

## ОСОБЕННОСТИ СТРУКТУРЫ И ЭКСПРЕССИИ ГЕНА СФИНГОМИЕЛИНСИНТАЗЫ 1 (*SGMS1*) ЧЕЛОВЕКА

© И. Б. Филиппенков,<sup>1</sup> Л. В. Дергунова, А. В. Рожкова,  
О. Ю. Сударкина, С. А. Лимборская

Институт молекулярной генетики РАН, Москва, 123182;

<sup>1</sup> электронный адрес: [Filippenkov@img.ras.ru](mailto:Filippenkov@img.ras.ru)

Представлен обзор результатов исследования структурно-функциональной организации гена сфингомиелинсинтазы 1 (*SGMS1*) человека в Отделе молекулярных основ генетики человека Института молекулярной генетики РАН. Ген *SGMS1* кодирует жизненно важный фермент, обеспечивающий синтез сфингомиелина и диацилглицерина из фосфатидилхолина и церамида, что определяет его участие в регуляции внутриклеточного везикулярного транспорта, метаболизме холестерина, пролиферации клеток, апоптозе и других значимых процессах. Наши исследования показали, что ген *SGMS1* расположен на хромосоме 10, имеет размер 320 тыс. п.н. и содержит более 20 экзонов. Детальное исследование структуры гена *SGMS1* позволило нам выявить множество его транскриптов. Были обнаружены изоформы мРНК, отличающиеся участком 5'-нетранслируемой области (5'-НТО) и кодирующие полноразмерный белок, а также транскрипты, возникающие в результате альтернативной комбинации экзонов, затрагивающих кодирующую область гена и 3'-НТО. В числе продуктов гена *SGMS1* нами обнаружены новые транскрипты — циклические РНК, которые содержат преимущественно последовательности мультиэкзонной 5'-НТО гена. Они консервативны и экспрессируются преимущественно в мозге. Циклические РНК гена *SGMS1* имеют большое количество сайтов связывания микроРНК, что может определить функциональное значение этих молекул. В работе описываются самые последние данные о структурно-функциональной организации и особенностях экспрессии гена *SGMS1* человека.

Ключевые слова: сфингомиелинсинтаза 1, структура гена *SGMS1*, альтернативные транскрипты, циклические РНК, тканеспецифический характер экспрессии, микроРНК

Принятые сокращения: НТО — нетранслируемая область, п. н. — пара нуклеотидов, ПЦР — полимеразная цепная реакция, SMS1 — фермент сфингомиелинсинтаза 1, *SGMS1* — ген сфингомиелинсинтазы 1 человека.

Более 20 лет тому назад, в конце прошлого тысячелетия, наш коллектив работал в актуальной по тем временам области, а именно в области изучения генов, экспрессирующихся в мозге человека. Исследования были связаны с получением клонотек кДНК на основе полиаденилированной РНК из лобной коры, мозжечка и продолговатого мозга (Буякова и др., 1992; Дергунова и др., 1998, 2003; Vladychenskaya et al., 2002). Одна из таких кДНК была выделена из клонотеки продолговатого мозга человека и получила наименование НМОВ33 (human medulla oblongata, номер клона 33). Она соответствовала мРНК гена, кодирующего неизвестный белок. Этот ген имел размер около 320 тыс. п. н., включал в себя 11 экзонов и был локализован на хромосоме 10 в области 10q11.2 (рис. 1). Нуклеотидная последовательность мРНК этого гена, названного нами МОВ (позднее ему было присвоено номенклатурное название ТМЕМ 23), была зарегистрирована в базе данных ГенБанк под номером BN000143 (Vladychenskaya et al., 2004). Ген МОВ кодировал гипотетический пептид длиной 413 аминокислотных остатков, первичная структура которого полностью совпала с аминокислотной последовательностью фермента сфингомиелинсинтазы 1 (SMS1), идентифици-

рованной вскоре двумя независимыми группами исследователей (Huitema et al., 2004; Yamaoka et al., 2004). Таким образом, было выяснено, что обнаруженный нами ген является геном *SGMS1*.

SMS1 является жизненно важным ферментом. Он катализирует перенос остатка фосфорилхолина с фосфатидилхолина на церамид, приводя к синтезу сфингомиелина и диацилглицерина. Это определяет его участие в регуляции внутриклеточного везикулярного транспорта, метаболизме холестерина, пролиферации клеток, апоптозе и других жизненно важных процессах.

В наших дальнейших исследованиях была уточнена экзон-интронная структура гена *SGMS1* (рис. 2). Ген содержит в своем составе более 20 экзонов и кодирует большое число альтернативных транскриптов (Rozhkova et al., 2011; Filippenkov et al., 2015). По принципу организации все транскрипты гена *SGMS1* можно условно разделить на четыре группы. Транскрипты первой группы кодируют полноразмерный белок SMS1, но различаются структурой 5'-НТО (рис. 3, а). Это достигается использованием альтернативных промоторов (рис. 2) (Рожкова и др., 2015). Наиболее часто во всех тканях человека используется дистальный промотор, который обеспечивает

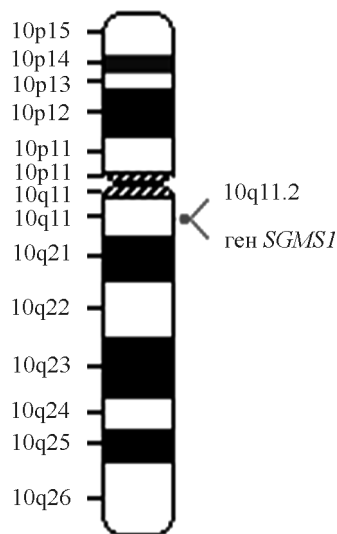


Рис. 1. Хромосома 10 человека с указанием участка, на котором локализован ген сфингомиелинсинтазы 1 (*MOB*, *TMEM23*, *SGMS1*).

синтез транскриптов, включающих в себя первый экзон. Транскрипты с других промоторов, локализованных в интронах гена, представлены значительно слабее.

Транскрипты второй группы укорочены с 3'-конца. Они не содержат экзонов 8—11, но включают в свой состав экзон 7d — участок интрона VII, прилегающий к экзону 7 (рис. 3, б). Их появление в клетке связано, по-видимому, с преждевременным полиаденилированием (рис. 2) (Dergunova et al., 2013). Содержание таких транскриптов в разных тканях очень невелико: максимальный уровень экспрессии не достигает 30 % относительно транскриптов первой группы, однако наблюдается тканеспецифический характер их экспрессии. Транскрипты этой группы не могут обеспечивать синтез полноразмерного белка, но теоретически могли бы определять укороченные варианты. При анализе белковых молекул SMS1 из тканей человека нам не удалось обнаружить укороченный белок (Sudarkina et al., 2015), однако в клеточной системе HeLa генно-инженерная конструкция, содержащая один из укороченных транскриптов, обеспечивала синтез белкового продукта уменьшенного размера (Rozhkova et al., 2011). Мы предполагаем, что синтез укороченных с 3'-конца транскриптов может быть связан с негативной регуляцией содержания фермента SMS1 в тех или иных клеточных структурах.

Транскрипты третьей группы представлены наименее изученными формами, функциональная значимость кото-

рых остается невыясненной (рис. 3, в). Общее для них состоит в том, что между последовательностями экзонов 7 и 8 присутствуют различные участки интрона VII (Rozhkova et al., 2011). Эти интронные вставки содержат преждевременные терминирующие кодоны, что препятствует синтезу полноразмерного белка. Уровень их экспрессии во всех тканях низок (менее 5 % от основного транскрипта).

Особый интерес представляют транскрипты гена *SGMS1* четвертой группы, так как они имеют не линейную, а циклическую структуру (Filippenkov et al., 2015). Циклические формы РНК были выявлены в клетках высших организмов совсем недавно. Они зачастую образуются в белок-кодирующих генах в процессе альтернативного сплайсинга (Salzman et al., 2012, 2013; Jeck et al., 2013; Memczak et al., 2013; Zhang et al., 2013; Guo et al., 2014). Такие транскрипты устойчивы к РНКазе R, которая избирательно разрушает линейные формы РНК (Suzuki et al., 2006). В составе транскриптов гена *SGMS1* циклические РНК были выявлены нами с помощью метода Нозерн-гибридизации суммарной РНК лобной коры, предварительно подвергнутой обработке РНКазой R. Они гибридизовались с зондом, комплементарным 5'-НТО мРНК гена *SGMS1*, и детектировались в области длин 200—500 п. н. Для выяснения структуры данных РНК нами была создана серия противоположно ориентированных ПЦР-праймеров, соответствующих всем экзонам мРНК гена *SGMS1*. С их помощью нам удалось получить последовательности, которые соответствовали 11 циклическим РНК (рис. 4). Выявленные циклические РНК содержат от 1 до 5 экзонов. В некоторых случаях наряду с экзонными в них имеются интронные участки (Filippenkov et al., 2015). Обращает на себя внимание то, что в составе циклических форм чаще всего присутствуют последовательности мультиэкзонной 5'-НТО, которые не участвуют в кодировании основного белкового продукта. Возможно, что основное предназначение этих экзонов как раз и состоит в том, чтобы обеспечивать существование выявленных циклических РНК.

Нами было показано, что циклические РНК гена *SGMS1* обладают консерватизмом. Они экспрессируются в ортологических генах крысы и мыши. Вместе с тем у человека и исследуемых животных они представлены преимущественно в отделах мозга, что может указывать на их функциональное значение в этих тканях (Filippenkov et al., 2015).

Вопрос о функциональной значимости циклических РНК остается до конца не выясненным. Однако имеются данные, показывающие наличие в циклических РНК большого числа сайтов связывания с микроРНК. Известным примером является циклическая РНК CIRs-7 гена CDR1, которая активно экспрессируется в мозге человека

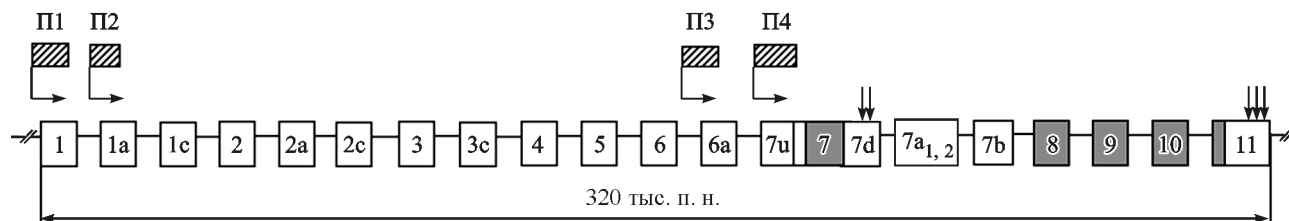


Рис. 2. Экзон-интронная структура гена сфингомиелинсинтазы 1.

Экзоны показаны в виде пронумерованных прямоугольников; интроны даны в виде прямых линий, соединяющих экзоны. Темно-серым цветом отмечены экзоны, кодирующие белок. Промоторные области (P1—P4) обозначены заштрихованными прямоугольниками со стрелками, расположенными над экзонами. Вертикальные стрелки показывают локализацию сайтов полиаденилирования.

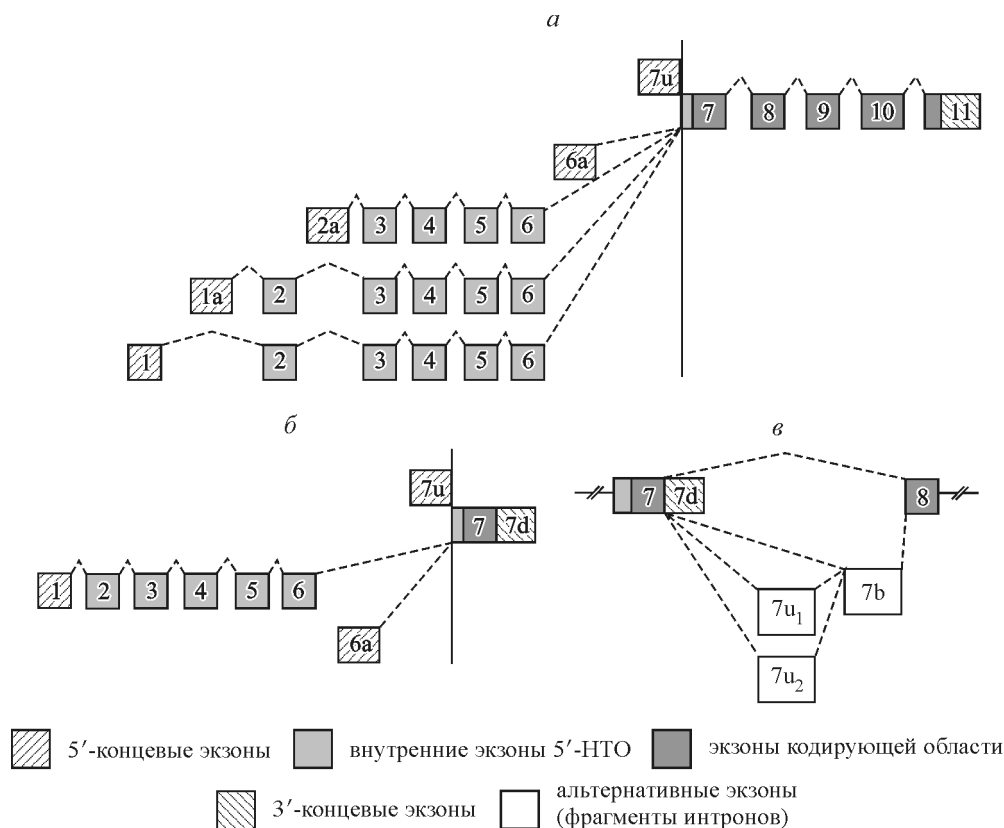


Рис. 3. Схема структурной организации транскриптов гена сфингомиелинсинтазы 1.

*a—в* — экзонная организация транскриптов гена *SGMS1*, различающихся 5'-НТО (*a*), с различной 3'-НТО (*b*) и содержащих участки интрона VII (*в*). Штриховые линии обозначают варианты соединения экзонов в транскриптах.

и мышцы, содержит около 70 сайтов связывания с miR-7 и подавляет ее функцию (Hansen et al., 2011, 2013). С помощью биоинформатического анализа мы проанализировали нуклеотидные последовательности циклических РНК гена *SGMS1* с целью поиска сайтов связывания микроРНК. Оказалось, что наибольшее число сайтов связывания находится внутри экзонов 5—7 (Filippenkov et al., 2015). Можно предположить, что циклические РНК, которые содержат в своем составе данные экзоны, могут как губка связывать на себе значительное количество микроРНК, удаляя их тем самым из клеточного пула. С учетом регуляторных свойств микроРНК это может су-

щественно сказываться на регуляции функционирования гена *SGMS1*.

Для изучения особенностей функционирования гена *SGMS1* на трансляционном уровне нами был разработан метод получения белкового экстракта тканей, обогащенного SMS1 (Сударкина, Дергунова, 2015). Сравнительный анализ экспрессии гена *SGMS1* на трансляционном и транскрипционном уровнях не выявил корреляции между уровнем кодирующих транскриптов и количеством белка в разных тканях человека (Sudarkina et al., 2015). Это может свидетельствовать о наличии сложной системы тканеспецифической регуляции гена *SGMS1*.

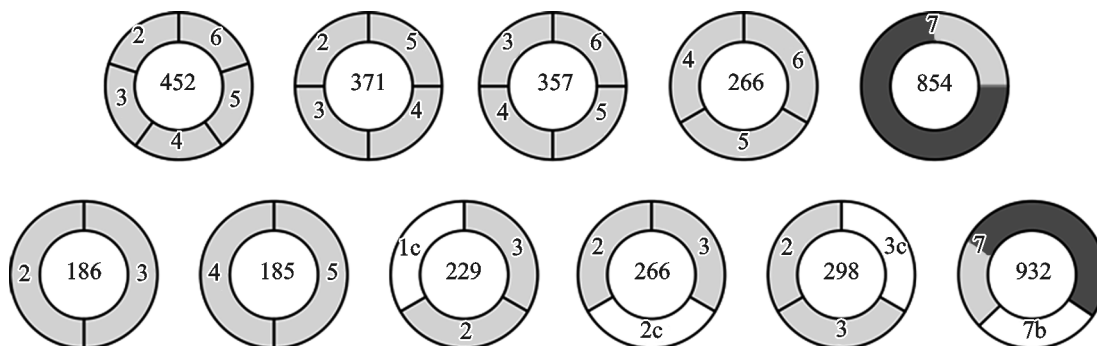


Рис. 4. Структурная организация циклических РНК гена *SGMS1*.

Цветом выделены экзоны, вошедшие в состав циклических РНК: светло-серым — внутренние экзоны 5'-НТО, темно-серым — экзоны кодирующей области, белым — фрагменты интронов. Цифры внутри кольца — номера экзонов, цифры в центре — длина циклических РНК в нуклеотидах.

В заключение необходимо отметить, что изучаемый нами ген *SGMS1* помимо кодирующей последовательности содержит значимую информацию, необходимую для регуляции его экспрессии. В частности, это может быть обусловлено функционированием разнообразных транскриптов гена, в том числе циклической природы. На примере *SGMS1* можно видеть, насколько сложно устроены гены высших организмов. По существу, каждый такой ген представляет собой сложную генетическую структуру, не только обеспечивающую синтез белкового продукта, но и снабженную многоуровневой системой, тонко регулирующей функционирование гена на разных этапах и в различных условиях.

Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (проекты 99-04-48555, 02-04-48809, 05-04-49566 и 11-04-00843), программы РАН «Молекулярная и клеточная биология» и Программы поддержки ведущих научных школ.

### Список литературы

- Бужкова О. И., Баринова Е. И., Дергунова Л. В., Хаспеков Г. Л., Чивилев И. В., Лимборская С. А. 1992. Выделение и анализ мозгоспецифических последовательностей из библиотеки кДНК разных отделов мозга человека. *Генетика*. 28 (5) : 40—46. (Buiakova O. I., Barinova O. I., Dergunova L. V., Khaspekov G. L., Chivilev I. V., Limborskaia S. A. 1992. Isolation and analysis of brain-specific sequences from cDNA libraries for various segments of the human brain. *Genetika*. 28 (5) : 40—46.)
- Дергунова Л. В., Владыченская И. П., Полукарова Л. Г., Раевская Н. М., Леликова Г. П., Лимборская С. А. 1998. Особенности структуры, экспрессии и хромосомная локализация нуклеотидных последовательностей Hmob3 и Hmob33, полученных из клонотеки кДНК продолговатого мозга человека. *Молекуляр. биол.* 32 (2) : 249—254. (Dergunova L. V., Vladychenskaya I. P., Polukarova L. G., Raevskaia N. M., Lelikova G. P., Limborskaia S. A. 1998. Features of the structure, expression and chromosomal mapping of the nucleotide sequences of Hmob3 and Hmob33, obtained from a human medulla oblongata cDNA library. *Mol. Biol. (Moscow)*. 32 (2) : 249—254.)
- Дергунова Л. В., Раевская Н. М., Владыченская И. П., Лимборская С. А. 2003. Структурно-функциональный анализ кДНК последовательностей Hfb1, Hmob3 и Hmob33 как пример исследования мозгоспецифических генов человека. *Молекуляр. биол.* 37 (2) : 315—324. (Dergunova L. V., Raevskaia N. M., Vladychenskaya I. P., Limborskaia S. A. 2003. Structural and functional analyses of the Hfb1, Hmob3 and Hmob33 cDNAs as an example of human brain-specific gene studies. *Mol. Biol. (Moscow)*. 37 (2) : 315—324.)
- Рожкова А. В., Филиппенков И. Б., Сударкина О. Ю., Лимборская С. А., Дергунова Л. В. 2015. Альтернативные промотеры, локализованные в интронах гена *SGMS1*, участвуют в регуляции его экспрессии в тканях человека. *Молекуляр. биол.* 49 (2) : 325—333. (Rozhkova A. V., Filippenkov I. B., Sudarkina O. Yu., Limborska S. A., Dergunova L. V. 2015. Alternative promoters located in *SGMS1* gene introns participate in regulation of its expression in human tissues. *Mol. Biol. (Moscow)*. 49 (2) : 287—294.)
- Сударкина О. Ю., Дергунова Л. В. 2015. Получение белкового экстракта тканей человека, обогащенного сфингомиелинсинтазой I. *Молекуляр. генет. микробиол. вирусол.* 33 (2) : 38—41. (Sudarkina O. Yu., Dergunova L. V. 2015. Preparation of human tissue protein extracts enriched with the sphingomyelin synthase I. *Mol. Gen. Microbiol. Virusol.* 33 (2) : 38—41.)
- Дергунова Л. В., Рожкова А. В., Сударкина О. Ю., Лимборская С. А. 2013. The use of alternative polyadenylation in the tissue-specific regulation of human *SMS1* gene expression. *Mol. Biol. Rep.* 40 : 6685—6690.
- Филиппенков И. Б., Сударкина О. Ю., Лимборская С. А., Дергунова Л. В. 2015. Circular RNA of the human sphingomyelin synthase 1 gene: multiple splice variants, evolutionary conservatism and expression in different tissues. *RNA Biol.* 12 : 1030—1042.
- Guo J. U., Agarwal V., Guo H., Bartel D. P. 2014. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol.* 15 : 409.
- Hansen T. B., Jensen T. I., Clausen B. H., Bramsen J. B., Finisen B., Damgaard C. K., Kjems J. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature*. 495 : 384—388.
- Hansen T. B., Wiklund E. D., Bramsen J. B., Villadsen S. B., Statham A. L., Clark S. J., Kjems J. 2011. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* 30 : 4414—4422.
- Huitema K., van den Dikkenberg J., Brouwers J. F., Holthuis J. C. 2004. Identification of a family of animal sphingomyelin synthases. *EMBO J.* 23 : 33—44.
- Jeck W. R., Sorrentino J. A., Wang K., Slevin M. K., Burd C. E., Liu J., Marzluff W. F., Sharpless N. E. 2013. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA*. 19 : 141—157.
- Memczak S., Jens M., Elefsinioti A. et al. 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*. 495 : 333—338.
- Rozhkova A. V., Dmitrieva V. G., Zhapparova O. N., Sudarkina O. Y., Nadezhdina E. S., Limborska S. A., Dergunova L. V. 2011. Human sphingomyelin synthase 1 gene (*SMS1*): organization, multiple mRNA splice variants and expression in adult tissues. *Gene*. 481 : 65—75.
- Salzman J., Chen R. E., Olsen M. N., Wang P. L., Brown P. O. 2013. Cell-type specific features of circular RNA expression. *PLoS Genet.* 9 : e1003777.
- Salzman J., Gawad C., Wang P. L., Lacayo N., Brown P. O. 2012. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS ONE* 7 : e30733.
- Sudarkina O. Y., Filippenkov I. B., Brodsky I. B., Limborska S. A., Dergunova L. V. 2015. Comparative analysis of sphingomyelin synthase 1 gene expression at the transcriptional and translational levels in human tissues. *Mol. Cell. Biochem.* 406(1—2) : 91—99.
- Suzuki H., Zuo Y., Wang J., Zhang M. Q., Malhotra A., Mayeda A. 2006. Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucl. Acids Res.* 34 : e63.
- Vladychenskaya I. P., Dergunova L. V., Dmitrieva V. G., Limborska S. A. 2004. Human gene *MOB*: structure specification and aspects of transcriptional activity. *Gene*. 338 : 257—265.
- Vladychenskaya I. P., Dergunova L. V., Limborska S. A. 2002. *In vitro* and *in silico* analysis of the predicted human *MOB* gene encoding a phylogenetically conserved transmembrane protein. *Genet. Anal. Biomol. Eng.* 18 : 263—268.
- Yamaoka S., Miyaji M., Kitano T., Umehara H., Okazaki T. 2004. Expression cloning of a human cDNA restoring sphingomyelin synthesis and cell growth in sphingomyelin synthase-defective lymphoid cells. *J. Biol. Chem.* 279 : 18 688—18 693.
- Zhang Y., Zhang X. O., Chen T., Xiang J. F., Yin Q. F., Xing Y. H., Zhu S., Yang L., Chen L. L. 2013. Circular intronic long noncoding RNAs. *Mol. Cell.* 51 : 792—806.

PECULIARITIES OF THE STRUCTURE AND EXPRESSION OF HUMAN  
SPHINGOMYELIN SYNTHASE 1 GENE (*SGMS1*)

I. B. Filippenkov,<sup>1</sup> L. V. Dergunova, A. V. Rozhkova, O. Yu. Sudarkina, S. A. Limborska

Institute of Molecular Genetics RAS, Moscow, 123182;

<sup>1</sup> e-mail: Filippenkov@img.ras.ru

The article is a review of the results of the study of the structural and functional organization of the human sphingomyelin synthase 1 gene (*SGMS1*) in Human Molecular Genetics Department of Institute of Molecular Genetics RAS. *SGMS1* gene encodes an essential enzyme which is involved in the synthesis of sphingomyelin and diacylglycerol from phosphatidylcholine and ceramide, which determines its participation in the regulation of intracellular vesicular transport, cholesterol metabolism, cell proliferation, apoptosis and other significant processes. Our research has shown that the *SGMS1* gene is located on the chromosome 10, has a size of 320 kb and contains more than 20 exons. A detailed study of the *SGMS1* gene's structure allowed us to identify the variety of its transcripts. mRNA isoforms with different fragments of 5' untranslated region (5' UTR) and encoding the full length protein, as well as transcripts resulting from alternative combinations of exons and containing the coding region of the gene and 3' UTR have been discovered. We have found new transcripts among the products of *SGMS1* gene — circular RNAs, which mostly contained sequences of multi-exon 5' UTR of the gene. They are conservative and predominantly expressed in the brain. Circular RNAs of *SGMS1* gene had a large number of binding sites for a microRNA that may determine the functional significance of these molecules. The review describes the latest information about the structural and functional organization of the human gene *SGMS1* as well as the features of its expression.

**Key words:** sphingomyelin synthase 1, *SGMS1* gene structure, alternative transcripts, circular RNAs, tissue-specific expression, microRNAs.

---