the DNA before the sequencing begins? What is used to do this?**

Sodium hydroxide is used to separate the two strands of DNA to allow for ligation and synthesis by sequencing.

21. **What is the flow cell and why is it important to thoroughly clean this before sequencing?**

A flow cell is a glass slide or silicon chip with a series of microchannels or wells that provide a solid surface for the sequencing reactions. Cleaning the flow cell thoroughly reduces the risk of contamination from foreign DNA or other substances.

22. **How long does the sequencer run for and what team needs to be notified when the process is complete?**

Approximately 20–24 hours. Bioinformatics need to be notified when the process is complete.