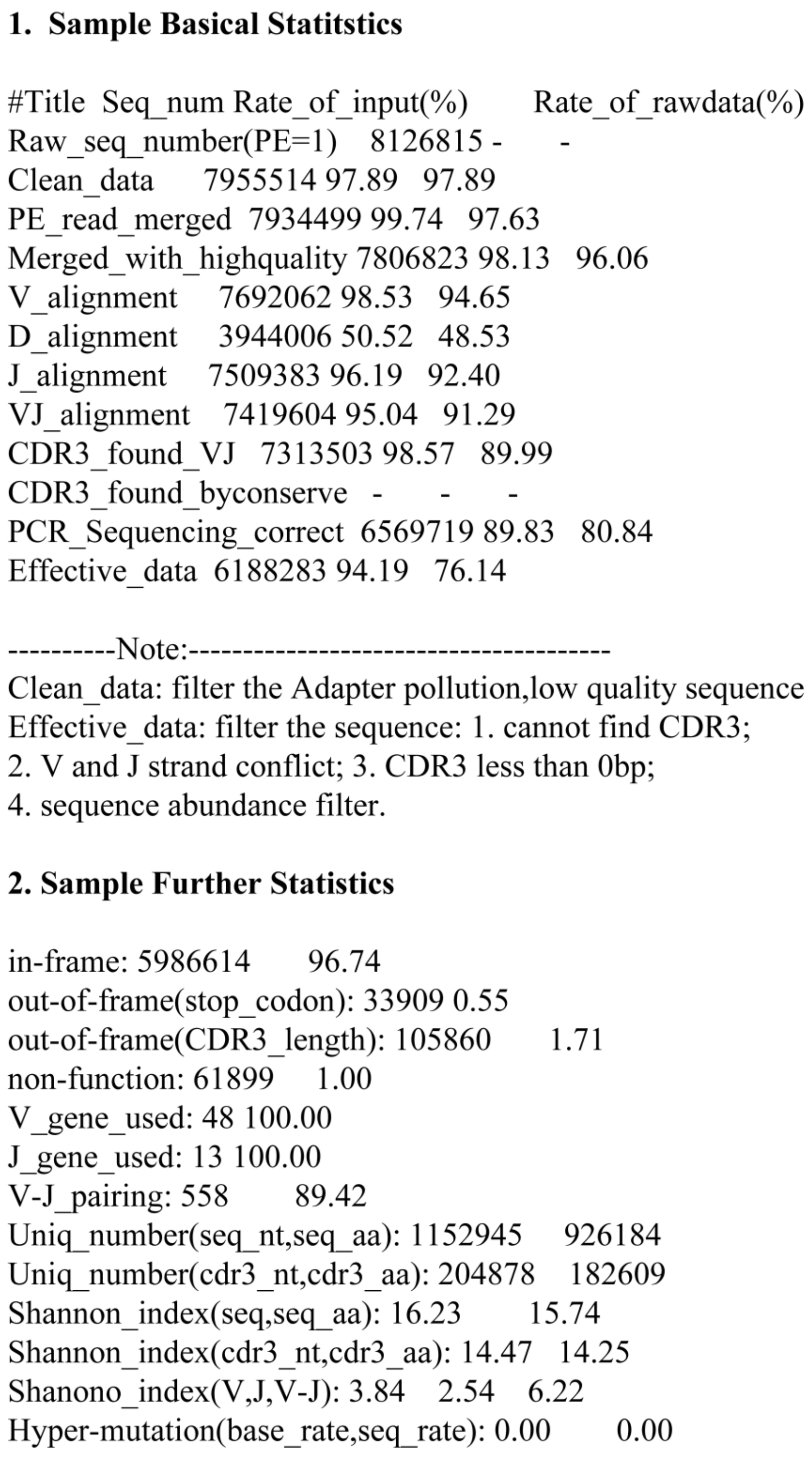

1. Sample Basical Statitstics

```
#Title  Seq_num Rate_of_input(%)        Rate_of_rawdata(%)
Raw_seq_number(PE=1)    8126815 -       -
Clean_data      7955514 97.89   97.89
PE_read_merged  7934499 99.74   97.63
Merged_with_highquality 7806823 98.13   96.06
V_alignment     7692062 98.53   94.65
D_alignment     3944006 50.52   48.53
J_alignment     7509383 96.19   92.40
VJ_alignment    7419604 95.04   91.29
CDR3_found_VJ   7313503 98.57   89.99
CDR3_found_byconserve   -       -       -
PCR_Sequencing_correct  6569719 89.83   80.84
Effective_data  6188283 94.19   76.14

----------Note:--------------------------------------
Clean_data: filter the Adapter pollution,low quality sequence
Effective_data: filter the sequence: 1. cannot find CDR3;
2. V and J strand conflict; 3. CDR3 less than 0bp;
4. sequence abundance filter.
```

2. Sample Further Statistics

```
in-frame: 5986614       96.74
out-of-frame(stop_codon): 33909 0.55
out-of-frame(CDR3_length): 105860       1.71
non-function: 61899     1.00
V_gene_used: 48 100.00
J_gene_used: 13 100.00
V-J_pairing: 558        89.42
Uniq_number(seq_nt,seq_aa): 1152945     926184
Uniq_number(cdr3_nt,cdr3_aa): 204878    182609
Shannon_index(seq,seq_aa): 16.23        15.74
Shannon_index(cdr3_nt,cdr3_aa): 14.47   14.25
Shanono_index(V,J,V-J): 3.84    2.54    6.22
Hyper-mutation(base_rate,seq_rate): 0.00        0.00
```

Figure S3. Outputs of IMonitor, H-B-01 as an example. Sample basic statistics show the data procedure, from raw data to effecive data, such as paired-end reads merged, V(D)J alignment rate. Sample further statistics, show the multiple statistics based on effective data.