Figure S2

A Error-correction based on base-call quality

Ref:	AGCGTTGACGTGGACTT			
Read1:	GCGTTG A CGTGGACT	error correction	Read1:	GCGTTG A CGTGGACT
Read2:	AGCGTTG <mark>T</mark> CGTGGAC		Read2:	AGCGTTGACGTGGAC
Read3:	CGTTG <mark>T</mark> CGTGGACTT	minimum mean	Read3:	CGTTGACGTGGACTT
Read4:	AGCGT/TG A CGTGGAC	base-call qualit =10	Read4:	AGCGTTG A CGTGGAC
Read5:	GCCTTG A CGTGGACT	(-xa 10)	Read5:	GCGTTG A CGTGGACT
	K 4	()(2,10)		
	Q: 20 Q: 6 ->	Mean Q: 8		

B Error-correction based on allele frequency

```
Ref:
       AGCGTTGACGTGGACTT
        GCGTTGACGTGGACT
Read1:
                                              Read1:
                                                       GCGTTGACGTGGACT
                             error correction
Read2: AGCGTTGTCGTGGAC
                                              Read2: AGCGTTGTCGTGGAC
                                                        CGTTGTCGTGGACTT
Read3:
         CGTTGTCGTGGACTT
                                              Read3:
                             minimum allele
Read4: AGCGTTGACGTGGAC
                                              Read4: AGCGTTGACGTGGAC
                             frequency =0.2
                                              Read5:
                                                       GCGTTGACGTGGACT
Read5:
        GCGTTGCCGTGGACT
                             (-xf 0.2)
Read6:
        GCGTTGACGTGGACT
                                              Read6:
                                                       GCGTTGACGTGGACT
```

C Error-correction based on allele frequency for pooled sample data

Read5 (sample B): allele C → frequency=0.5 → not corrected

```
Ref:
       AGCGTTGACGTGGACTT
Read1:
        GCGTTGACGTGGACT
                                               Read1:
                                                        GCGTTGACGTGGACT
                              error correction
Read2: AGCGTTGCCGTGGAC
                                               Read2: AGCGTTGACGTGGAC
                                               Read3:
Read3:
          CGTTGACGTGGACTT
                                                         CGTTGACGTGGACTT
                             minimum allele
Read4: AGCGTTGACGTGGAC
                                               Read4: AGCGTTGACGTGGAC
                             frequency =0.4
Read5:
        GCGTTGCCGTGGACT
                                               Read5:
                                                        GCGTTGCCGTGGACT
                             (-ms hetero -xf 0.3)
Read6:
        GCGTTGACGTGGACT
                                               Read6:
                                                        GCGTTGACGTGGACT
Read2 (sample A): allele C → frequency=0.25 → corrected
```

Figure S2. Algorithm of Coval-Refine error correction.

In the error correction mode of Coval-Refine, the error correction is conducted according to two criteria specified by user. One is based on base-call quality (A): correcting mismatch bases when the mean quality of non-reference bases supported at the site is lower than a minimum base quality (10 in this case), and the other based on allele frequency (B): correcting mismatch bases when the frequency of the non-reference allele is lower than a minimum allele frequency (0.2 in this case). The allele frequency-based correction for pooled sample data can be performed for reads of each sample by specifying –ms option (C), where read names of sample A are marked in blue, and those of sample B in red. Mismatch bases that meet either criterion are corrected to match the corresponding reference base.