**Статья:** 2003 - K. Malde, E. Coward, I. Jonassen - Fast sequence clustering using a sufﬁx array algorithm

Для активного участия в семинаре следует:

* владеть основными понятиями, которые используют, когда говорят о подстроках: подстрока, суфикс, префикс и т.п. (изучаем [статью на Википедии](http://en.wikipedia.org/wiki/Substring) и убеждаемся, что всё понятно);
* узнать, что такое [range minimum query](http://en.wikipedia.org/wiki/Range_Minimum_Query);
* посмотреть на [суффиксное дерево](http://www.google.com/url?q=http%3A%2F%2Fen.wikipedia.org%2Fwiki%2FSuffix_tree&sa=D&sntz=1&usg=AFQjCNEpcLfb_Itp5eXAYLGTEyLpudQ-og) в задаче “[Finding the Longest Multiple Repeat](http://rosalind.info/problems/lrep/)”;
* (опционально) узнать о суффиксном дереве больше, например, из [этой презентации](http://www.cs.ucf.edu/~shzhang/Combio12/lec3.pdf);
* (опционально) тем, кто уже знает про героя статьи -- суффиксный массив, рекомендована к прочтению [статья о его построении за O(n)](http://www.cs.helsinki.fi/u/tpkarkka/publications/icalp03.pdf).