

**LC-MS/MS идентификация в сыворотке крови человека пептидных  
фрагментов, транслированных с lncRNA**

**Павлович Полина Викентьевна**

*Студент (бакалавр)*

Московский физико-технический институт, Москва, Россия

*E-mail: pavlovich@phystech.edu*

Полный геном человека был секвенирован в 2003 году в рамках международного проекта «Геном человека» (International Human Genome Sequencing Consortium, 2001). Однако анализ отдельных его участков все еще не закончен: появляются новые гены. Было обнаружено, что в клетках человека присутствуют транскрипты, покрывающие примерно 70% генома, и многие из них не являются известными тРНК, рРНК или мРНК (Laurent, 2015). Среди них выделяют отдельный класс длинных некодирующих РНК (lncRNA) - это транскрипты длиннее 200 нуклеотидов, с которых, как известно на сегодняшний день, трансляция не происходит. Однако не все «lncRNA» являются действительно белок некодирующими. Например, данные масс-спектрометрического анализа выявили присутствие 250 новых пептидов мыши, кодируемых предположительно lncRNAs (Prabakaran, 2014). Таким образом можно предположить, что часть lncRNA человека может являться транскриптами ранее не известных генов.

В нашей работе используются масс-спектрометр Sciex TripleTOF 5600+ для исследования пептидома сыворотки крови человека. Идентификация данных масс-спектрометрического исследования проводится против объединенной базы данных (БД), состоящей из БД UniProt по человеку и БД по длинным некодирующим РНК, LNCipedia. LNCipedia представляет собой интегрированную базу хорошо аннотированных транскриптов из различных источников (high confidence set, 79769 transcripts).

Проверку вероятности трансляции транскриптов из LNCipedia проводили с помощью различных биоинформатических алгоритмов: blastx, hmmscan, CPAT, CNCI, CPC. Кроме того, была проанализирована встречаемость транскриптов из LNCipedia в двух клеточных локализациях: ядре и цитоплазме - для 18 клеточных линий по данным ENCODE.

Анализ шести образцов сыворотки крови здоровых доноров позволил обнаружилось 12 пептидов, которые не являются фрагментами известных белков человека, но транскрипты которых содержатся в БД LNCipedia. Анализ соответствующих транскриптов показал отсутствие их кодирующей способности. В докладе будут обсуждаться планируемые шаги по валидации идентифицированных пептидов с помощью биоинформатических и химических методов, в том числе при помощи синтетических аналогов.

**Источники и литература**

- 1) <http://www.nature.com/nature/journal/v409/n6822/pdf/409860a0.pdf>
- 2) <http://www.sciencedirect.com/science/article/pii/S0168952515000542>
- 3) <http://www.nature.com/ncomms/2014/141118/ncomms6429/full/ncomms6429.html>

**Слова благодарности**

Выражаю свою благодарность Ивановой О.М., Капранову Ф.В., Говоруно В.М.