dbcAmplicons: A modular, highly multiplexed design for Illumina amplicon sequencing

Matthew L. Settles[1], Dan New[2], Bioinformatics Core Staff[1] , Alida Gerritsen[2]
[1] Genome Center Bioinformatics Core, University of California at Davis
[2] Genomics Resources Core, University of Idaho

UCDAVIS GENOME CENTER Bioinformatics Core

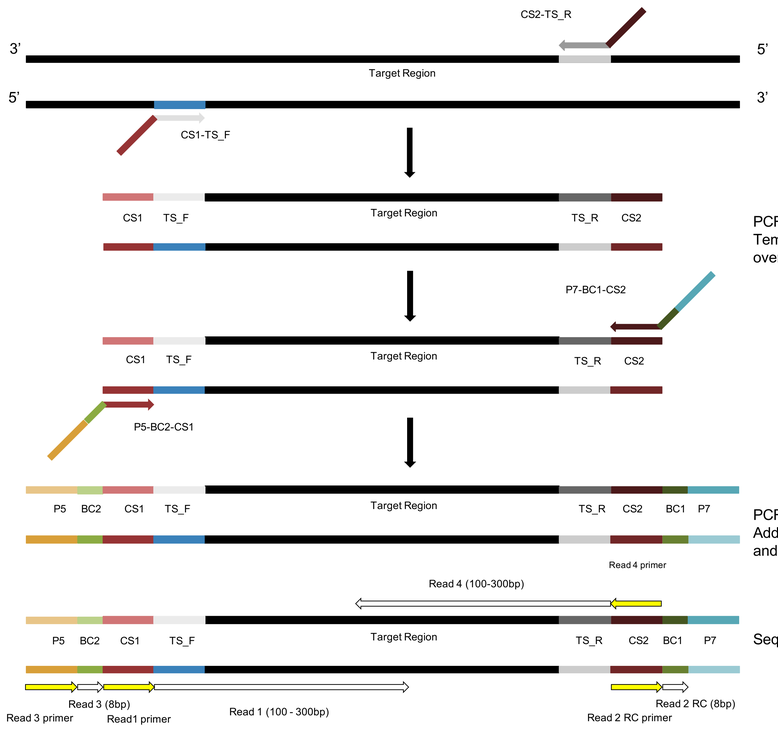
http://bioinformatics.ucdavis.edu

16S
Bacterial

ITS
LSU
Fungal

18S
Eukaryotic

Two-Step PCR



Two-step PCR reaction: the first PCR extracts out the target specific region and the second PCR adds on adapters and barcodes. Target specific primers include universal sequences CS1 and CS2 and possible phase-shifting bases, the second PCR extends the universal sequences with adapters and barcodes. The two steps allow for maximum flexibility in target specific primer usage and the ability to swap out targets, or include multiple targets in the same sequencing reaction without needing to purchase a large number of barcoded target specific primers.

Performed in Researcher's Lab

Target Specific Primer Design → Primer Validation and Optimization → DNA → Dual-barcoded Two-step PCR → Pooling And Quantification → Illumina Sequencing four read setup

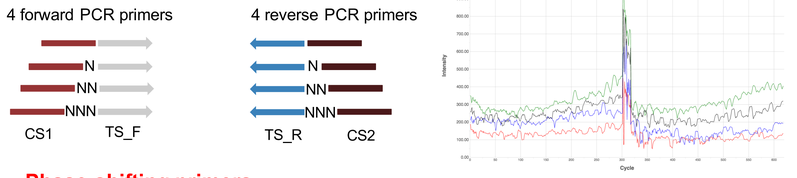
Performed in Core Facility

Python dbcAmplicons application

Illumina Sequencing four read setup → Demultiplex by barcodes and primer sequences → Downstream Analysis

Two-step PCR
highly multiplexed and modular amplicon workflow

Appropriate nucleotide diversity and cluster density are important for high quality data. Low nucleotide diversity in combination with high cluster density can lead to poor data quality and/or low data yield. Including phase-shifting primers (below) and adding in ~15% of a shotgun library produces high quality sequencing reactions.

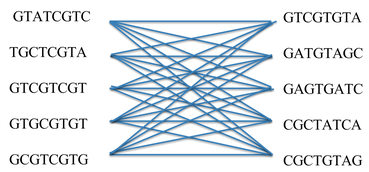


Phase-shifting primers

Dual barcoding allows for massively multiplexing of samples using only a relatively few primers

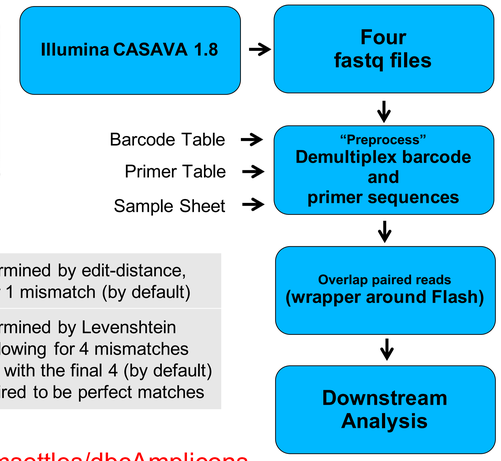
5 Pairs of Barcodes allows for multiplexing of **25 samples**.
32 Pairs can multiplex **1024 samples** in the same sequencing reaction

Pairing of BC1 and BC2 uniquely identifies sample



dbcAmplicons
A python pipeline for analysis of dbcAmplicon

dbcAmplicons Python Application



Barcode	Match determined by edit-distance, allowing for 1 mismatch (by default)
Primers	Match determined by Levenshtein Distance allowing for 4 mismatches (by default) with the final 4 (by default) bases required to be perfect matches

https://github.com/msettles/dbcAmplicons

Population Community Profiling (i.e. microbial, bacterial, fungal, etc.)

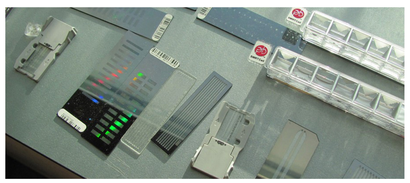
Classify	Wrapper around the MSU Ribosomal Database Project (RDP) Classifier for Bacterial and Archaeal 16S rRNA sequences, and Fungal 28S rRNA and ITS
Abundance	Reduce RDP classifier results to abundance tables and/or biom formated bile, rows are taxa and columns are samples ready for additional community analysis
Ecological Analysis	Analysis within R packages such as Vegan, Vegetarian, etc.

Benefits

- Maximum Flexibility, fewer target specific primers needed, multiple possible.
- Dual barcoding, allowing for massively multiplexing of samples
- dbcAmplicons software for analysis

Drawbacks

- Two – step PCR reaction
- Sequence the target specific primer



Strong collaborative ties between core facilities with unparalleled resources

Data Analysis

Intelligently analyzing data from genomics, and other projects, to help drive research forward

Research Computing

Helping build **high-performance software and hardware** bioinformatics solutions

Training

Providing **acclaimed training workshops** that will equip people with in-demand bioinformatics skills

For more information, contact us at ***bioinformatics.core@ucdavis.edu***. All project inquiries are free, and we encourage you to meet with us to see how we can assist your research program.