automatic adapter detection

DRASearch

Send Feedback | Search Home | DRA Home

SRX262175 FASTQ SRA

Experiment Detail	
Title	mother_trio_2
Design Description	barcoded_small_RNA_cDNA_PMID_23440203 3' adapter: TCGATTCGTATGCCGTCTTCTGCTTG
Organism	Homo sapiens

Library Description	
Name	
Strategy	miRNA-Seq
Source	TRANSCRIPTOMIC
Selection	size fractionation
Layout	SINGLE
Construction Protocol	

Navigation			
Submission	SRA065726	FTP	
Study	SRP018255		
Sample	SRS409117		
Run	SRR822468	FASTQ	SRA

Tuschl exRNA-seq

DRASearch

Send Feedback | Search Home | DRA Home

SRX272861 FASTQ SRA

Experiment Detail	
Title	GSM1131187: CGC2 (miRNA); Homo sapiens; RNA-Seq
Design Description	
Organism	Homo sapiens

Library Description	
Name	
Strategy	RNA-Seq
Source	TRANSCRIPTOMIC
Selection	size fractionation
Layout	SINGLE
Construction Protocol	Total RNA was extracted using the miRNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. Small RNA fraction (< 200nt) was eluted separately from large RNA fraction. Small RNA libraries were prepared using Bioo Scientific NEXTflex Small RNA Sequencing Kit and NEXTflex small RNA Barcode Primer Set A (Bioo Scientific Corporation) following the manufacturer's protocol. Size selection of 140-160 bp fragments corresponding to mature miRNAs and adapters was carried out by gel electrophoresis with 10% Mini-PROTEAN TBE Precast gels (Bio-Rad Laboratories). 0.66 pM of individual indexed samples were pooled for template hybridization and clustering/amplification was performed using TruSeq SR Cluster Kit v2-cBot-GA (Illumina).

Navigation			
Submission	SRA074994	FTP	
Study	SRP021911		
Sample	SRS417370		
Run	SRR836168	FASTQ	SRA
	SRR836169	FASTQ	SRA

ovary smallRNA-seq

strategy

- map reads to genome using 20nt seed (i.e. don't penalise for mismatches due to adapter)
- retain only unique maps from sense strand with 22 matched bases at the start of the read
- remove first 22 bases and any 3' poly-A's and select top two potential adapter sequences
- keep the longest of these two adapter sequences

command line

```
bowtie2 --no-head -p 8 --local -D 15 -R 2 -N 0 -L 20 -i S,1,0.75 -k 2 --upto 10000000 -x hg19 -U SRR822433.fastq
| awk '{if ($5==255) print $0}' | awk '{if ($2==0) print $3"\t"$4"\t"$6"\t"$10}' | grep "[[:space:]]22M[0-9]
[0-9]S" | awk '{print substr($4,23)}' | sed 's/[A]*$//' | sort | uniq -c | sort -rnk 1 | awk '{if ($1 > 99) print
$0}' | tee > SRR822433.potentialAdapters | head -n 2 | awk '{print length, $0}' | sort -nr | head -1 | awk
'{print $3}'
```

1) map using bowtie
2) use only unique mapped reads on +ve strand
3) use only reads with 22 matched reads at 5' end
4) remove first 22 bases and any 3' A's
5) count unique potential adapters and save to disk
6) of the top 2 most abundant candidates, keep the longest

```
## SRR822433
## actual adapter = TGTGTTCGTATGCCGTCTTCTGCTTG
   38650 TGTGTTCGTATGCCGTCTTCTGCTTG
   21276 GTGTTCGTATGCCGTCTTCTGCTTG
    3884 ATGTGTTCGTATGCCGTCTTCTGCTTG
    3188 TTGTGTTCGTATGCCGTCTTCTGCTTG
     410 AATGTGTTCGTATGCCGTCTTCTGCTTG
     340 TGTTCGTATGCCGTCTTCTGCTTG
     315 GTGTGTTCGTATGCCGTCTTCTGCTTG
     249 ATTGTGTTCGTATGCCGTCTTCTGCTTG
     155 TTTGTGTTCGTATGCCGTCTTCTGCTTG
     132 TGTGTTCGTATGCCGACTTCTGCTTG
```

```
## SRR822462
## actual adapter = TCCACTCGTATGCCGTCTTCTGCTTG
   90524 TCCACTCGTATGCC
   13724 TTCCACTCGTATGC
   12082 CCACTCGTATGCCG
    3721 ATCCACTCGTATGC
    2711 TTTCCACTCGTATG
     763 CTCCACTCGTATGC
     684 GATCCACTCGTATG
     505 GTCCACTCGTATGC
     485 CTTCCACTCGTATG
     425 ATTCCACTCGTATG
```

```
## SRR822462
## actual adapter = TCTCCTCGTATGCCGTCTTCTGCTTG
  153092 TCTCCTCGTATGCCGTCTTCTGCTTG
   53881 CTCCTCGTATGCCGTCTTCTGCTTG
   10001 ATCTCCTCGTATGCCGTCTTCTGCTTG
    8028 TTCTCCTCGTATGCCGTCTTCTGCTTG
    1329 AATCTCCTCGTATGCCGTCTTCTGCTTG
     685 TCTCCTCGTATGCCGTCTTTTGCTTG
     629 ATTCTCCTCGTATGCCGTCTTCTGCTTG
     502 GTCTCCTCGTATGCCGTCTTCTGCTTG
     482 TTTCTCCTCGTATGCCGTCTTCTGCTTG
     400 TCTCCTCGTATGCCGTCTTCTGCTTGG
```