Goal: formalize naming conventions so the metadata labels clearly and succinctly describe key parameters of each CRISPR screen experiment analysis.

Key areas to update:

* Last name of main experimenter
* Date of analysis performed or date of screen?
* Virus or viral replicon (ex: DENV2 vs DENVrep)
* Replicate - single vs treated as biological replicates? R1, R2, R3…or R123
* Should prefix somehow denote low vs high? Low vs unsorted?
* Name output analysis files with enough detail so the files aren’t override when you run new similar analysis such as low vs high, low vs unsorted
* Compatible with MAGECK RRA and VISPR?
  + Any new metadata for VISPR?
* How does this interface with the virus host factor database?

Action items:

Prioritize documenting metadata instead of editing csv fields

Csv file

[Analysis] contrast

[Analysis] reps included (R1, R2, or R1R2)

[Analysis] Optional extra analysis field (ex: dates of analysis + description of analysis “NT guides removed”)

Multiple rows:

1.[File] Last name (first author)

2.[File] Last name (last author)

3.[File] Virus name/type

4.[File] Host name

5.[File] Treatment

6.[File] Library

7.[File] Split (Full/splitA/splitB)

8.[File] Rep

Example fastq file name ( fields 1-8):

Cheng\_Kistler\_EBOV\_A549\_Infected\_Brunello\_Full\_rep1.fastq.gz

Analysis folder name ( fields 1-7, 9-10):