

Широкомасштабный биоинформатический поиск генов, играющих важную роль в развитии мозга и регенерации придатков тела

В основе предлагаемого проекта лежит идея о том, что появление или утрата в эволюции крупных признаков, резко меняющих план строения тела и механизмы индивидуального развития организмов, могут быть связаны не только с перестройкой функциональных связей и первичных структур генов, но также с появлением или утратой генов.

Изучение связи между геномными и фенотипическими изменениями в эволюции – одна из основных проблем современной биологии. С одной стороны, этот вопрос является ключевым для детального понимания механизмов эволюции и эмбриогенеза, а с другой – для выявления преобразований в геноме, обусловивших то или иное преобразование фенотипических признаков, что важно для перспективной разработки подходов к управлению биологическими процессами.

Накопившиеся к настоящему времени данные по геномному секвенированию и биоинформатические методы их анализа, а также данные молекулярной биологии развития, выявили высокую консервативность генетических сетей, отвечающих за работу основных молекулярных механизмов в эмбриогенезе позвоночных животных из разных классов. В связи с этим, в научном мире сложилось мнение, что основной причиной даже крупных фенотипических отличий между разными классами современных позвоночных являются какие-то достаточно тонкие модификации одних и тех же генных сетей, связанные либо с перестройкой регуляторных элементов, управляющих генной экспрессией, либо с точечными мутациями в кодирующих участках генов. Также считается, что в эволюции позвоночных, как и других организмов, существенную роль может играть процесс увеличения числа генов, за счёт геномных дупликаций с последующей дивергенцией образовавшихся гомологов. Существенно меньшее значение придается роли полного исчезновения генов. Хотя известно, что наряду с процессом увеличения числа генов, в эволюции шёл и обратный процесс – потери генов, который может рассматриваться также как потенциальный двигатель эволюционных изменений (Olson, Am. J. Hum. Genet. 1999, 64, 18-23).

Большой интерес представит *биоинформатический* широкомасштабный поиск других генов, исчезновение или появление которых в эволюции может коррелировать с эволюцией переднего мозга и понижением способности к регенерации. Также на математических и компьютерных моделях будут исследованы сигнальные системы и функции найденных белков, основные метаболические пути, связанные с ними. Будут построены филогенетические деревья совместной эволюции найденных генов, соответствующих видов, регуляторных систем и их факторов. Разработанные биоинформатические методы будут использованы также для идентификации генов, исчезновение/появление которых могло определять и другие макро-эволюционные изменения.

В настоящее время доступны программы, позволяющие находить в геномных базах данных гены хордовых видов, ортологи которых имеются в двух сравниваемых организмах (например, Ensembl Biomart: <http://www.ensembl.org/biomart/martview>). Результирующие семейства содержат не более двух белков. В нашем же случае необходимо проводить поиск ортологичных генов одновременно во многих геномах позвоночных с тем, чтобы иметь возможность исключить из анализа гены, присутствие или отсутствие которых специфично только для одного вида или даже класса позвоночных. Например, некоторые гены амфибий, возможно, имеются только у них и связаны со специфической физиологией.

Семейства белков будут определяться для большей достоверности по локальной геномной синтении и по аминокислотным заменам, вставкам и делециям в белках, кодируемых сравниваемыми генами. Разделение паралогичных генов представляет

большую трудность, которая преодолевается специальным методом, описанным в (*Математическая биология и биоинформатика* (2013), Т. 8, №1, с. 1-9.). Также этот метод может быть использован как единственный при наличии лишь фрагментарных баз данных (например, баз EST). Насколько нам известно, доступные программы не позволяют осуществлять такой комплексный широкомасштабный поиск. Именно поэтому предполагается создание и отработка соответствующего алгоритма. Он будет, в частности, применён для белков у видов хордовых, в том числе: не являющихся позвоночными *Branchiostoma floridae* (ланцетник) и *Ciona intestinalis*; рыб *Danio rerio*, *Oreochromis niloticus*, *Takifugu rubripes*; лягушки *Xenopus tropicalis*; ящерицы *Anolis carolinensis*; птиц *Gallus gallus*, *Meleagris gallopavo*, *Taeniopygia guttata*; утконоса *Ornithorhynchus anatinus*; сумчатого млекопитающего *Sarcophilus harrisii* и некоторых плацентарных млекопитающих.

В результате, с помощью этого алгоритма будет создан список всех генов, исчезнувших у высших позвоночных, но имеющих у низших позвоночных. Созданный алгоритм поиска генов будет иметь и важное самостоятельное значение, поскольку он будет использован в дальнейшем для решения других задач, предполагающих идентификацию генов исчезнувших/возникших на определенном этапе эволюционного процесса.

Для найденных генов и содержащих их видов будут построены эволюционные деревья и сценарии эволюции (совместные вложения деревьев генов в дерево видов), включающие все эволюционные события (дупликации, потери, возникновения, горизонтальные переносы и т.д.) вместе с временными слоями на дереве видов. Это позволит судить о времени и последовательности возникновения и исчезновения генов, типах эволюционных событий, которые привели к этому. Ссылки: (Lyubetsky et al., *Biology Direct* (2012) 7:48; Горбунов и др. *Молекулярная биология* (2012) 46:1, 176–183; Горбунов и др. *Проблемы передачи информации* (2011) 47:3, 64-79; Gorbunov et al., *In Silico Biology* (2010) 10 0012; Горбунов и др. *Молекулярная биология* (2009) 43:5, 946-958; Горбунов и др., *Молекулярная биология* (2007) 41:5, 918-925; Селиверстов и др. *Молекулярная биология* (2007) 41:3, 572-574; Lyubetsky et al. In: *Bioinformatics of Genome Regulation and Structure II*. Springer Science & Business Media, Inc. (2006) 189-204; Seliverstov et al., *BMC Microbiology* (2005) 5:54; Горбунов и др., *Молекулярная биология* (2005) 39:5, 847-858).

С помощью разрабатываемых нами моделей регуляции и известных баз данных (например, Pfam, ScanProsite) будет проведена классификация белков, кодируемых найденными генами, идентификация сигнальных путей, в которые они участвуют, рассмотрены процессы которые могут регулироваться этими генами в развитии мозга и при регенерации. В работе (Mikhailov et al., *BioEssays* (2009) 31:07, 758-768) нами рассмотрена эволюция таксономической группы Metazoa. Близкие вопросы, связанные с транскрипциях в митохондриях животных рассмотрены в работе (Lyubetsky et al., *Biology Direct* (2012) 7:26). Гиббсовский подход в задаче эволюции регуляторного сигнала экспрессии гена рассмотрен в (Любецкий и др., *Проблемы передачи информации* (2008) 43:4, 52-71). Отметим другие наши работы, связанные с исследованием регуляции и её эволюции (Lyubetsky et al., *Biology Direct* (2011) 6:3; Lyubetsky et al., *Biology Direct* (2010) 5:34; Vitreschak et al., *Ribonucleic acids molecular biology* (2008) 14:4, 717-735; Seliverstov et al., *Biophysics* (2006) 51, Suppl. 1, 518-522; Seliverstov et al. *Journal of Bioinformatics and Computational Biology* (2006) 4:4, 783-793.