Figure S2:

**Reference preparation**

- Combined genome, ERCC and R-luc sequences
  - C-to-T and G-to-A conversion of forward and reverse strand
- Single annotation of tRNA
- Single annotation of rRNA
  - C-to-T conversion of forward strand

**Mapping strategy**

- All reads mapped to Genome, ERCC and R-luc spike-in
- All reads mapped to tRNA and rRNA sequences separately
- Obtain BAM files
- Uniquely mapping converted reads were replaced by original unconverted read

**Filter application in each library**

- Combine bsRNA-seq fraction libraries into composite library
  - Retain converted reads with ≤3 C per read
  - Site calling at each cytosine position
    - \( S/N90 \); Signal to noise ratio filter 0.9
  - \( 30RC \); Minimum coverage of ≥30 reads per site
  - \( 5C \); Minimum frequency of ≥5 non-converted C at each site
  - \( 80CT \); Remove variable site requiring ≥80% C+T at each site

**Fraction non-conversion analysis**

- Deconvolute high confidence sites into fractions

**Fraction clustering analysis**

- Require coverage in 9 out of 12 samples
  - ≥10 read average coverage across respective fraction
- F1234 clustering
- F234 clustering
  - \( \text{Mfuzz} (m=2) \) clustering acquiring \( c=9 \) clusters

**High confidence candidate site selection**

- Minimum average non-conversion of ≥10% across replicates

**Differentially methylated sites**

- Fraction A non-conversion
- Fraction B non-conversion
  - etc
- ≥10 read average coverage across fractions
- \( \text{Logistic regression test} \)
  - Relative non-conversion change ≥10%
  - \( q\text{-value} ≤0.05 \)
- Differentially non-converted sites