

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

****Hypothetical data****

```

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

****Hypothetical data****