Custom amplicon sequencing on Illumina

NGI can not guarantee results for any custom library constructs, i.e. construct using third-party components such as custom sequence primers and adapters. It is up to each user to make sure the library is correctly made, but we will do our best to help with the design process. This document is a summary of the NGI guidelines for making Illumina-compatible libraries from amplicons.

Please email us on support@ngisweden.se if you have questions.

---

1. Overview

The following strategy is mainly used for 16S sequencing (see e.g. reference 2 for details of 16S specific primers), but can of course be used for any kind of amplicons. It's important to be aware of some of the limitations of Illumina sequencing when planning an experiment, so here are a few things to keep in mind:

1. The total length of the finished library should not exceed 1kb since clustering on the Illumina flowcell will not be efficient.
2. If several different amplicons are pooled together the size distribution should be similar since shorter libraries cluster with a higher efficiency.
3. At the moment of writing, the longest possible reads on the Illumina MiSeq is 2x300bp. If the whole amplicon sequence is needed it's recommended to aim for ca 50bp overlap of the reads, meaning ca 550bp.
4. Low diversity libraries such as amplicons require a spike in of a high diversity library (e.g. PhiX control DNA) to 5-10% molar ratio, meaning that ca 10% of the reads will be lost.
5. When pooling indices each color-set of nucleotides (green [G/T] and red [A/C]) need to be represented in each cycle of each index for high quality index reads. For example:

Good Examples				Bad Examples			
Index 1		Index 2		Index 1		Index 2	
705	GGACTCCT	503	TATCCTCT	705	GGACTCCT	502	CTCTCTAT
706	TAGGCATG	503	TATCCTCT	706	TAGGCATG	502	CTCTCTAT
701	TAAGGCGA	504	AGAGTAGA	701	TAAGGCGA	503	TATCCTCT
702	CGTACTAG	504	AGAGTAGA	702	CGTACTAG	503	TATCCTCT
	✓✓✓✓✓✓✓✓		✓✓✓✓✓✓✓✓		✓✓✓✓✓✓✓✓		✓✓✓✓XXXX
	✓= signal in both color						
	X= signal missing in one color channel						

The basics of the strategy used here is that the amplicon of interest is amplified with adapter containing primers. A second PCR is then used to add Illumina specific adapters with unique barcode sequences to these amplicons.

2. Primer design

This is a summary of the nested PCR strategy described by Anders Andersson et al (2).

Primers for PCR #1 (amplicon specific)

Forward:

5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-**[amplicon specific sequence]**

Reverse:

5′-AGACGTGTGCTCTTCCGATCT-**[amplicon specific sequence]**

Primers for PCR #2 (adding Illumina specific adapters and indexes)

Forward:

5′-AATGATACGGCGACCACCGAGATCTACAC-**[i5 index]**-ACACTCTTTCCCTACACGACG

Reverse:

5′-CAAGCAGAAGACGGCATACGAGAT-**[i7 index]**-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

IMPORTANT! Please note that the i7 index should be in reverse complement when taken from the lists (see below).

Example index sequences

Table 1. 8 base indexes from Nextera XT index kit (2)

Index 1 (i7)	Sequence	Index 2 (i5)	Sequence
N701	TAAGGCGA	S501	TAGATCGC
N702	CGTACTAG	S502	CTCTCTAT
N703	AGGCAGAA	S503	TATCCTCT
N704	TCCTGAGC	S504	AGAGTAGA
N705	GGACTCCT	S505	GTAAGGAG
N706	TAGGCATG	S506	ACTGCATA
N707	CTCTCTAC	S507	AAGGAGTA
N708	CAGAGAGG	S508	CTAAGCCT
N709	GCTACGCT		
N710	CGAGGCTG		
N711	AAGAGGCA		
N712	GTAGAGGA		

3. References

1. 16S Metagenomic Sequencing Library Preparation (Illumina Part # 15044223 Rev. A)
2. Strategy used by Anders Andersson group at SciLifeLab can be found on Github (https://github.com/EnvGen/LabProtocols).