Bioinformatics I/O

Tips && tricks from a cluster of bioinformaticians

Illumina adapter and primer sequences

Illumina Adapter and Primer Sequences

Illumina libraries are normally constructed by ligating adapters to short fragments (100 – 1000bp) of DNA. The exception to this is if Nextera is used (see end of this post) or where PCR amplicons have been constructed that already incorporate the P5/P7 ends that bind to the flowcell.

Illumina Paired End Adapters (cannot be used for multiplexing)

Top adapter
5' ACACTCTTTCCCTACACGACGCTCTTCCGATC*T 3'

Bottom adapter
5' P-GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG 3'

Note that the last 12nt of the adapters are complementary (when the bottom adapter is viewed 3'-5' as below) hence the name 'forked adapters'. The adapters are annealed together then ligated to *both ends* of the library DNA. The bottom adapter is 5'-phosphorylated in order to promote ligation. The top adapter has a phosphorothioate bond (*) before the terminal T, ensuring that exonucleases cannot digest the T overhang that pairs to the A-tail added to library fragments.

Structure of Illumina forked PE adapter

```
5' ACACTCTTTCCCTACACGACGCTCTTCCGATC*T 3'
                      ||||||||||||      <-Ligated to library fragment
3' GAGCCGTAAGGACGACTTGGCGAGAAGGCTAG-P 5'
```

PCR with partially complementary primers then extends the ends and resolves the forks, adding unique termini that bind to the oligos on the surface of the flow cell (P5 blue/P7 red, also see diagram at foot of page).

PE PCR Primer 1.0 (P5) (same as universal adapter)
5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T 3'

PE PCR Primer 2.0 (P7)
5' CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T 3’

Structure of Illumina TruSeq™ indexed forked adapters

```
5’       AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T
                                                  ||||||||||||   <ligate
3’GTTCGTCTTCTGCCGTATGCTCTA-index-CACTGACCTCAAGTCTGCACACGAGAAGGCTAG-P    here
```

The last 12nt of the adapters are complementary, allowing them to anneal and form the forked structure. The adapter is ligated to both ends of the A-tailed DNA library, generating larger floppy overhangs than with the paired-end adapters on the first page. Note that while the top adapter is identical to the Illumina Universal oligo, the bottom adapter is different to the PE adapter in the purple highlighted section. The adapter already has the index and complete P7/P5 ends.

PCR with the following primers resolves the forked ends to generate products with no floppy overhangs. The sequences that bind to the flow cell (P5 blue/P7 red) finish up at opposite ends of the library fragments.

PCR Primer 1.0 (P5)
5’ AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA 3’

PCR Primer 2.0 (P7)
5’ CAAGCAGAAGACGGCATACGAGAT 3’

The following oligos (provided in the MiSeq reagent cartridge) are used to prime the sequence reads. Note that the index read primer is complementary to the Read 2 sequencing primer (see diagram below). This is used to sequence the hexamer index tag in the forward direction after read 1 is complete, before the reverse strand is synthesised by bridge amplification.

Multiplexing Read 1 Sequencing Primer
5' ACACTCTTTCCCTACACGACGCTCTTCCGATCT

Multiplexing Index Read Sequencing Primer
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC

Multiplexing Read 2 Sequencing Primer
5' GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

When ordering primers for use in Illumina libraries, make certain to include the modifications (e.g. 5’-phosphorylation and phosphorothioate bonds on the 3' terminal nucleotide) and ensure the oligos are PAGE purified. Even small amounts of n-1 primers will lead to messy out-of-phase sequencing and cause clusters to fail filtering. Costs per oligo for 0.2µmole synthesis scale and PAGE are in region of £40.

TruSeq™ DNA Sample Prep Kit v2

There are currently two versions of the kit, each with 12 different adapters that incorporate unique index tags – allowing samples to be multiplexed on the same sequencing run.

Kit A contains indexes: 2, 4, 5, 6, 7, 12, 20, 21, 22, 23, 25, 27.
Kit B contains indexes: 1, 3, 8, 9, 10, 11, 13, 14, 15, 16, 18, 19.

NOTE that all the indexed adapters should be 5'-Phosphorylated. For unknown reasons adapters 13-27 have an extra 2 bases (these are not used for the indexing). Illumina also reserve certain numbers e.g. 17, 24 and 26.
The 6-base index tag sequences are in italics below.

TruSeq Universal Adapter
5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T

TruSeq Adapter, Index 1
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***ATCACG***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 2
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***CGATGT***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 3
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***TTAGGC***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 4
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***TGACCA***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 5
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***ACAGTG***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 6
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GCCAAT***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 7
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***CAGATC***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 8
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***ACTTGA***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 9
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GATCAG***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 10
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***TAGCTT***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 11
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GGCTAC***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 12
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***CTTGTA***ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 13
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***AGTCAA****CA*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 14
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***AGTTCC****GT*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 15
5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***ATGTCA****GA*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 16

5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***CCGTCC****CG*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 18
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GTCCGC****AC*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 19
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GTGAAA****CG*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 20
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GTGGCC****TT*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 21
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GTTTCG****GA*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 22
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***CGTACG****TA*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 23
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***GAGTGG****AT*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 25
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***ACTGAT****AT*ATCTCGTATGCCGTCTTCTGCTTG
TruSeq Adapter, Index 27
5′ GATCGGAAGAGCACACGTCTGAACTCCAGTCAC***ATTCCT****TT*ATCTCGTATGCCGTCTTCTGCTTG

NEBNext® DNA Library Prep

The NEB kit uses a short adapter which is supplied as a single self-complimentary oligo forming a stem-loop. It has a Uracil base that is later cleaved and removed by Uracil Glycosylase and base excision repair enzyme mix (USER).

NEBNext adaptor for Illumina

5′ P-GATCGGAAGAGCACACGTCTGAACTCCAGTC-U-ACACTCTTTCCCTACACGACGCTCTTCCGATC*T 3′

Oligo is designed to self-anneal forming a stem-loop structure as below. This may help to prevent formation of adapter dimers during ligation.

```
 /ACACTCTTTCCCTACACGACGCTCTTCCGATC*T 3′
U                    ||||||||||||       <-Ligated to library fragments
 \-CTGACCTCAAGTCTGCACACGAGAAGGCTAG-P 5′
```

```
                           5′  ACACTCTTTCCCTACACGACGCTCTTCCGATC*T------
Uracil cleaved and removed                          ||||||||||||
to form forked structure    3′   CTGACCTCAAGTCTGCACACGAGAAGGCTAG--------
```

The Index tags and the P5/P7 ends are added by PCR using universal and tagged primers. The end result is exactly the same as TruSeq.

NEBnext Universal primer
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T

NEBnext Indexed primers 1 – 12 (6-mer indexes)

Index 1 CAAGCAGAAGACGGCATACGAGAT***CGTGAT***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 2 CAAGCAGAAGACGGCATACGAGAT***ACATCG***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 3 CAAGCAGAAGACGGCATACGAGAT***GCCTAA***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 4 CAAGCAGAAGACGGCATACGAGAT***TGGTCA***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 5 CAAGCAGAAGACGGCATACGAGAT***CACTGT***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 6 CAAGCAGAAGACGGCATACGAGAT***ATTGGC***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 7 CAAGCAGAAGACGGCATACGAGAT***GATCTG***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 8 CAAGCAGAAGACGGCATACGAGAT***TCAAGT***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 9 CAAGCAGAAGACGGCATACGAGAT***CTGATC***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 10 CAAGCAGAAGACGGCATACGAGAT***AAGCTA***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 11 CAAGCAGAAGACGGCATACGAGAT***GTAGCC***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T
Index 12 CAAGCAGAAGACGGCATACGAGAT***TACAAG***GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T

SURESELECT (POST-CAPTURE INDEXING)

This begins with a shorter bottom adapter that is extended to add the P5 end in the pre-capture PCR. The post-capture PCR step adds the index and P7 end. Note The NEB adapter is more efficient than the InPE adapter in my comparative tests.

InPE adapter (indexing paired end adapter)

```
5' ACACTCTTTCCCTACACGACGCTCTTCCGATC*T  3’
                        ||||||||||||      <-Ligated to library fragments
            3’ TCTGCACACGAGAAGGCTAG-P 5’
```

PRE-CAPTURE PCR

PCR Primer 1.0 [Tm 70deg]
5’ AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACG*A

Multiplex PCR Primer 2.0 [Tm 67deg]
5' GTGACTGGAGTTCAGACGTGTGCTCTTCCGAT*C

POST-CAPTURE PCR Indexing Primers
2nd PCR reaction (post-capture amplification) adds indexes and P7 sequence.

Universal Primer [Tm 75deg]
5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T

PCR Primer, Index 1 5’ CAAGCAGAAGACGGCATACGAGAT***CGTGAT***GTGACTGGAGTT*C
PCR Primer, Index 2 5’ CAAGCAGAAGACGGCATACGAGAT***ACATCG***GTGACTGGAGTT*C
PCR Primer, Index 3 5’ CAAGCAGAAGACGGCATACGAGAT***GCCTAA***GTGACTGGAGTT*C

PCR Primer, Index 4 5′ CAAGCAGAAGACGGCATACGAGAT***TGGTCA***GTGACTGGAGTT*C
PCR Primer, Index 5 5′ CAAGCAGAAGACGGCATACGAGAT***CACTGT***GTGACTGGAGTT*C
PCR Primer, Index 6 5′ CAAGCAGAAGACGGCATACGAGAT***ATTGGC***GTGACTGGAGTT*C
PCR Primer, Index 7 5′ CAAGCAGAAGACGGCATACGAGAT***GATCTG***GTGACTGGAGTT*C
PCR Primer, Index 8 5′ CAAGCAGAAGACGGCATACGAGAT***TCAAGT***GTGACTGGAGTT*C
PCR Primer, Index 9 5′ CAAGCAGAAGACGGCATACGAGAT***CTGATC***GTGACTGGAGTT*C
PCR Primer, Index 10 5′ CAAGCAGAAGACGGCATACGAGAT***AAGCTA***GTGACTGGAGTT*C
PCR Primer, Index 11 5′ CAAGCAGAAGACGGCATACGAGAT***GTAGCC***GTGACTGGAGTT*C
PCR Primer, Index 12 5′ CAAGCAGAAGACGGCATACGAGAT***TACAAG***GTGACTGGAGTT*C

Guidelines for Low-Level Pooling

Some sequencing experiments require the use of fewer than 12 index sequences in a lane with a high cluster density. In such cases, a careful selection of indexes is required to ensure optimum cluster discrimination by having different bases at each cycle of the index read. Illumina recommends the following sets of indexes for low-level pooling experiments

Pool of 2 samples:
Index 6 GCCAAT
Index 12 CTTGTA

Pool of 3 samples:
Index 4 TGACCA
Index 6 GCCAAT
Index 12 CTTGTA

Pool of 6 samples:
Index 2 CGATGT
Index 4 TGACCA
Index 5 ACAGTG
Index 6 GCCAAT
Index 7 CAGATC
Index 12 CTTGTA

Nextera Sample Preparation

The sequences in Nextera libraries are different to all the other workflows.

Nextera® transposase sequences (FC-121-1031, FC-121-1030)

The Tn5 transposase cuts the sample DNA and adds the following sequence at either end of each fragment, with the highlighted sequence next to the library insert.

5′ TCGTCGGCAGCGTC*AGATGTGTATAAGAGACAG*
Read 1 >

5′ GTCTCGTGGGCTCGG*AGATGTGTATAAGAGACAG*
Read 2 >

Nextera® Index Kit – PCR primers (FC-121-1012, FC-121-1011)

PCR with the following primers adds the P5 and P7 termimi that bind to the flowcell and also the dual 8bp index tags (denoted by the i5 and i7 below).

5′ AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC

5′ CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGG

If trimming adapters from Nextera runs should cut the reads at CTGTCTCTTATACACATCT instead of the usual AGATCGGAAGAGC. Use of cutadapt, trim_galore or similar program is recommended with custom adapter specified.

Gavin Wilkie

This entry was posted in Deep Sequencing and tagged Illumina, next-generation sequencing, NGS, primers, tagging, TrueSeq on January 26, 2015 [http://bioinformatics.cvr.ac.uk/blog/illumina-adapter-and-primer-sequences/] by Gavin Wilkie.

Anup • 2 years ago

Anyone knows how QXT and Nextera Adapters are attached during Transposase treatment? Any note on the architecture of the adaptors?

3 ^ | v • Reply • Share ›

Maike • 2 years ago

How can the Nextera Sequences be different from all the other ones? I was trying to understand how the different preps work and as far as I understood the padding region is where illumines sequencing primers bind. Wouldn't they need to share some similarity at least?

1 ^ | v • Reply • Share ›

sam • 2 years ago

Does anyone know the annealing conditions that can generate the PE duplex adapter. I have ordered the oligos separately and would want to duplex them in house.

1 ^ | v • Reply • Share ›

Kshitiz ➚ sam • 2 years ago

Mix 2 microlitre of each complementary oligo in 1X ligase buffer. In thermocycler, set up following cycle 98 deg celsius X 2min and then, ramp 0.2 deg celsius till 4 degree celsius. Your oligos will be annealed

^ | v • Reply • Share ›

Amin Zia • 2 years ago

Why is it necessary to phosphorylate the 5' of those adapters when the DNA fragment has already gone through end-repair/phosphorylation which I suppose adds the phosphate group to the 5' end?

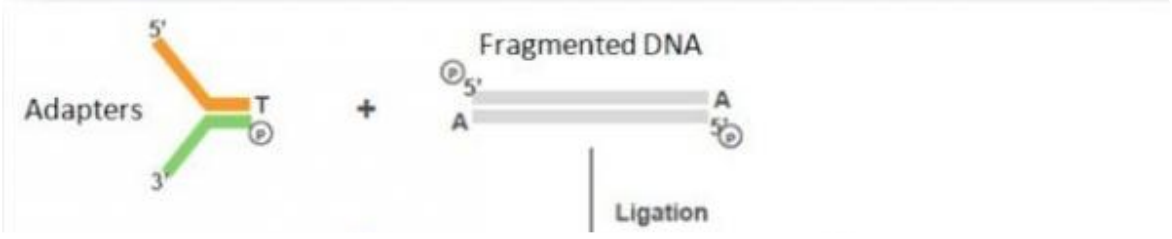
1 ^ | v • Reply • Share ›

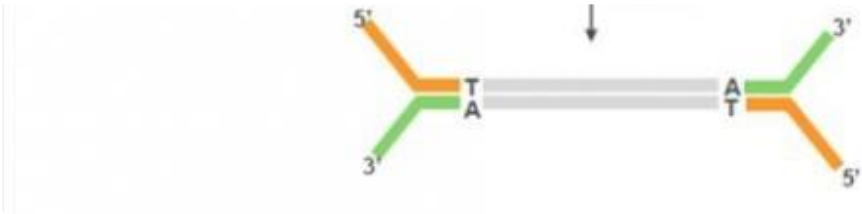
Gavin Wilkie ➚ Amin Zia • 2 years ago

The Illumina adapter is a partially double stranded DNA adapter, which is ligated to double stranded DNA fragments. The adapter has both a 5' and a 3' end.

The 3' end of the adapter (top strand, orange in diagram below) is added to the 5' end of the DNA fragment (this is already phosphorylated as the poster has pointed out).

However, the 5' end of the adapter (bottom strand, green in diagram below) is added to the 3' end of the DNA (not the phosphorylated 5' end). The 3' end of the DNA has a 3'OH group, and the 5' end of the adapter has to be phosphorylated in order to be incorporated. If the adapter 5' end is not phosphorylated, the bottom strand of the adapter will not be ligated.





1 • Reply • Share ›

Albert Palleja • 3 years ago

Hi Gavin great post. Is there any link where one can download the standard illumina adapters as a fasta file? Well, is it save enough to assume that if your reads have been sequence by illumina technology, screning against the standard illumina ones it should fine or should I contact for every project the sequencing center and ask for the adapters?

1 • Reply • Share ›

Renato Puga • 3 years ago

Nice shot! tks

1 • Reply • Share ›

Quan Gu • 3 years ago

It is helpful for all ILLMUNIA users!

1 • Reply • Share ›

bljog • 4 years ago

Thanks for your first post Gavin. Hopefully we can convince you to do some more in the future. This is useful info to have here. Cheers.

1 • Reply • Share ›

Ouhao Lin • 13 hours ago

hi, is the adapter contains seq_primer, index and oligos ligated to the flowcell surface(P5 or P7)

• Reply • Share ›

Scott Riley • 10 months ago

Anyone use the newest 3RAD protocol? If so, which P5 sequence do you use for the PCR step with the i7?

• Reply • Share ›

John Chiang • a year ago

Can we use a universal adaptor first for ligation and then use different index PCR primers to generate library? Is this approach more cost effective?
thanks

John

• Reply • Share ›

Amin Zia • 3 years ago

Great post. I was wondering if there is any protocol to ligate only 5' adapters or only 3' adapters

Great post. I was wondering if there is any protocol to ligate only 5' adapters or only 3' adapters to dsDNA.

^ | v • Reply • Share ›

bljog ➚ Amin Zia • 3 years ago

Gavin says:
No, I'm not aware of any protocol to ligate adapters to one side only.

The standard Illumina protocols for DNA library preparation ligate the same partially double-stranded adapter to both ends. This is then resolved into the different P5 or P7 adapters at each end of the molecule by PCR.

1 ^ | v • Reply • Share ›

liu logen ➚ bljog • 3 years ago

it seems that PCR could not ensure different ends(P5 and P7) of Target DNA, its possibility is 50%??

1 ^ | v • Reply • Share ›

Gavin Wilkie ➚ liu logen • 2 years ago

It is 100%, you will see if you draw the PCR intermediates and final structures that the adapter is resolved to a different termial sequence on each end of the fragment. This is why the Illumina Y-shaped adapters are very efficient

^ | v • Reply • Share ›