Data Processing

- raw reads
  - Illumina pass filter
- filtered raw reads
  - Fastqc
  - Read trimming (if required)
- filtered (and trimmed) raw reads
  - Tophat
- bam files
  - HT-Seq

Mapping to Genomes

- raw count table
  - Table S2 (T. cruzi)
  - Table S3 (Human)

Data assessment and normalization

- filtered count table
  - $R\ Qnorm$
  - R log2
  - hierarchical clustering
    - Figure S7
  - PCA
    - Figure 1C
- normalized count table
  - R Deseq size factor
  - R log2
  - global gene expression profiles
    - Figures 1B, 2
  - Pearson correlation - heatmap and standardized median correlation

DE Analysis

- log2(normalized counttable) without outlier samples
  - R dist
  - R makeSVD
  - R boxplot
  - R cor

- statistical modeling with batch factor
  - voomMod
- subtraction of control from infected samples
  - statistical modeling with batch factor
    - T. cruzi
  - limma-ImFit
  - limma-eBayes

Enrichment Analyses

- Collapse paralogues (T. cruzi only)
- Gene Ontology Enrichment
  - Table S5A (T. cruzi)
  - Table S8 (T. cruzi)
  - Table S9 (Human)
  - K-means clustering
    - Table S7
    - Table S10 (T. cruzi)
    - Table S11 (Human)

NGCB Sequence Read Archive (SRA)
Bioprojects: PRJNA251582 (human);
PRJNA251583 (T. cruzi)
Accession numbers listed in Table S1