

Amplify *pfama1* gene with heminested primers containing multiplex identifying (MID) sequences

- The heminested (two-step) PCR approach increases ability to amplify *pfama1* from low parasitemia samples.
- Nested PCR is performed in duplicate to improve haplotype frequency estimates and identify potential false haplotypes due to PCR errors.
- MID sequences are “barcodes” located at the 5'-end of the forward PCR primer that provide a unique ID when PCR amplicons are pooled (in this case 24 separate MID).

Visualize PCR amplicons via agarose gel electrophoresis

- PCR amplicons are visualized using gel electrophoresis and positives are selected based on the correct size band.
- PCR amplicons are purified to remove primers, PCR enzymes, salts, and dNTPs.
- Nucleic acid concentrations for positive PCR amplicons are determined to ensure that the same concentration of nucleic acid from each amplicon is added during index preparation.

Index MID amplicon sets

- Amplicon pools are constructed so that each unique MID is represented only once in each of the pools.
- A unique index (another barcode) is then ligated to each pool, making each initial PCR amplicon uniquely identifiable based on the combination of MID and index barcode sequence.

Prepare deep sequencing libraries

- A sequencing library is then prepared for each indexed amplicon pool according to the manufacturer's protocol adding appropriate sequencing adapters (in this case IonTorrent).

Run amplicon deep sequencing on the Ion Torrent PGM Platform

- Library nucleic acid concentration is determined for each library so that equal concentrations of each library are added to the Ion 318 chip.

Perform data analysis using the SeekDeep Bioinformatics Pipeline

- Extractor step: demultiplexes samples based on MID sequence/barcodes, performs quality filtering steps.
- Qluster step: determines haplotypes and frequency data from sequencing reads for each sample.
- ProcessClusters step: removes chimeric sequences, compares replicates, determines haplotypes/haplotype frequencies for each sample, and produces final population haplotypes.