

СИСТЕМА ХРАНЕНИЯ ПРОФИЛЕЙ ЭЛЕКТРОСТАТИЧЕСКОГО ПОТЕНЦИАЛА ГЕНОМНОЙ ДНК

Темлякова Е.А., Дзелядин Т.Р., Камзолова С.Г., Сорокин А.А.

Институт Биофизики Клетки РАН, Россия, 142290, Пущино, Институтская ул. 3, Тел.: (4967)739319, E-mail: evgenia.templ@gmail.com

Расшифровка полного текста геномов большого количества бактерий и развитие высокопроизводительных экспериментальных методов привело к быстрому росту количества информации о транскрипционных факторах, местах их связывания и взаимном влиянии. В этих условиях для анализа влияния физических свойств ДНК, например электростатического потенциала, на скорость инициации транскрипции необходимо грамотно организовывать все имеющиеся данные, как биологические, так и физические в общую базу данных для более эффективной работы.

В данной работе нами представлена база данных, спроектированная для хранения профилей электростатического потенциала вдоль двойной спирали ДНК, и продемонстрировано ее использование для хранения, поиска и анализа промоторных последовательностей *E.coli*. Отличительным свойством предложенной базы данных является то, что весь профиль хранится как единый объект, который с точки зрения СУБД полностью подобен строке или числу. Такие объекты СУБД может сравнивать друг с другом и осуществлять быструю выборку на основании индексов [1], [2]. В базу данных загружена информация о 1227 известных промоторах *E.coli*. Для каждого промотора сохранена нуклеотидная последовательность, а также вычислен и загружен в базу профиль электростатического потенциала промоторной ДНК [3]. Кроме того, каждый промотор связан с генами, транскрипцию которых он регулирует, а также с записями о сайтах посадки транскрипционных факторов, влияющих на функционирование промотора.

База данных доступна по адресу: <http://promodel.icb.psn.ru/eldna>.

Работа поддержана грантом РФФИ №11-04-01436-а.

Литература.

1. Hellerstein J., Naughton J., Pfeffer A., Generalized search trees for database systems // *IN PROC. 21 ST INTERNATIONAL CONFERENCE ON VLDB*, 1995. Стр.562-573.
2. Sorokin A., Selkov G., Goryanin I., A user-defined data type for the storage of time series data allowing efficient similarity screening // *European Journal of Pharmaceutical Sciences*, **46(4)**, 2012. Стр.272-274.
3. R.V.Polozov, T.R.Dzhelyadin, A.A.Sorokin, N.N.Ivanova, V.S.Sivozhelezov, S.G.Kamzolova. Electrostatic potentials of DNA. Comparative analysis of promoter and nonpromoter nucleotide sequences// *J. Biomol. Struct. Dyn.* **16(6)**, 1999. Стр.1135-1143.