Custom multiplexing on the Illumina GA using barcoding adapters

Barcoding adapters used by Cronn et al. (2008; Nucl Acids Res vol. 36 (19), e122; supp. table 2] – *‘barcoding’ sequences are indicated by turquoise highlighting:*

Tag CCT

```
Adapt CCT1: 5'-pGGAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG
Adapt CCT2: 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCT
```

Tag GGT

```
Adapt GGT1: 5'-pCCAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG
Adapt GGT2: 5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGT
```

etc. for other tags *[note that Cronn et al. carried out single-end reads on an Illumina GA I]*

Representation of above in an adapter-tagged library fragment [shown for CCT adapter only; vertical dashes represent base-pairing]*:*

```
5′- ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTnnnnnnnnAGGAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG
                        ¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦
3′-GTTCGTCTTCTGCCGTATGCTCGAGAAGGCTAGAGGAnnnnnnnnTCCTCTAGCCTTCTCGCAGCACATCCCTTTCTCACA
```

∴ barcoding tag becomes incorporated at 5’ end of library fragment strands, and is removed by software prior to alignment of sequences

Generic strategy for use of custom multiplex adapters for paired-end sequencing on the Illumina GAII

The unique barcoding tag is represented by 5’-XYZ-3’, where 3’-xyz-5’ is the complementary sequence; note that the tag and its complement must be preceded by a T (complement A) to allow extension during amplification – see PCR steps below, with the PEMx1 adapter also carrying the 3’-T overhang for annealing to A-tailed library fragments. In this schematic the unique part of the tag (‘XYZ’) is 3nt, but can be longer as required. Compared to a non-barcoded sample, sequence reads will contain the additional 4nt sequence ‘XYZT’ at the 5’ end, and this should be removed prior to alignment to the reference sequence. Note that PEMx1 and PEMx2 are identical to Illumina PE1 & 2 respectively except for the additional 4nt added by the barcode; PCR and sequencing primers are standard Illumina primers.

Paired-end multiplex (barcoding) adapter sequences [* = phosphorothioate bond; p = phosphate group; complementary regions are shown underlined]

```
PEMx1 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCTXYZ*T
PEMx2 5′-pzyxAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG
```

Ds paired-end multiplex (barcoding) adapter (‘Y’ structure):

‘Left arm’:

```
PEMx1 5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCTXYZ*T
                            ¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦
PEMx2 3′-GAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzp
```

‘Right arm’:

```
 pzyxAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG-3′
  ¦¦¦¦¦¦¦¦¦¦¦¦¦
T*ZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-5′
```

PCR amplification & sequencing of libraries tagged with custom multiplex (barcoding) adapters

Adapter-tailed library fragments *(yellow highlighting = library fragment insert; turquoise highlighting = barcoding tag):*

```
5′-ACACTCTTTCCCTACACGACGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG-3′
                      ||||||||||||||||||||||||||||||||||||||||||||
3′-GAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-5′
```

Denaturation, annealing of PCR primer PE1 to left-arm multiplex (barcode) adapter PEMx2:

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
                               |||||||||||||||||||||||||||||||||
                            3′-GAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-5′
```

note that the ‘A’ in the adaptor PEMx2 sequence (blue) is required for annealing at the 3’ end of primer PE1

Extension of PCR primer PE1 *(newly-synthesized sequence in purple):*

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3′
                               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                            3′-GAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-5′
```

Denaturation, annealing of PCR primer PE2 to newly-synthesized sequence on right arm:

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3′
                                                                                |||||||||||||||||||||||||||||||||
                                                                                TCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA-5′
```

Extension of PCR primer PE2 to form ‘asymmetric product’ *(newly-synthesized sequence in brown):*

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3′
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3′-GTTCGTCTTCTGCCGTATGCTCTAGCCAGAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA-5′
```

Second round of PCR – annealing & extension of PCR primers PE1 and PE2 to strands of ‘asymmetric product’ from round 1:

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3′-GTTCGTCTTCTGCCGTATGCTCTAGCCAGAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA-5′
```

and

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3′
                                                                                |||||||||||||||||||||||||||||||||
                                                                                TCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA-5′
```

⇩

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT-3′
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3′-GTTCGTCTTCTGCCGTATGCTCTAGCCAGAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA-5′
```

and

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3′
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3′-GTTCGTCTTCTGCCGTATGCTCTAGCCAGAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA-5′
```

Subsequent rounds of PCR will result in exponential amplification of full-length strands

Annealing & extension of PE sequencing primers PE1 and PE2

```
5′-CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTXYZTnnnnnnnnAzyxAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT-3′
                                                                                ¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦
Sequencing primer PE1:                                                       ←←←TCTAGCCTTCTCGCAGCACATCCCTTTCTCACA-5′


Sequencing primer PE2:  5′-CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT→→→
                           ¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦¦
3′-GTTCGTCTTCTGCCGTATGCTCTAGCCAGAGCCGTAAGGACGACTTGGCGAGAAGGCTAGAxyzAnnnnnnnnTZYXTCTAGCCTTCTCGCAGCACATCCCTTTCTCACATCTAGAGCCACCAGCGGCATAGTAA-5′
```

extension of both primers results in incorporation of barcode ‘XYZT’ at 5’ end of insert read; note that in the Illumina multiplexing system, the barcode is read in a separate dedicated sequence reaction, subsequent to read 1.