

Suffix Arrays

- Even though Suffix Trees are $O(n)$ space, the constant hidden by the big-Oh notation is somewhat “big”: ≈ 20 bytes / character in good implementations.
- If you have a 10Gb genome, 20 bytes / character = 200Gb to store your suffix tree. “Linear” but large.
- Suffix arrays are a more efficient way to store the suffixes that can do most of what suffix trees can do, but just a bit slower.
- Slight space vs. time tradeoff.