

KEY INSIGHTS

What "Coverage Loss" Means:

Each sequencing error affects a LOCAL window of k-mers:

- k=21: error affects 21 k-mers
- k=41: error affects 41 k-mers

You don't lose entire reads!  
You lose CLASSIFICATION ABILITY in that local genomic region.

Why k=21 is Better:

ARMS markers, k=21:  
~22M lost k-mers/Mb

ARMS markers, k=41:  
~48M lost k-mers/Mb

k=41 loses 2× MORE coverage per megabase!

Practical Impact:

With k=21: ~81% coverage retained  
With k=41: ~66% coverage retained

For a 10 Mb region:  
k=21: 8.1 Mb classifiable ✓  
k=41: 6.6 Mb classifiable ✗

You lose 1.5 Mb more with k=41!

Coverage Loss Analysis: Error-Affected K-mers per Megabase  
The Correct Metric for ONT Sequencing Performance

