

ОПТИМИЗАЦИЯ ДИЗАЙНА ДНК-СЕНСОРОВ ДЛЯ ДЕТЕКЦИИ НУКЛЕИНОВЫХ КИСЛОТ ПРИ ПОМОЩИ МЕТОДА АДДИТИВНОЙ РЕГРЕССИИ

Брюккель А.Г. (ИТМО), Горбенко Д.А. (ИТМО), Баранникова М.В. (МФТИ), Шепелев И.А.
(АГНИ)

Научный руководитель – кандидат биологических наук, Кошель Е.И. (ИТМО)

Введение

Методы детекции нуклеиновых кислот (НК) на основе гибридизации широко используются в диагностике генетических заболеваний, детекции патогенов. Бинарные ДНК-зонды позволяют обнаруживать целевую НК с высокой чувствительностью и селективностью. Они превосходят ПЦР и ДНК-чипы по простоте анализа и не требуют дорогих реагентов и обученных специалистов [1]. Среди ДНК-зондов наиболее широкое распространение получили бинарные дезоксирибозимы (BiDz) [2], способные к детекции от 50-100 пМ аналита. Вместе с тем для них характерна сложность и высокая себестоимость процесса дизайна. В настоящее время для создания сенсора необходимо от 4 до 10 дизайнов на одну последовательность аналита. Учитывая затраты на доставку олигонуклеотидов и их тестирование, выбор такого сенсора может потребовать до 100 тысяч рублей и от 3 до 9 месяцев работы (в некоторых случаях больше - до 1 года).

Основная часть

Целью исследования было сокращение экспериментов отработки дизайнов BiDz *in vitro*. Был проведен анализ данных, полученных в одной лаборатории и разработана предиктивная математическая модель, которая описывает взаимосвязь численных параметров, учитываемых в дизайне BiDz (изменение свободной энергии Гиббса (ΔG), длины аналит-связывающих плеч (L_a и L_b), длины анализируемых последовательностей (L_{tar}), температуры плавления аналит-связывающих плеч ($T_{пл}$), температуры инкубации ($T_{и}$) и т.п.) Выбор инструмента для создания математической модели был ограничен в связи с очень малым размером выборки (50 дизайнов BiDz). Для решения задачи был выбран алгоритм аддитивной регрессии SISSO, сочетающий в себе “compressed sensing” и символьную регрессию [3]. По результатам анализа данных созданы две математические модели, учитывающие следующие переменные: ΔG , L_a и L_b , L_{tar} и $T_{пл}$ для аналит-связывающих плеч. Входные данные включали длины L_a и L_b , ΔG для аналит связывающих плеч, а также L_{tar} и ΔG для формирования их комплексов. Выходным параметром было ΔG в зависимости от соотношения L_a к L_b . Согласно прогнозу модели, локальные минимумы ΔG достигнуты при соотношениях длин аналит-связывающих плеч 18 (L_a) к 18 (L_b), 36 (L_a) и 13 (L_b), а также 35 (L_a) и 7 (L_b). Эти точки выбраны для тестирования, а альтернативные соотношения длин 18 (L_a) и 13 (L_b), 26 (L_a) и 10 (L_b) использованы в качестве контрольных (выбраны вручную). Для верификации предсказаний модели были созданы 5 дизайнов. В качестве аналита использована последовательность вируса простого герпеса человека (ген UL9). Полученные пары протестированы *in vitro* согласно стандартным операционным процедурам [4,5]. Для оценки пригодности сенсоров использованы следующие показатели: отношение сигнала к фону (присутствие аналита / отсутствие аналита) после инкубации 1 и 3 часа с пороговым соотношением 1,5, а также лимит детекции после инкубации 1 час с пороговым значением 100 пМ [6].

После инкубации в течение 1 часа сигнал в присутствии аналита для пары 18 (L_a) к 18 (L_b) был в 5 раз выше, чем в пустом эксперименте, для пары 18 (L_a) и 13 (L_b) - в 1,7 раза выше. Для пары 26 (L_a) и 10 (L_b) отношение сигнал/фон составило 1,9, а для пары 36 (L_a) и 13 (L_b) - 5,12. Для пары 35 (L_a) и 7 (L_b) отношение сигнал/фон составило 1,4. После 3-часовой инкубации отношение сигнал/фон увеличилось и составило 6,1 для пары 18 (L_a) к 18 (L_b) и 2,5 для пары 18 (L_a) и 13 (L_b). Пары аналит-связывающих плечей с разницей в 26 (L_a) и 10 (L_b) также показали отношение сигнал/фон 2,5, в то время как сенсор 36 (L_a) и 13 (L_b) достиг значения 5,6. Сенсор с парой 35 (L_a) и 7 (L_b) достиг приемлемого значения 1,8 после 3-часовой инкубации. Лимит детекции был определен для пар 18 (L_a) к 18 (L_b), 36 (L_a) и 13 (L_b), 18 (L_a) и 13 (L_b), 26 (L_a) и 10 (L_b). Соотношение 35 (L_a) и 7 (L_b) не использовалось для лимита детекции из-за недостаточной разницы сигнала к фону. Лимит детекции для пар с минимальными значениями ΔG был определен как 87 пМ для 18 (L_a) к 18 (L_b) и 28 пМ для 36 (L_a) и 13 (L_b). Для пар с соотношением 18 (L_a) и 13 (L_b) и 26 (L_a) и 10 (L_b) лимит детекции был значительно ниже - 670 пМ и 840 пМ, соответственно.

Выводы

Произведена сборка данных по ранее созданным в лаборатории дизайнам ViDz. Проанализированы такие численные параметры, как ΔG , L_a и L_b , L_{tag} и $T_{пл}$. Найдены корреляционные функции для ΔG и L_a к L_b , а также для ΔG и $T_{пл}$ аналит-связывающих плечей. Применена модель аддитивной регрессии на основе алгоритма SISSO для описания локальных минимумов ΔG при различных длинах L_a и L_b . Проведено сравнение модельных комбинаций плеч с парами, подобранными вручную. Результаты тестирования *in vitro* показали преимущество модельных значений по отношению сигнал/фон и пределу детекции. В среднем, значение сигнала к фону после 3 часов инкубации для модельных пар составило 6, тогда как немодельные показали соотношение 2,5. Лимит детекции для модельных пар был зафиксирован в пределах 100 пМ (87 пМ для пары 18 (L_a) и 18 (L_b) и 28 пМ 36 (L_a) и 13 (L_b)), тогда как для немодельных пар этот параметр составил 600-900 пМ (670 пМ для 18 (L_a) и 13 (L_b) и 840 пМ для 26 (L_a) и 10 (L_b)).

Список литературы

1. Yuwen L, Zhang S, Chao J. Recent Advances in DNA Nanotechnology-Enabled Biosensors for Virus Detection// Biosensors (Basel) – 2023 – Т.13. – С.822.
2. Kolpashchikov DM. A binary deoxyribozyme for nucleic acid analysis. // Chem Biochem. –2007– Т.8., С. 2039-2042
3. Ouyang, R., Curtarolo, S., Ahmetcik, E., Scheffler, M., Ghiringhelli, L. M. SISSO: A compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates // Phys. Review Materials – 2018– Т. 2. – С. 083802
4. Ateiah M., Standard operating procedure “Initial testing of Binary deoxyribozymes and DNA nanomachines”, SOP – code/XX/2023, https://docs.google.com/document/d/16f2h_ihMxyNWki-5Qp-mP3bQO-6FHSM0/edit (accessed 2024-02-01)
5. Ateiah M. Standard operating procedure, “Limit of detection” SOP – code/XX/2023, <https://docs.google.com/document/d/1g-Ld3hQDPfR78KOLbdyL13AHkIN7FgQY/edit> (accessed 2024-02-01)
6. Stancescu M.; Fedotova T.A.; Hooyberghs J.; Balaeff A.; Kolpashchikov D.M. Non-equilibrium Hybridization Enables Discrimination of a Point Mutation within 5-40 °C. // J Am Chem Soc. – 2016 –Т. 138 – С. 13465-13468.