

what is the negative control?-- random group of cells
what about bat-specific or human-specific sequences?

B

Differential k-mer

- Not in genome assembly
- Unannotated gene
- Non-orthologous gene
- Known ortholog

```

graph TD
    A[Interesting  
k-mer] --> B[Not in genome  
assembly]
    A --> C[In reference  
genome]
    C --> D[Not in a gene]
    C --> E[In a gene]
    E --> F[Unknown  
homolog]
    E --> G[Known  
homolog]

```

D

Phylogenetic tree comparing Bat and Human sequences. The tree shows two main branches: Bat (left) and Human (right). The Bat branch includes sequences: Utah, Rhode, Rhode, Texas, Maine, Virginia, Kansas, Arizona, Maryland, and Missouri. The Human branch includes sequences: Utah, Iceland, Iceland, Korea, Vietnam, Vietnam, Virginia, Kansas, Kansas, Illinois, Illinois, and Maryland. The scale bar at the bottom ranges from 0.0 to 3.0 for Bat and 0 to 5 for Human.

****Hypothetical data****