Description of Data Objects

- R Objects:
  - Core Data Files
    - All_QC.rds
      - This file contains all quality control information for each cell from STAR and RSEM metrics.
    - ASE.bygene.rds
      - This file contains the allele specific expression of each cell by gene determined by number of reads covering heterozygous SNPs using ASEreadcounter.
    - TPM.nolim.rds
      - The file contains expression (Transcripts Per Million) levels of each cell for all genes using RSEM.
    - TPM.bygene.rds
      - This file contains expression (Transcripts Per Million) levels of each cell for a subset of reliable genes using RSEM.
  - Meta Data and Annotations
    - controlSampleIDs.rds
      - This file contains a list of all control cells IDs.
    - geneRanges_Nikos.rds
      - This file contains annotations of the genomic ranges of each gene.
    - grouped_control_aneuploidies.rds
      - This file contains a list of manually selected cells that demonstrate control aneuploid chromosomes.
    - armRanges.rds
      - This file contains annotations of the genomic ranges of each chromosome arm.
- CSVs:
  - annotation_list.csv
    - This file contains experimental meta data and annotations for each cell.
  - centromeres.csv
    - This file contains annotations of genomic loci of centromeres.
  - manual_review_results.csv
    - This file is an annotation list that records our manual interpretations of each cell after analysis. This file is used only for summary visualizations.
- VCF and DAT files:
  - RPE-1.hets.chr1-X.BA.SNPs_only.recode.vcf.gz
    - This file is a list of heterozygous SNPs.
  - RPE1_Haplotype_update.dat
    - This file is a list of haplotype assignments for each heterozygous SNP.
- Samples
  - Directory of txt files with the IDs of raw sequencing files that are read into the preprocessing pipeline.