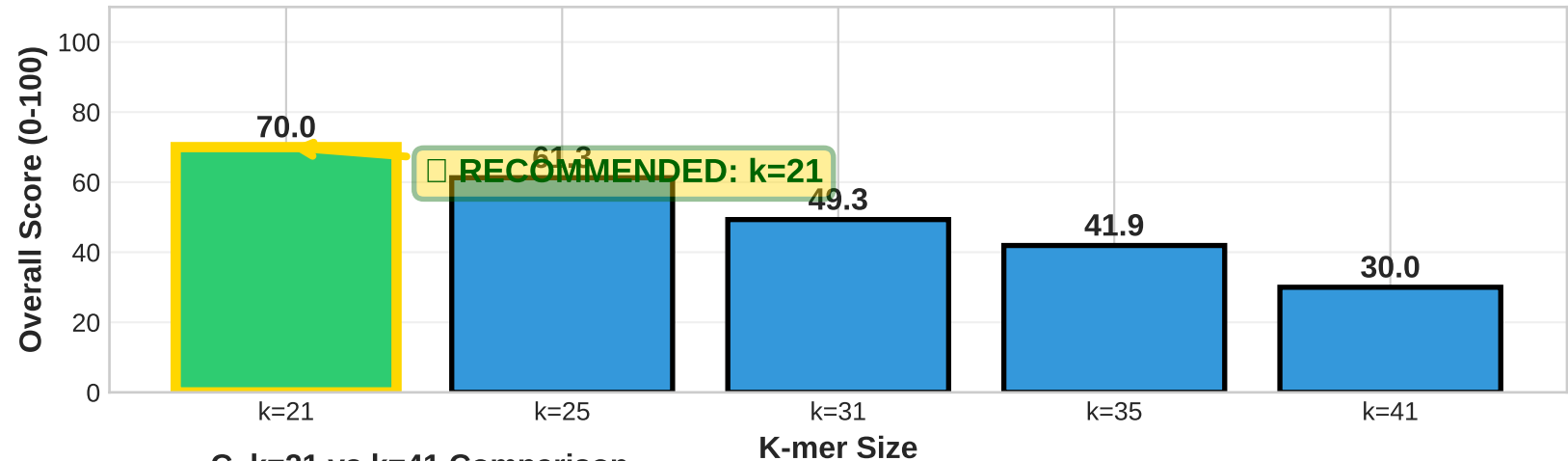
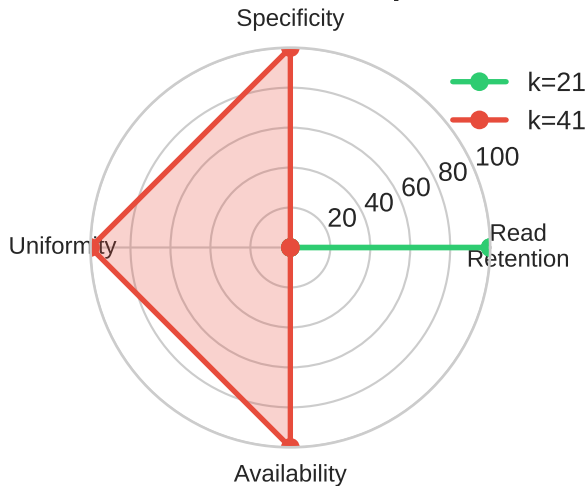


# Comprehensive K-mer Size Evaluation & Recommendation

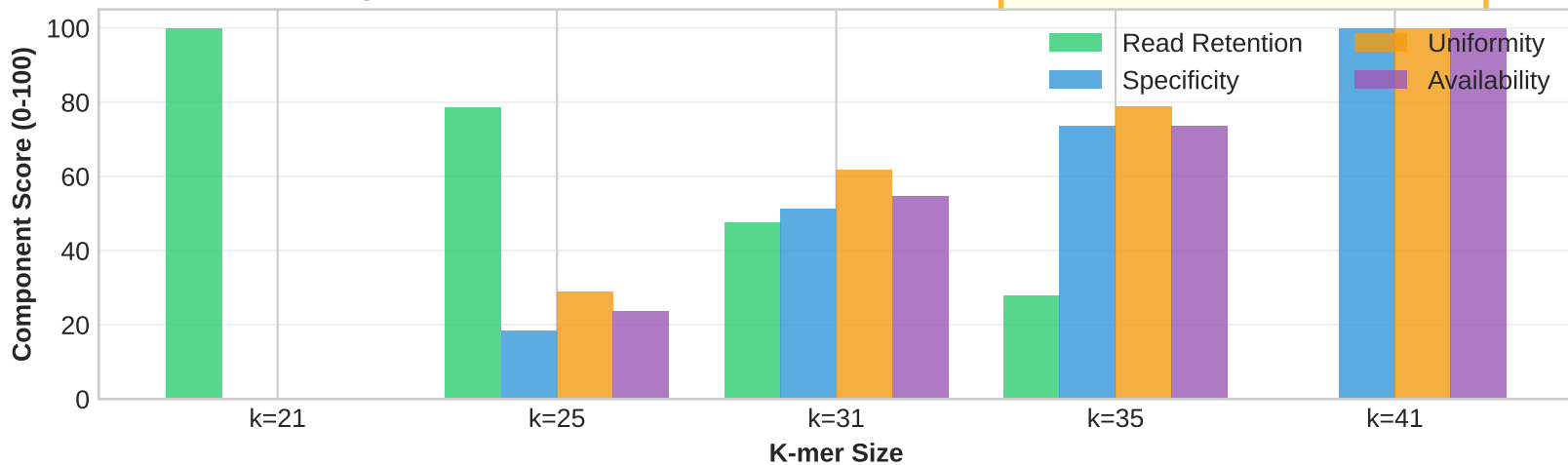
A. Overall Performance Score ■ FINAL RECOMMENDATION



C. k=21 vs k=41 Comparison



D. Individual Component Scores



E. Complete Decision Matrix

| K-mer Size | Usable Reads (%) | False Pos. Rate (%) | Marker Count | Uniformity CV (%) | Overall Score | Rank |
|------------|------------------|---------------------|--------------|-------------------|---------------|------|
| k=21       | 81.10            | 0.1627              | 1,756,428    | 38.0              | 70.0          | #5   |
| k=25       | 77.93            | 0.1527              | 2,025,001    | 35.8              | 61.3          | #4   |
| k=31       | 73.36            | 0.1348              | 2,378,600    | 33.2              | 49.3          | #3   |
| k=35       | 70.47            | 0.1226              | 2,594,793    | 31.9              | 41.9          | #2   |
| k=41       | 66.34            | 0.1082              | 2,896,034    | 30.3              | 30.0          | ❑ #1 |

## SCORING CRITERIA

Optimized for ONT sequencing

- ❑ Read Retention: 70%  
\*\*\* MOST CRITICAL  
81% vs 66% usable reads  
= 15% practical difference!
- ❑ Specificity: 20%  
All <0.2% (excellent!)  
Minimal practical impact
- ❑ Uniformity: 5%  
All reasonably uniform
- ❑ Availability: 5%  
All have enough markers

★ Overall Score: Optimized for maximum read retention