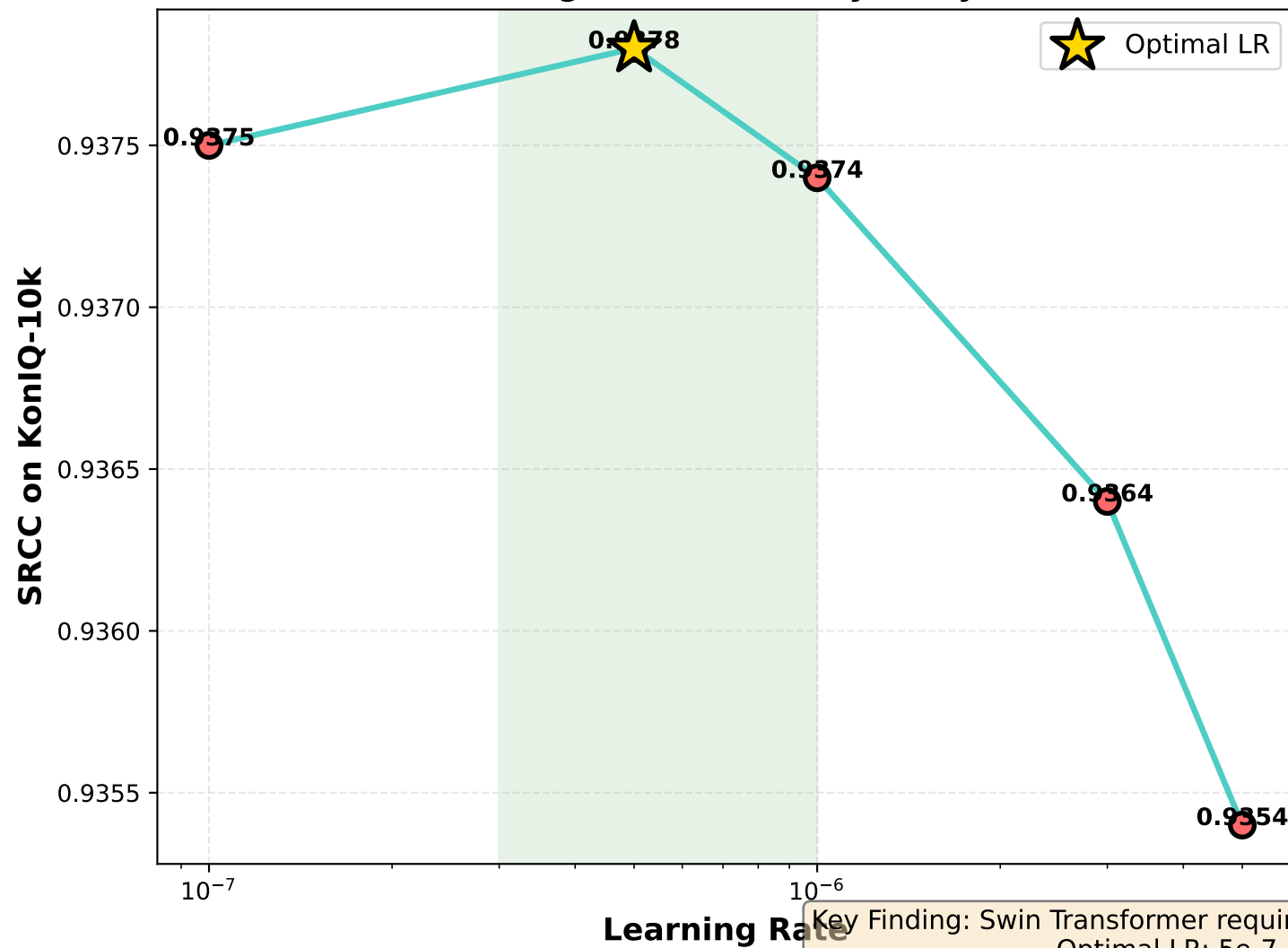
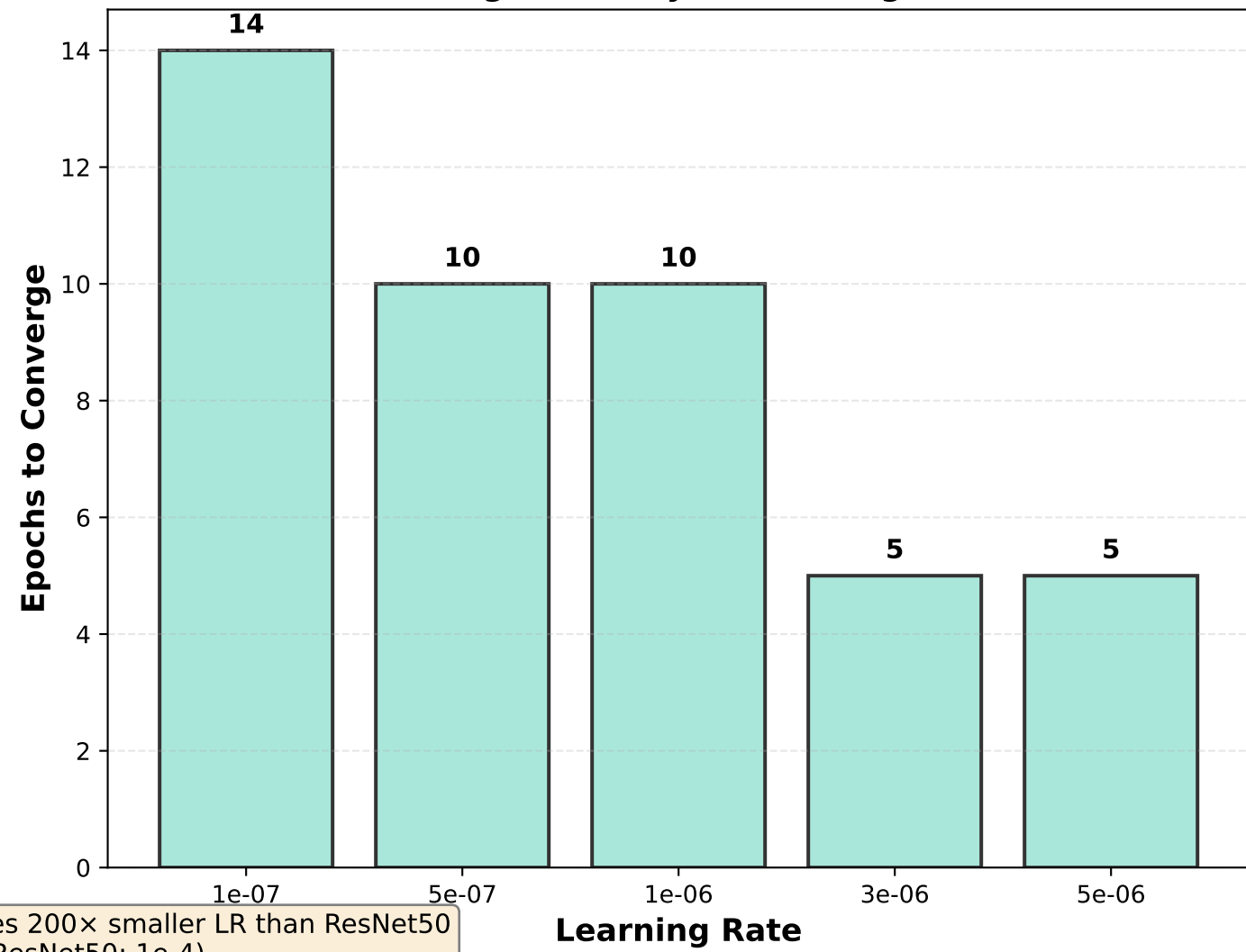


### Learning Rate Sensitivity Analysis



### Training Efficiency vs Learning Rate



Key Finding: Swin Transformer requires 200× smaller LR than ResNet50  
Optimal LR: 5e-7 (ResNet50: 1e-4)