

12 cases with TOC Link to HDFS File 3 normals  
5 Abnormals  
 $\nwarrow$   $\nearrow$  not stochastic

Spade Gating  $\rightarrow$  **Maturing Erythroid Events**

for  $\forall$  Antigens:

Select [Backbone] + [antigen]  $\Rightarrow$  Pseudotube

Maturing Erythroid Events [case] [Pseudotube]  $\rightarrow$  K-means Clustering  $\Rightarrow$   $\left( \frac{\text{Clusters}}{\text{Centroid, num Events}} \right)$  Normalize this  
 $\downarrow$   
 $N = 40$

Take EMD of Clusters [case<sub>i</sub>] & Cluster [case<sub>j</sub>]:

Distance Mtx  $\langle 40 \times 40 \rangle$  = pairwise distance of centroids.

$H_i$  = normalize numerals

$H_j$  = normalized num events

$EMD(D, H_i, H_j) = \text{Score}(\text{case}_i, \text{case}_j)$ : This is the similarity score  
 (High  $\rightarrow$  dis-similar)  
 (Low  $\rightarrow$  similar)

For each antigen we should have 15 pairs of differences.

Implementation:

- ① (Case<sub>i</sub>, Case<sub>j</sub>) pairs
  - ② Erythroid Extraction  $\times 2$
  - ③ Pick Antigen/Make Backbone List.
  - ④ Cluster on Backbone List
  - ③ Cluster on  $\forall$  antigens
  - ④ Select Backbone List
- } Will need a clustering program or function
- ⑤ Take EMD of Backbone clusters
  - ⑥