

БИОИНФОРМАТИЧЕСКИЙ АНАЛИЗ СЕМЕЙСТВА ГЕНОВ ЦИСТЕИНОВЫХ ПРОТЕАЗ МОЛЛЮСКА *BIOMPHALARIA GLABRATA*

П.Ю. Пинчук

Витебск, ВГУ имени П.М. Машерова

В настоящее время использование биоинформатических методов является необходимым для обработки полученных данных, определения молекулярных механизмов заболеваний, увеличения эффективности сканирования генов. Более того, при изучении возможных фенотипических последствий генетической патологии необходима крайне строгая оценка последствий изменения последовательностей ДНК на молекулярном и клеточном уровнях. Следовательно, задача разработки биоинформатических технологий для оценки функциональных последствий геномных вариаций является актуальной для современной биомедицины. Целью исследования явился биоинформатический анализ цистеиновых протеаз человека (*Homo sapiens*) и моллюска *Biomphalaria glabrata*.

Материал и методы. Материалом для исследования послужили нуклеотидные и аминокислотные последовательности цистеиновых протеаз. В частности, Катепсина В (ЕС:3.4.22.1), Катепсина Z (ЕС:3.4.18.1), Катепсина L (ЕС:3.4.22.15) и Катепсина О (ЕС:3.4.22.42) моллюска *Biomphalaria glabrata*.

Поиск аминокислотных последовательностей ферментов, относящихся к семейству цистеиновых протеаз характерных для моллюска, осуществлялся в базе данных белковых последовательностей – UniProt (<https://www.uniprot.org/>). Поиск, исследуемых последовательностей для человека осуществляла с помощью базового инструмента – BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Веб-приложение Protein Parameters (<https://www.protparam.net/index.html>) использовался для получения значений изоэлектрической точки, алифатического индекса и молекулярной массы. Анализ белков, классифицируя их по семействам и предсказывая домены проводился в программе InterPro 92.0. Шаблоны для построения 3D-моделей ферментов были взяты из банка данных белков PDB (<https://www.rcsb.org/>). По шаблону фермента человека с помощью ресурса SWISS-MODEL (<https://swissmodel.expasy.org/>) строились 3D-модели фермента моллюска [1].

Результаты и их обсуждение. В таблице 1 представлены данные, полученные в ходе идентификации семейства генов цистеиновых протеаз *Biomphalaria glabrata*.

Таблица 1 – Идентификация семейства генов цистеиновых протеаз *Biomphalaria glabrata*

Имя гена	Алифатический индекс	Количество аминокислот	Молекулярная масса (Да)	Изоэлектрическая точка	Локализация	Домены
Катепсин В	66,25	333	36540,2	6,58	клеточная мембрана	C1A (PF00112) 87-330
Катепсин Z	69,15	351	39220,2	7,18	лизосома	C1A (PF00112) 113-350
Катепсин L	70,00	326	35696,8	5,50	лизосома; клеточная мембрана	C1A (PF00112) 110-325
Катепсин О	78,98	343	38848,7	6,27	лизосома	C1A (PF00112) 124-334

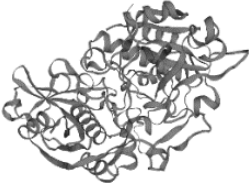
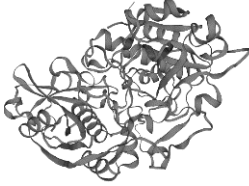
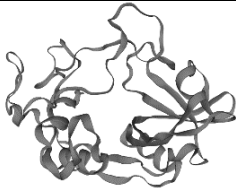
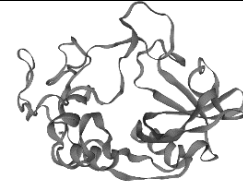
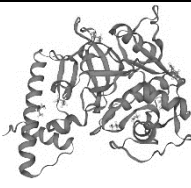


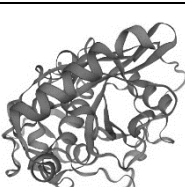
Из анализа данных таблицы следует, что наиболее термостабильный фермент – Катепсин О (алифатический индекс – 78,98). Молекулярная масса исследуемых ферментов колеблется от 35696,8 Да до 39220,2 Да. Все ферменты обладают кислотными свойствами. За исключением Катепсина В, ферменты локализованы в лизосомах.

При хороших показателях выравнивания аминокислотных последовательностей исследуемых ферментов можно говорить о гомологии, о сходстве строения и функции исследуемых протеинов и смоделировать третичную структуру. При этом установлено, что белки, имеющие не менее 40-50% гомологии в аминокислотных последовательностях, будут иметь похожие третичные структуры. Следовательно, можно предполагать, что и функции таких белков могут быть похожими [2].

При сравнении аминокислотных последовательностей цистеиновых протеаз человека и моллюска *Biomphalaria glabrata* было установлено, что наибольший процент гомологии характерен для фермента Катепсин Z – 66,26 %. Для остальных ферментов уровень гомологии колеблется в промежутке 40-53 %, а именно у Катепсина L – 52,25 %, Катепсина В – 48,82 % и Катепсина О – 41,64 %.

Для дальнейшего анализа были построены 3D-структуры исследуемых ферментов человека (*Homo sapiens*) и моллюска *Biomphalaria glabrata* и проведён их сравнительный анализ результаты, которого представлены в таблице 2.

Таблица 2 – 3D – структуры лизосомальных ферментов человека (*Homo sapiens*) и моллюска *Biomphalaria glabrata*

Фермент	Характеристика		<i>Homo sapiens</i>	<i>Biomphalaria glabrata</i>
Cathepsin Z (EC: 3.4.18.1)	GMQE	0,77		
	QMEAN	-1,27		
	Identity	69,01%		
Cathepsin B (EC: 3.4.22.1)	GMQE	0,39		
	QMEAN	-0,53		
	Identity	62,69%		
Cathepsin L (EC:3.4.22.15)	GMQE	0.82		
	QMEAN	0.81		
	Identity	54,72%		
Cathepsin O (EC:3.4.22.42)	GMQE	0.58		
	QMEAN	0.69		
	Identity	40,15%		

Примечание: GMQE – глобальная оценка качества модели; QMEAN – составная оценка, основанная на различных геометрических свойствах и представляет, как глобальные, так и локальные оценки абсолютного качества на основе модели; Identity – гомология, идентичность.

Установлено, что процент гомологии третичных структур варьирует от 40,15 % до 69,01 %. В результате выравнивания установлено, что наибольший процент гомологии характерен для фермента Катепсин Z (69,01 %), наименьший процент гомологии выявлен у фермента Катепсин O (40,15 %). У ферментов Катепсин L и Катепсин B процент гомологии равен 54,72 % и 62,69 % соответственно.

Закключение. Практическое значение достаточно высокой степени гомологии цистеиновых протеаз у людей и легочных пресноводных моллюсков обосновывает целесообразность формирования аквакультуры моллюсков, для получения из их тканей белковых ферментативных препаратов протеолитического действия в рамках задач биофармацевтики, косметологии и пищевой промышленности.

1. Милеева, П.Ю. Сравнительная характеристика лизосомальных ферментов человека и легочных пресноводных моллюсков / П.Ю. Милеева // Химия, физика, биология, математика: теоретические и прикладные исследования: сб. ст. по материалам XLIII Междунар. науч.-практ. конф. «Химия, физика, биология, математика: теоретические и прикладные исследования», Москва, 18 декабря 2020 г. / Изд. «Интернаука». – Москва, 2020. – №12(32). – С.7-11.

2. Choithia C., Lesk A.M. The relation between the divergence of sequence and structure in proteins// Embo J. 1986 №4 (5). P.823-826.