

©Коллектив авторов

## ИНТЕРАКТОМИКА СХХС-БЕЛКОВ, УЧАСТВУЮЩИХ В ЭПИГЕНЕТИЧЕСКОЙ РЕГУЛЯЦИИ ГЕННОЙ ЭКСПРЕССИИ

*П.В. Ершов, Е.О. Яблоков\*, Ю.В. Мезенцев, А.С. Иванов*

<sup>1</sup>Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича,  
119121, Москва, ул. Погодинская, 10; \*эл. почта: evgeyablokov1988@mail.ru

Регуляция экспрессии генов является крайне сложным и многокомпонентным биологическим явлением. Белки, содержащие СХХС-домен “цинковые пальцы” (СХХС-белки), являются мастер-регуляторами экспрессии многих генов и обладают консервативными функциями метилирования оснований ДНК и гистоновых белков. СХХС-белки функционируют в составе сложных мультибелковых комплексов, что указывает на фундаментальную важность изучения посттрансляционной регуляции посредством модуляции спектра белок-белковых взаимодействий (ББВ) как в норме, так и при патологии. В работе обобщены собственные и литературные данные о роли СХХС-белков и их белковых партнёров в неопластических процессах. Особое внимание уделено интерактоме белка CFP1, кодируемого геном *CXXC1*, расположенным на 18-ой хромосоме человека. CFP1 не проявляет ферментативной активности и реализует эпигенетическую регуляцию экспрессии через связывание с хроматином и определённый спектр ББВ.

**Ключевые слова:** интерактомика; СХХС-белки; молекулярный фишинг; гель-фильтрация; эпигенетическая регуляция

**DOI:** 10.18097/PBMC20226805339

### ВВЕДЕНИЕ

Бинарные и мультибелковые комплексы выполняют средообразующую, сигнальную, регуляторную и ряд других функций для обеспечения жизнедеятельности клеток и тканей организма человека. Нарушение сетей белок-белковых взаимодействий (ББВ) в условиях патологий приводит к появлению aberrантных белковых комплексов, перераспределению потоков биологической информации [1-3] и модуляции активности транскрипционных мастер-регуляторов (главные факторы транскрипции, регулирующие экспрессию множественных нижележащих генов, уровень экспрессии которых может определять направления передачи биологической информации в клеточных путях) [4, 5]. Поэтому исследование интерактома белков, которые участвуют в регуляции транскрипционного аппарата клеток, имеет фундаментальное и прикладное значение, способствуя совершенствованию способов фармакологической коррекции нарушений в сетях ББВ при заболеваниях [6]. Среди мастер-регуляторов особое внимание привлекает группа, представленная одиннадцатью белками, содержащими СХХС-домен (СХХС-белки) (табл. 1), которые регулируют экспрессию генов посредством эпигенетических модификаций хроматина [7-9]. СХХС-белки в значительной мере отличаются друг от друга по первичной аминокислотной последовательности и выполняемым функциям [10], но содержат одинаковый домен СХХС “цинковые пальцы” [11]. Этот домен состоит из двух цистеин-содержащих аминокислотных мотивов CxxC-x17-CxxC и нужен для связывания с хроматином в областях неметилированных CpG-островков. Каждый мотив хелатирует один ион цинка. Белки СХХС1, СХХС3, СХХС4 и СХХС5 не проявляют

ферментативной активности; ферментативной активностью обладают СХХС2 и СХХС8 (гистон H3 Lys36-деметилазная активность), СХХС9 (метилирование цитозинового основания ДНК), СХХС6 (гидроксиметилирование цитозинового основания ДНК), а также СХХС7 и СХХС10 (гистон H3 Lys4-метилазная активность) (табл. 1).

Специфичность паттерна эпигенетической регуляции при развитии патологий [12, 13] может обеспечиваться через модуляцию спектра ББВ и посттрансляционные модификации, что также верно и в отношении СХХС-белков, которые принимают непосредственное участие в данном биологическом процессе. Поэтому исследования интерактома СХХС-белков как совокупного спектра взаимодействующих друг с другом индивидуальных белков и белковых комплексов важны с точки зрения получения новых знаний о: а) прямых и непрямых белковых партнёрах, формирующих мультибелковые комплексы; б) модифицирующих белках, таких как протеинкиназы и убиквитиназы; в) “каркасных” белках (от англ. scaffold), которые, в частности, были идентифицированы в качестве белковых партнёров СХХС5 [14].

Целью данной работы был анализ известных источников интерактомной информации о СХХС-белках (в том числе и о CFP1), принимающих участие в эпигенетической регуляции генной экспрессии в организме человека в норме и при патологиях.

### 1. АССОЦИАЦИЯ СХХС-БЕЛКОВ С ПАТОЛОГИЧЕСКИМИ ПРОЦЕССАМИ

Литературные данные свидетельствуют о прямом или опосредованном вовлечении генов, кодирующих СХХС-белки, в ряд патологий, главным образом,

# ИНТЕРАКТОМИКА СХХС-БЕЛКОВ

Таблица 1. Номенклатура целевых генов, кодирующих СХХС-белки организма человека

Обозначение гена	Каноничный символ гена* / синонимы	Название гена, англ / рус	HGNC ID**
<i>CXXC1</i>	<b><i>CXXC1</i></b> / <i>HsT2645, PCCX1, hCGBP, PHF18, CGBP, SPP1, CFP1, ZCGPC1</i>	CXXC finger protein 1 / белок «цинковых пальцев» 1	24343
<i>CXXC2</i>	<b><i>KDM2B</i></b> / <i>FBXL10, PCCX2, Fbl10, JHDM1B</i>	lysine demethylase 2B / лизин деметилаза 2B	13610
<i>CXXC3</i>	<b><i>MBD1</i></b> / <i>PCM1</i>	methyl-CpG binding domain protein 1 / белок 1 метил-CpG-связывающий домен	6916
<i>CXXC4</i>	<b><i>CXXC4</i></b> / <i>IDAX</i>	CXXC finger protein 4 / белок «цинковых пальцев» 4	24593
<i>CXXC5</i>	<b><i>CXXC5</i></b> / <i>HSPC195, RINF, WID</i>	CXXC finger protein 5 / белок «цинковых пальцев» 5	26943
<i>CXXC6</i>	<b><i>TET1</i></b> / <i>LCX, KIAA1676, bA119F7.1</i>	tet methylcytosine dioxygenase 1 / tet метилцитозин диоксигеназа 1	29484
<i>CXXC7</i>	<b><i>KMT2A</i></b> / <i>TRX1, HRX, ALL-1, HTRX1, MLL1A, MLL1, ALL1, HTRX</i>	lysine methyltransferase 2A / лизин метилтрансфераза 2A	7132
<i>CXXC8</i>	<b><i>KDM2A</i></b> / <i>FBXL11, KIAA1004, FBL11, LILINA, FBL7, FLJ00115, JHDM1A</i>	lysine demethylase 2A / лизин деметилаза 2A	13606
<i>CXXC9</i>	<b><i>DNMT1</i></b> / <i>MCMT</i>	DNA methyltransferase 1 / ДНК метилтрансфераза 1	2976
<i>CXXC10</i>	<b><i>KMT2B</i></b> / <i>KIAA0304, MLL2, TRX2, HRX2, WBP7, MLL1B, MLL4</i>	lysine methyltransferase 2B / лизин метилтрансфераза 2B	15840
<i>CXXC11</i>	<b><i>RTP5</i></b> / <i>C2orf85, FLJ33590, Z3CXXC5</i>	receptor transporter protein 5 (putative) / белок 5 рецептор-транспортер	26585

Примечание: \* – каноничный символ гена выделен полужирным шрифтом; \*\* – согласно данным портала HGNC [8].

в неопластические процессы. Это следует из простого анализа числа публикаций в данной области за последние пять лет: *CXXC11* (5 публикаций); *CXXC1* (8); *CXXC4* (24); *CXXC5* (45); *CXXC3* (61); *CXXC8* (84); *CXXC2* (89); *CXXC10* (174); *CXXC6* (250); *CXXC9* (800); *CXXC7* (свыше 1000 публикаций) [15]. Наибольший научный интерес к белкам *CXXC7* и *CXXC9* обусловлен тем, что они рассматриваются с точки зрения молекулярных мишеней для прототипов лекарств с целью фармакологической коррекции эпигенетических нарушений [16, 17]. В зависимости от биологического контекста гены *CXXC* вносят разнонаправленный вклад в нормальные и патологические процессы, что можно проследить на следующих примерах из литературы. Так, впервые было обнаружено, что слияние генов *CXXC7* и *MAML2* (mastermind-like protein 2) характерно для двух агрессивных гистологических подтипов тимомы B2 и B3 и оно может лежать в основе резистентности к терапии [18]. Связь соматических полиморфных вариантов с новообразованиями с высокой летальностью была установлена в отношении генов *CXXC10* [19] и *CXXC9* [20]. Интересно, что спектр мутаций в гене *CXXC6* имел значимость для прогноза ответов на лечение ингибиторами контрольных точек иммунитета и общей выживаемости пациентов с разными нозологиями [21]. Как правило, индукция экспрессии большинства генов *CXXC* в опухолевых тканях коррелирует с развитием высокопролиферативных и инвазивных новообразований [22, 23], о чём также свидетельствуют эксперименты по нокдауну генов [14, 24, 25]. Также было показано, что паттерн

экспрессии ассоциирован с неблагоприятным прогнозом онкологического заболевания и метастатическими осложнениями [26, 27]. Однако гиперэкспрессия гена *CXXC4* в опухолевых клетках приводила к увеличению их восприимчивости к воздействию иммунной системы организма [28], активации апоптоза [29], снижению резистентности к тамоксифену [30], а индукция экспрессии *CXXC9* сопровождалась гиперметилованием CpG-динуклеотидов в промоторах генов опухолевых супрессоров, приводя к торможению роста опухолей [31].

Экспрессия генов *CXXC* регулируется рядом транскрипционных факторов (ТФ). Согласно данным портала GeneCards [32, 33], экспрессию всех одиннадцати генов *CXXC* могут регулировать четыре десятка ТФ, среди которых ATF1, CREM, EGR1, EGR2, HDAC1, HDAC2, MLLT1, SIN3A, SP1 и ZBTB17 являются опухолеассоциированными (KEGG ID hsa05202, hsa05166, hsa05203) [34]. В свою очередь, СХХС-белки регулируют экспрессию генов-драйверов неопластической трансформации и могут способствовать “переключению” сигнальных путей. По данным портала GeneCards, экспрессию генов *CXXC3*, *CXXC9* и *CXXC10* регулируют десять ТФ (рис. 1), которые участвуют в передаче сигналов (KEGG ID hsa05200). В свою очередь, каждый из *CXXC3*, *CXXC9*, *CXXC10* белков может регулировать достаточно большое число генов с высокой частотой мутаций (табл. 2), но только пять из них (*ATR*, *CDC73*, *BRCAl*, *LZTR1* и *PRKAR1A*) являются общими мишенями для трёх рассматриваемых СХХС-белков [33, 35]. На рисунке 1



**Рисунок 1.** Пример участия СХХС-белков в сигнальных каскадах регуляции экспрессии генов, ассоциированных с неопластической трансформацией клеток: *ATR* (ген *ATR* Ser/Thr протеинкиназы), *CDC73* (цикл клеточного деления 73), *BRCA1* (ген *BRCA1* ассоциированный с репарацией ДНК), *LZTR1* (ген регулятор 1 фактора транскрипции “лейциновая молния”) и *PRKAR1A* (альфа субъединица сАМР-зависимой протеинкиназы тип 1). Сигнальные и клеточные пути (по данным портала Reactome), в которые вовлечены ряды генов: R-HSA-1640170, клеточный цикл; R-HSA-5358351, Hedgehog; R-HSA-5685942, репарация ДНК путём гомологичной рекомбинации; R-HSA-157118, NOTCH (семейство трансмембранных белков, содержащие повторяющиеся внеклеточные домены EGF и DSL); R-HSA-170834, TGF-β (трансформирующий фактор роста бета); R-HSA-195721, WNT.

**Таблица 2.** Данные об опухолеассоциированных драйверных генах (гены, которые содержат определённые мутации, обеспечивающие преимущество опухолевым клеткам в процессе эволюции), экспрессия которых регулируется СХХС-белками

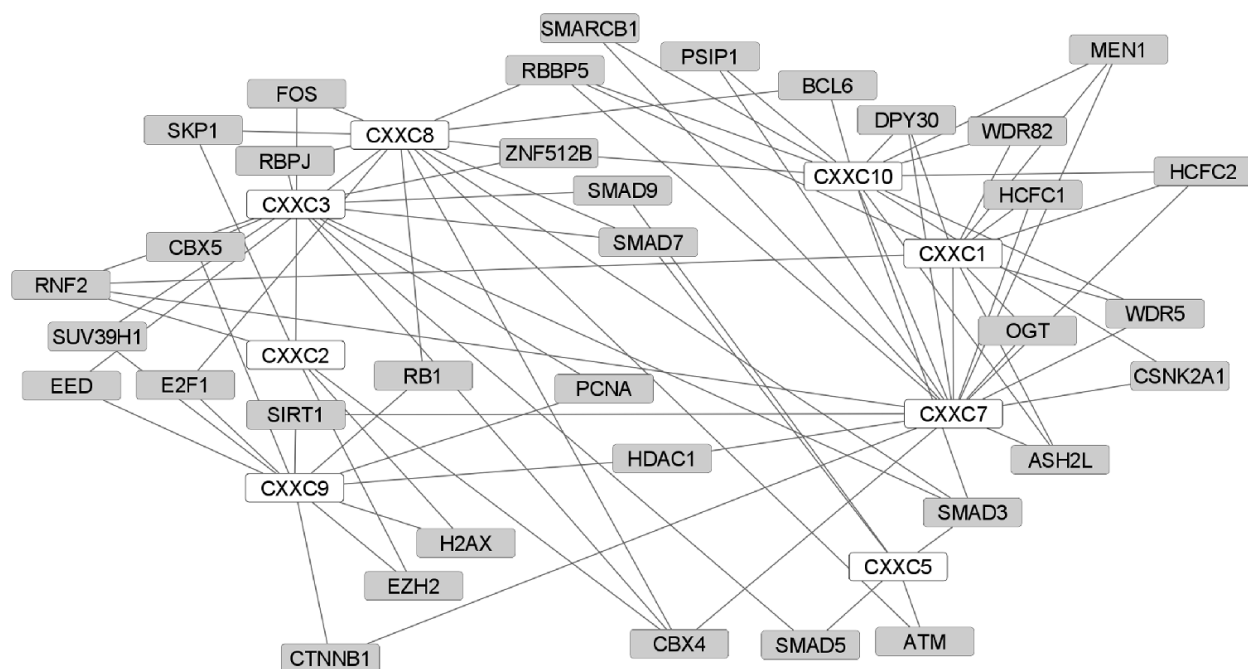
СХХС-белки	Гены-мишени*
CXXC10	<i>ATR, BARD1, BAP1, CTNNA1, SMAD4, ATM, AXIN2, CDC73, CDKN1B, DICER1, BRCA1, BRCA2, CYLD, LZTR1, MEN1, MSH2, MSH6, PRKAR1A, RB1, SPOP, SDHA, TGFB2, TP53, VHL</i>
CXXC3	<i>ATR, BARD1, BAP1, MLH1, FAT1, PTCH1, SETBP1, AXIN2, CDH1, CDC73, CDKN2A, BRCA1, LZTR1, MEN1, MSH2, PDGFRA, PRKAR1A, SPOP, TGFB2, HRAS, VHL</i>
CXXC9	<i>ATR, CTNNA1, PTCH1, APC, CDC73, BRCA1, LZTR1, PRKAR1A</i>

Примечание: \* — поиск генов-мишеней и регулирующих их экспрессию транскрипционных факторов был выполнен при помощи онлайн-каталога Cancer Gene Census [35] и портала GeneCards [33].

схематично показано, что “расположенные выше” ТФ, которые регулируют экспрессию генов СХХС, управляются одним спектром сигнальных путей, тогда как СХХС-белки регулируют гены, представленные в иных сигнальных и клеточных путях. Таким образом, можно полагать, что ряд СХХС-белков, осуществляя эпигенетическую регуляцию экспрессии генов-мишеней, выступает в роли молекулярных “переключателей” сигнальных путей в опухолевых клетках (рис. 1).

## 2. ОБЩИЕ ОСОБЕННОСТИ ИНТЕРАКТОМИКИ СХХС-БЕЛКОВ

Данные о ББВ с участием СХХС-белков были получены с помощью классических экспериментальных подходов, таких как панорамное интерактомное профилирование, ко-иммунопреципитация, аффинное выделение в комбинации с масс-спектрометрией, а также подхода с двухстадийным разделением мультибелковых комплексов при сочетании гель-электрофореза в нативных и денатурирующих условиях [36]. Важным моментом валидности этих и других подходов является интерпретация массивов данных о ББВ, которая осложняется проблемой различий между прямыми и непрямыми белковыми партнёрами. Поэтому усовершенствование биоинформатических и системнобиологических алгоритмов [37], а также создание электронных ресурсов по каталогизации ББВ способствует значимому увеличению ценности интерактомных данных, полученных экспериментально. Информация о физических взаимодействиях между СХХС-белками и их белковыми партнёрами была собрана при помощи следующих электронных ресурсов: Funcoup 5.0 [38], Mentha [39], APID [40], MINT [41], HuRI [42], IID [43], Bioplex Explorer 3.0 [44], Wiki-Pi [45], HIPPIE 2.0 [46], HINT (High-quality INteractomes) [47] и STRINGdb v.11.5 [48]. В финальный список были включены только те белковые партнёры, которые встречались в трёх и более ресурсах и имели сходство по профилю субклеточной локализации с соответствующим СХХС-белком. Таким образом, была собрана информация о белковых партнёрах СХХС1 (37 белков); СХХС2 (16 белков); СХХС3 (41 белок); СХХС4 (5 белков); СХХС5 (7 белков); СХХС6 (0 белков); СХХС7 (50 белков); СХХС8 (19 белков); СХХС9 (65 белков); СХХС10 (36 белков) и СХХС11 (50 белков). Сравнение списков белков позволило выявить 35 общих белковых партнёров (около 15% от суммарного числа белковых партнёров) для восьми СХХС-белков, которые образуют сеть ББВ (рис. 2). Следует сказать, что для СХХС4 и СХХС11 не было найдено общих белковых партнёров с другими СХХС-белками. СХХС11 и его белковые партнёры отличались цитоплазматической субклеточной локализацией, что было нетипичным для других десяти СХХС-белков с локализацией в ядре. Анализ генной онтологии и молекулярных путей показал, что белковые партнёры СХХС-белков вовлечены в опухолеассоциированные сигнальные пути (FoxO, Wnt, TGF-β), а также в патогенез вирусных и системных невирусных заболеваний (табл. 3) [49].



**Рисунок 2.** Сеть экспериментально подтверждённых ББВ с участием ряда СХХС-белков и их общих белковых партнеров. Для визуализации ББВ была использована программа Cytoscape v.3.8.2. Данные о ББВ были извлечены с помощью интерактивных ресурсов Funcoup 5.0, Mentha, APID, MINT, HuRI, IID, Bioplex Explorer 3.0, Wiki-Pi, HIPPIE 2.0, HINT (High-quality INTeractomes) и STRINGdb v.11.5.

**Таблица 3.** Участие белковых партнеров СХХС-белков в молекулярных путях

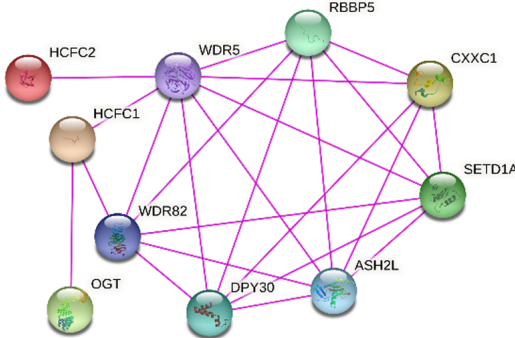
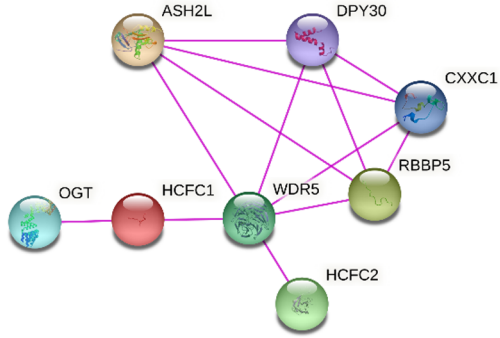
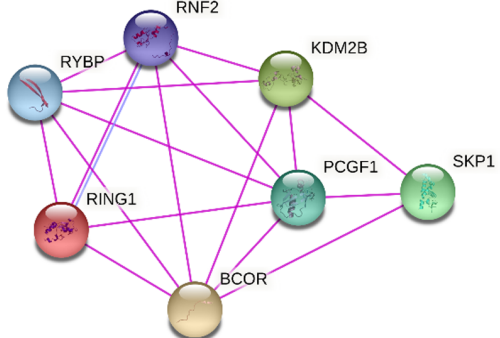
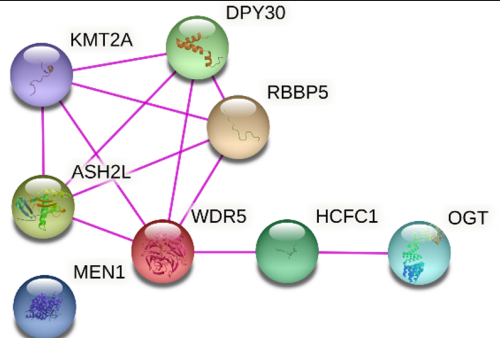
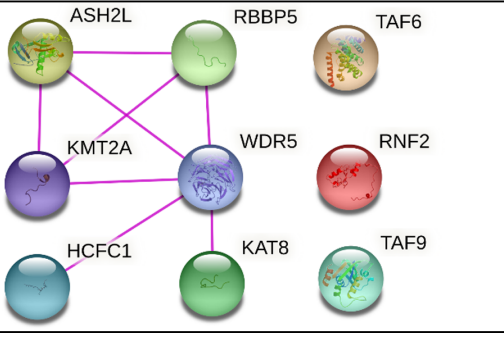
KEGG ID	Термины KEGG*	Белки
hsa04068 hsa05220 hsa05225	Сигнальный путь FoxO Хроническая миелоидная лейкемия Гепатоцеллюлярная карцинома	CXXC8, CXXC9, CXXC10
hsa04110	Клеточный цикл	CXXC2, CXXC7, CXXC8, CXXC9
hsa04310	Сигнальный путь Wnt	CXXC2, CXXC7, CXXC9
hsa04350	Сигнальный путь TGF- $\beta$	CXXC2, CXXC3, CXXC5, CXXC8
hsa04520 hsa05206	Адгезивные соединения МикроРНК в опухолях	CXXC7, CXXC9, CXXC10
hsa05202	Дисрегуляция транскрипции генов	CXXC7, CXXC8, CXXC9, CXXC10
hsa05215	Рак предстательной железы	CXXC1, CXXC8, CXXC9
hsa04934	Синдром Кушинга	CXXC1, CXXC7, CXXC8, CXXC10
hsa05168	Инфекция простого вируса герпеса	CXXC1, CXXC7, CXXC8, CXXC9, CXXC10

Примечание: \* – анализ обогащения списков генов терминами KEGG (Kyoto Encyclopedia of Genes and Genomes) выполнен при помощи сервера Webgestalt [44]: минимальное число генов в категории – 5; поправка на множественное сравнение – метод Бенжамини-Хохберга; уровень значимости –  $FDR \leq 0,1$ .

Белки СХХС1, СХХС2, СХХС3, СХХС7, СХХС9 и их белковые партнёры участвуют в формировании мультибелковых комплексов, регулирующих аппарат транскрипции путём генетического импринтинга и посттрансляционных модификаций хроматина (метилирование/деметилирование остатков лизина гистоновых белков H3, убиквитинилирование и деацетилирование белков) (табл. 4) [49-51]. Таким образом, СХХС-белки характеризуются преимущественным взаимодействием с белками, локализованными в клеточном ядре, образуя там сложные мультибелковые комплексы.

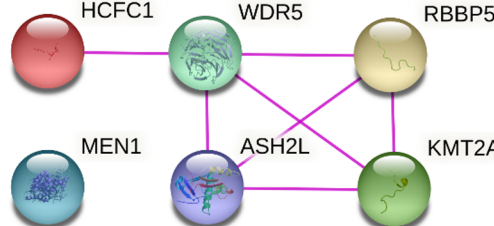
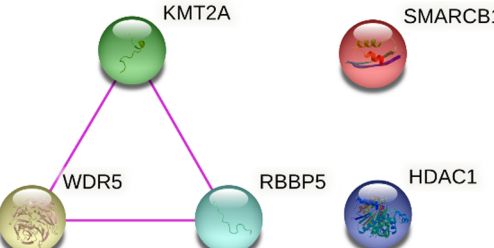
Предположительно, активность СХХС-белков как эпигенетических мастер-регуляторов по-разному модулируется в условиях нормы и патологии, но так или иначе связана с функционированием в составе сложных мультибелковых комплексов, композиция которых может быть контекст-зависимой. Например, модуляция активности белка может быть следствием взаимодействия со своим белковым партнёром. В таблице 5 мы привели потенциальные опухолеспецифичные ББВ с участием СХХС-белков. Для их селекции сначала были найдены дифференциально экспрессируемые

Таблица 4. Представленность CXXC-белков и их партнеров по взаимодействию в составе функционально значимых мультибелковых комплексов

ID Corum database*	Название / Композиция белкового комплекса**	Композиция мультибелкового комплекса согласно данным STRINGdb***
6469	Set1A complex <b>CXXC1</b> , ASH2L, DPY30, HCFC1, HCFC2, OGT, RBBP5, SETD1A, WDR5, WDR82	
6470	Set1B complex <b>CXXC1</b> , ASH2L, DPY30, HCFC1, HCFC2, OGT, RBBP5, WDR5	
1178	BCOR complex (ubiquitin E3 ligase) <b>CXXC2 (KDM2B)</b> , BCOR, PCGF1, RING1, RNF2, RYBP, SKP1	
6462	MLL1 complex <b>CXXC7 (KMT2A)</b> , ASH2L, DPY30, HCFC1, MEN1, OGT, RBBP5, WDR5	
5386	MLL1-WDR5 complex <b>CXXC7 (KMT2A)</b> , ASH2L, HCFC1, KAT8, RBBP5, RNF2, TAF6, TAF9, WDR5	

## ИНТЕРАКТОМИКА СХХС-БЕЛКОВ

Таблица 4. Представленность СХХС-белков и их партнеров по взаимодействию в составе функционально значимых мультибелковых комплексов

ID Corum database*	Название / Композиция белкового комплекса**	Композиция мультибелкового комплекса согласно данным STRINGdb***
1256	MLL-HCF complex <b>CXXC7 (KMT2A)</b> , ASH2L, HCFC1, MEN1, RBBP5, WDR5	
1257	ALL-1 supercomplex <b>CXXC7 (KMT2A)</b> , HDAC1, RBBP5, SMARCB1, SYMPK, TAF6, TAF9, WDR5	
5695	TIP5-DNMT-HDAC1 complex <b>CXXC9 (DNMT1)</b> , BAZ2A, DNMT3B, HDAC1	Не применимо
2233	Replication-coupled CAF-1-MBD1-ETDB1 complex <b>CXXC3 (MBD1)</b> , CHAF1A, SETDB1	Не применимо
2763	MBD1-Suv39h1-HP1 complex <b>CXXC3 (MBD1)</b> , CBX5, SUV39H1	Не применимо

Примечание: \* – CORUM – The comprehensive resource of mammalian protein complexes [45]; \*\* – анализ обогащения списков генов терминами CORUM выполнен при помощи сервера Webgestalt [44]: минимальное число генов в категории – 3; поправка на множественное сравнение – метод Бенжамини-Йекутиэлли; уровень значимости –  $FDR \leq 0,1$ ; \*\*\* – STRINGdb v. 11.5 [46]: организм – *Homo sapiens*; тип сети – physical subnetwork; значения рёбер сети – evidence; источник взаимодействий – experimental; score > 0,7. Название СХХС-белков выделено жирным

Таблица 5. Опухолеспецифичные ББВ с участием СХХС-белков

СХХС-белок/ белковый партнёр	Опухолевая ткань*	Коэкспрессия генов**	Прогностическая значимость (HR, высокий уровень экспрессии)***
CXXC2 / RNF2	TGCT	0,48	нет
CXXC3 / SMAD1	TGCT	0,66	нет
CXXC8 / RBBP5	TGCT	0,67	нет
CXXC9 / EZH2	BLCA	0,76	нет
CXXC9 / HELLS	BLCA	0,65	нет
CXXC9 / PCNA	BLCA	0,68	нет
CXXC9 / UHRF1	BLCA	0,74	нет
CXXC9 / EZH2	STAD	0,62	нет
CXXC9 / CHEK1	STAD	0,68	нет
CXXC9 / PCNA	STAD	0,60	нет
CXXC9 / UHRF1	STAD	0,69	нет

Примечание: \* – BLCA – уротелиальная карцинома; STAD – аденокарцинома желудка; TGCT – рак семенников; \*\* – коэффициенты корреляции по Спирмену  $r \geq 0,4$  и  $p \leq 0,05$ ; \*\*\* – общая выживаемость (ОВ) и выживаемость без прогрессирования (ВБП) по графикам Каплан-Майера ( $HR \leq 0,5$ ;  $HR \geq 2$ ;  $p \leq 0,05$ ;  $n \geq 35$ , где HR – отношение рисков, n – число пациентов в группах).

гены (ДЭ-гены) СХХС и гены, кодирующие их партнёров по взаимодействию в нормальной и опухолевых тканях ( $|\log_2 FC| = 0,9$  и  $p \leq 0,05$ ) при помощи GENT2 [52] и GEPIA2 [53]. Затем при помощи веб-инструмента GEPIA2 исследовали коэкспрессию и прогностическую значимость

экспрессии каждой пары ДЭ-генов, продукты которых образуют ББВ. Несмотря на наличие высокой корреляции коэкспрессии ДЭ-генов, связей между прогнозом заболевания и уровнями их экспрессии не было выявлено (табл. 5). Тем не менее, мы нашли, что гены CXXC9 (DNMT1), EZH2, HELLS,

*PCNA*, *UHRF1*, *CHEK1* характеризовались высокой экспрессией в некоторых типах опухолей, таких как аденокарцинома, папиллярная почечно-клеточная карцинома и мезотелиома, а паттерн экспрессии имел связь с неблагоприятным прогнозом общей выживаемости. Однако какой-либо информации о функционировании белков, кодируемыми данными генами, в составе единого мультибелкового комплекса пока нет в литературе. Стоит отметить, что некоторые СХХС-белки могут выступать в качестве партнёров по ББВ между собой. Например, в интерактомных базах данных есть записи о взаимодействии между СХХС7 и СХХС1, СХХС7 и СХХС10, а в литературе — о СХХС1 и СХХС9 [54]. Для последней пары Cherau с коллегами показали, что дестабилизация этого ББВ при помощи пептидного ингибитора увеличивала гибель опухолевых клеток при воздействии темозоломида, и авторы резонно предположили об участии ББВ в метилировании генов, ответственных за химиорезистентность. По данным портала TCGA (The Cancer Genome Atlas), средневзвешенная коэкспрессия генов *CXXC2*, *CXXC3*, *CXXC7* и *CXXC8* характеризуется высокими значениями коэффициентов корреляции Пирсона 0,75–0,84 ( $p < 0,05$ ) в большинстве тканей человека, что косвенным образом предполагает возможные ББВ, которые могут существовать и для других СХХС-белков.

Таким образом, сеть ББВ с участием СХХС-белков, где на каждый СХХС-белок в среднем приходится по 11 связей (от 5 до 21), состоит из высококоннективной центральной части (рис. 2), состоящей из прямых взаимодействий между собственно СХХС-белками и пула общих белковых партнёров, которые опосредованно связывают остальные СХХС-белки между собой.

### 3. НОВЫЕ ДАННЫЕ ОБ ИНТЕРАКТОМИКЕ БЕЛКА CFP1

В своих работах мы сделали акцент на расширении интерактомной информации о мастер-регуляторе CFP1 (CxxC-finger protein-1, Uniprot ID: Q9P0U4) процессов эпигенетической регуляции генной экспрессии [55, 56] как одним из наименее изученных СХХС-белков в контексте вовлечённости в неопластическую трансформацию клеток [57]. CFP1 кодируется геном *CXXC1* 18-ой хромосомы человека, что обуславливает также важность изучения его ББВ в рамках международного Проекта “Протеом Человека” (The Human Proteome Project, HPP). CFP1 является мультифункциональным и мультидоменным белком, влияющим на уровень метилирования ДНК и пост-трансляционные модификации гистоновых белков по типу метилирования и ацетилирования [56, 58]. CFP1 связывается с открытым хроматином в области CpG-островков посредством домена “цинковые пальцы” [59]. Известно, что в норме CFP1 принимает участие в клеточной пролиферации и дифференцировке [60, 61], гематопозе [62], регуляции клеточного цикла [63, 64], развитии Т-лимфоцитов

в тимусе [65] и гаметогенезе у млекопитающих [66]. Поэтому выяснение всего спектра белковых партнёров CFP1 важно с фундаментальной точки зрения для понимания аспектов регуляции через ББВ не только в нормальных, но и опухолевых клетках, где увеличивается вероятность возникновения aberrантных белковых взаимодействий.

Используя подход прямого молекулярного фишинга, основанного на сочетании аффинной хроматографии, масс-спектрометрии и поверхностного плазмонного резонанса [67], из общего лизата культуры опухолевых клеток HepG2, полученной из гепатоцеллюлярной карциномы человека, были выделены 32 потенциальных белковых партнёра CFP1 (табл. 6) [68–73]. Анализ литературных данных по гель-хроматографии клеточных лизатов и транскриптомному профилированию тканей показал, что, во-первых, 12 белков (ANP32A, ANXA7, CARM1, CDC42, CPSF1, EIF2B4, PGRMC2, PLIN3, PSME2, TIAL1, LGALS8 и DNAJC10) из общего числа выделенных нами потенциальных белков-партнёров кофракционируются с CFP1 (табл. 6), а во-вторых, ген *CXXC1* имеет высокую корреляцию коэкспрессии с соответствующими генами, кодирующими эти белки (табл. 6). Таким образом, в литературе есть косвенные указания на взаимодействие перечисленных выше белков с CFP1. Среди найденных белковых партнёров CFP1 наиболее однородной группой представлены РНК-связывающие белки (CPSF1, CSTF2, EIF2B4, NOB1, ANP32A, ANXA7 и DAZAP1). Общеизвестно, что РНК-связывающие белки не только имеют способность связываться друг с другом в разных комбинациях, определяя специализацию трансляционного аппарата [74], но и взаимодействовать с обычными транскрипционными факторами для модуляции генной экспрессии [75, 76]. Один из идентифицированных белковых партнёров CFP1 — DAZ-ассоциированный белок-1 (DAZAP1) — участвует в сплайсинге мРНК и является клинически значимым белком; нокдаун гена *DAZAP1* приводит к существенному ингибированию пролиферации, миграции и инвазии клеток гепатоцеллюлярной карциномы [77]. В базе данных Biogrid [78] и NURSA [79] мы обнаружили, что белковые партнёры CFP1 (NQO1, SOD1, TIA1, TIAL1 и CARM1) физически взаимодействуют с DAZAP1, при этом белки TIA1 и TIAL1 участвуют в регуляции трансляции и альтернативного сплайсинга [80]. Коммуникация аргинин-метилтрансферазы CARM1 с CFP1, которая модифицирует РНК-связывающие белки [81], осуществляется через ко-локализацию этих двух белков в составе мультибелкового комплекса с метилтрансферазами гистонов, участвующего в эпигенетической регуляции [82]. Таким образом, CFP1 и часть идентифицированных в наших экспериментах белковых партнёров вовлечены в смежные процессы регуляции экспрессии генов на транскрипционном и посттранскрипционном уровнях.

Ещё один белок из числа обнаруженных нами белков-партнёров CFP1 — ко-шаперон DNAJC10 (гомолог DnaJ подсемейства C10). Его основная

## ИНТЕРАКТОМИКА СХХС-БЕЛКОВ

Таблица 6. Список<sup>#</sup> потенциальных белковых партнеров CFP1, кодируемого геном *CXXC1*, выделенных из лизата клеточной линии HepG2

Символ гена	Коэкспрессия*	Кофракционирование**	Название гена	Uniprot ID	M <sub>r</sub> , кДа
ACAT2	—	—	Acetyl-CoA acetyltransferase, cytosolic	Q9BWD1	41,3
ANP32A	•	•	Acidic leucine-rich nuclear phosphoprotein 32 family member A	P39687	28,5
ANXA7	•	•	Annexin A7	P20073	52,7
CASP1	—	—	Caspase-1	P29466	45,2
CDC42	•	•	Cell division control protein 42 homolog	P60953	21,3
CPSF1	•	•	Cleavage and polyadenylation specificity factor subunit 1	Q10570	161,0
CSTF2	—	•	Cleavage stimulation factor subunit 2	P33240	61,0
EIF2B4	•	•	Translation initiation factor eIF-2B subunit delta	Q9UI10	57,6
HMGCS1	—	—	Hydroxymethylglutaryl-CoA synthase, cytoplasmic	Q01581	57,2
NOB1	•	—	RNA-binding protein NOB1	Q9ULX3	46,7
PDCD6	•	•	Programmed cell death protein 6	O75340	21,7
PEG10	—	•	Retrotransposon-derived protein PEG10	Q86TG7	80,2
PGRMC2	•	•	Membrane-associated progesterone receptor component 2	O15173	23,8
PLIN3	•	•	Perilipin-3	O60664	47,0
PRPF39	—	—	Pre-mRNA-processing factor 39	Q86UA1	78,4
PSME2	•	•	Proteasome activator complex subunit 2	Q9UL46	27,4
PTGES3	•	—	Prostaglandin E synthase 3	Q15185	18,7
QPRT	—	—	Nicotinate-nucleotide pyrophosphorylase	Q15274	30,8
S100A7	—	—	Protein S100-A7	P31151	11,5
SOD1	—	•	Superoxide dismutase [Cu-Zn]	P00441	16,0
ST13P4	—	—	Putative protein FAM10A4	Q8IZP2	27,4
TIAL1	•	•	Nucleolysin TIAR	Q01085	41,6
TJP2	—	•	Tight junction protein ZO-2	Q9UDY2	134,0
FABP1	—	—	Fatty acid-binding protein. liver	P07148	14,2
NQO1	—	—	NAD(P)H dehydrogenase [quinone] 1	P15559	30,9
SERPINB12	—	—	Serpin B12	Q96P63	46,3
CARM1	•	•	Histone-arginine methyltransferase CARM1	Q86X55	65,8
CTSB	—	•	Cathepsin B	P07858	37,8
DAZAP1	—	•	DAZ-associated protein 1	Q96EP5	43,4
TIA1	—	•	Nucleolysin TIA-1 isoform p40	P31483	43,0
DNAJC10	•	•	DnaJ homolog subfamily C member 10	Q8IXB1	91,0
LGALS8	•	•	Galectin-8	O00214	35,8

Примечание: \* – коэкспрессия генов (коэффициент корреляции Пирсона >0,8), кодирующих CFP1 и его потенциальные партнеры по взаимодействию, согласно portalу GeneFriends [68]; \*\* – белки, которые кофракционируются с CFP1, согласно литературным данным, гель-хроматографического профилирования клеточных лизатов HEK293 [64], HeLa-CCL2 [65], U2OS [66], HeLa-S3 и HEK293 [63]; # – 32 белка осталось после вычитания известных контаминирующих белков по данным CRAPome [67] и белков с несоответствующим CFP1 паттерном субклеточной локализации; всего было выделено из лизата клеток HepG2 144 белка ( $p < 0,01$  и Mascot score >30).

функция — катализ восстановления несвойственных дисульфидных связей в эндоплазматическом ретикулуме, что важно для регуляции оборота аномально свернутых гликопротеинов [83]. Так, нокдаун гена *DNAJC10* приводил к подавлению протеолитического расщепления мутантного

белка SFTPC с аномальной трёхмерной структурой без значимого эффекта в случае белка с последовательностью “дикого типа” [84]. Среди примерно двухсот ББВ с участием DNAJC10, аннотированных в STRINGdb (score >0,4), отсутствуют факторы транскрипции, но можно предположить,



что взаимодействие DNAJC10 с CFP1 способствует правильному образованию дисульфидных связей в последнем [85]. Упоминания о взаимодействиях между катепсином В (CTSB) и CFP1 не встречаются в литературе. Катепсины В, L, S и К в разной степени оказывают влияние на сигнальный путь TGF- $\beta$ /SMAD [86] и CTSB рассматривают в качестве молекулярной мишени для противоопухолевых агентов [87]. Новые данные о роли CTSB в регуляции TGF- $\beta$ /SMAD на животных моделях [88] хорошо согласуются с данными о физических ББВ между представителями групп CXXC- и SMAD-белков, например, взаимодействия CXXC5/SMAD3 [89] и CFP1/SMAD1 [90], и могут говорить об опосредованном взаимодействии катепсина В с CFP1 через SMAD-белки.

## ЗАКЛЮЧЕНИЕ

CXXC-белки функционируют в составе мультибелковых комплексов, играя роль эпигенетических мастер-регуляторов генов-мишеней, в том числе имея потенциал молекулярных “переключателей” сигнальных путей в клетках. CXXC-белки и определенная доля их белковых партнёров формируют высококоннективную сеть ББВ как через прямые связи между собой, так и взаимодействуя опосредованно — через общих белковых партнёров. По литературным данным, индукция экспрессии генов CXXC преимущественно ассоциируется с усилением опухолевых процессов и неблагоприятным прогнозом заболевания. Одним из наименее изученных CXXC-белков в контексте передачи сигналов в опухолевых клетках является мультифункциональный белок CFP1, интерактом которого представлен РНК-связывающими белками; белками, вовлечёнными в сплайсинг РНК; метилтрансферазами гистонов и ко-шаперонами. Перспективы практического использования фундаментальных данных об интерактоме CXXC-белков и белковых комплексов с их участием связаны с идентификацией инновационных противоопухолевых агентов, действующих через модуляцию аффинности и времени жизни клинически значимых ББВ.

## ФИНАНСИРОВАНИЕ

Работа выполнена в рамках Программы фундаментальных научных исследований в Российской Федерации на долгосрочный период (2021-2030 годы) (№ 122030100168-2).

## СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит каких-либо исследований с участием людей или с использованием животных в качестве объектов.

## КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

## ЛИТЕРАТУРА

1. Grechkin M., Logsdon B.A., Gentles A.J., Lee S.-I. (2016) Identifying network perturbation in cancer. *PLoS Comput. Biol.*, **12**(5), e1004888. DOI: 10.1371/journal.pcbi.1004888
2. Sewell J.A., Fuxman Bass J.I. (2017) Cellular network perturbations by disease-associated variants. *Curr. Opin. Syst. Biol.*, **3**, 60-66. DOI: 10.1016/j.coisb.2017.04.009
3. Vidal M., Cusick M.E., Barabási A.-L. (2011) Interactome networks and human disease. *Cell*, **144**(6), 986-998. DOI: 10.1016/j.cell.2011.02.016
4. Cassandri M., Smirnov A., Novelli F., Pitolli C., Agostini M., Malewicz M., Melino G., Raschella G. (2017) Zinc-finger proteins in health and disease. *Cell Death Discov.*, **3**, 17071. DOI: 10.1038/cddiscovery.2017.71
5. Conacci-Sorrell M., McFerrin L., Eisenman R.N. (2014) An overview of MYC and its interactome. *Cold Spring Harb. Perspect. Med.*, **4**(1), a014357. DOI: 10.1101/cshperspect.a014357
6. Lu H., Zhou Q., He J., Jiang Z., Peng C., Tong R., Shi J. (2020) Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduct. Target Ther.*, **5**(1), 213. DOI: 10.1038/s41392-020-00315-3
7. Blackledge N.P., Thomson J.P., Skene P.J. (2013) CpG island chromatin is shaped by recruitment of ZF-CxxC proteins. *Cold Spring Harb. Perspect. Biol.*, **5**(11), a018648. DOI: 10.1101/cshperspect.a018648
8. HGNC: HUGO Gene Nomenclature Committee. <https://www.genenames.org/>
9. Liu K., Min J. (2020) Structural basis for the recognition of non-methylated DNA by the CXXC domain. *J. Mol. Biol.*, **432**(6), 1674-1686. DOI: 10.1016/j.jmb.2019.09.025
10. Xu C., Liu K., Lei M., Yang A., Li Y., Hughes T.R., Min J. (2018) DNA sequence recognition of human CXXC domains and their structural determinants. *Structure*, **26**(1), 85-95.e3. DOI: 10.1016/j.str.2017.11.022
11. Pfam database: CXXC zinc finger domain. <http://pfam.xfam.org/family/pf02008>
12. Kuang Y., Xu H., Lu F., Meng J., Yi Y., Yang H., Hou H., Wei H., Su S. (2021) Inhibition of microRNA let-7b expression by KDM2B promotes cancer progression by targeting EZH2 in ovarian cancer. *Cancer Sci.*, **112**(1), 231-242. DOI: 10.1111/cas.14708
13. Leng X., Wang J., An N., Wang X., Sun Y., Chen Z. (2020) Histone 3 lysine-27 demethylase KDM6A coordinates with KMT2B to play an oncogenic role in NSCLC by regulating H3K4me3. *Oncogene*, **39**(41), 6468-6479. DOI: 10.1038/s41388-020-01449-y
14. Ayaz G., Turan G., Olgun Ç.E., Kars G., Karakaya B., Yavuz K., Demiralay Ö.D., Can T., Muyan M., Yaşar P. (2021) A prelude to the proximity interaction mapping of CXXC5. *Sci. Rep.*, **11**(1), 17587. DOI: 10.1038/s41598-021-97060-6
15. Piñero J., Ramírez-Anguaita J.M., Saüch-Pitarch J., Ronzano F., Centeno E., Sanz F., Furlong L.I. (2020) The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.*, **48**(D1), D845-D855. DOI: 10.1093/nar/gkz1021
16. Rabal O., San José-Enériz E., Agirre X., Sánchez-Arias J.A., de Miguel I., Ordoñez R., Garate L., Miranda E., Sáez E., Vilas-Zornoza A., Pineda-Lucena A., Estella A., Zhang F., Wu W., Xu M., Prosper F., Oyarzabal J. (2021) Design and synthesis of novel epigenetic inhibitors targeting histone deacetylases, DNA methyltransferase 1, and lysine methyltransferase G9a with *in vivo* efficacy in multiple myeloma. *J. Med. Chem.*, **64**(6), 3392-3426. DOI: 10.1021/acs.jmedchem.0c02255

17. Tregnago C., Benetton M., Da Ros A., Borella G., Longo G., Polato K., Francescato S., Biffi A., Pigazzi M. (2021) Novel compounds synergize with venetoclax to target KMT2A-rearranged pediatric acute myeloid leukemia. *Front. Pharmacol.*, **12**, 820191. DOI: 10.3389/fphar.2021.820191
18. Massoth L.R., Hung Y.P., Dias-Santagata D., Onozato M., Shah N., Severson E., Duncan D., Gillespie B.J., Williams N.F., Ross J.S., Vergilio J.-A., Harkins S.K., Glomski K., Nardi V., Zukerberg L.R., Hasserjian R.P., Louissaint A., Williams E.A. (2020) Pan-cancer landscape analysis reveals recurrent KMT2A-MAML2 gene fusion in aggressive histologic subtypes of thymoma. *JCO Precis. Oncol.*, **4**, PO.19.00288. DOI: 10.1200/PO.19.00288
19. Wong W.H., Junck L., Druley T.E., Gutmann D.H. (2019) NF1 glioblastoma clonal profiling reveals KMT2B mutations as potential somatic oncogenic events. *Neurology*, **93**(24), 1067-1069. DOI: 10.1212/WNL.00000000000008623
20. Li H., Liu J.-W., Sun L.-P., Yuan Y. (2017) A meta-analysis of the association between DNMT1 polymorphisms and cancer risk. *Biomed. Res. Int.*, **2017**, 3971259. DOI: 10.1155/2017/3971259
21. Wu H.-X., Chen Y.-X., Wang Z.-X., Zhao Q., He M.-M., Wang Y.-N., Wang F., Xu R.-H. (2019) Alteration in TET1 as potential biomarker for immune checkpoint blockade in multiple cancers. *J. Immunother. Cancer*, **7**(1), 264. DOI: 10.1186/s40425-019-0737-3
22. Lu B., Wei J., Zhou H., Chen J., Li Y., Ye L., Zhao W., Wu S. (2022) Histone H3K36me2 demethylase KDM2A promotes bladder cancer progression through epigenetically silencing RARRES3. *Cell Death Dis.*, **13**(6), 547. DOI: 10.1038/s41419-022-04983-7
23. Wang J., Li T., Li X., Zhang Y., Wang X. (2021) Expression pattern and regulatory effect of lysine-specific demethylase 2A gene in clear cell renal cell carcinoma. *BMC Urol.*, **21**(1), 108. DOI: 10.1186/s12894-021-00867-8
24. Fang Y., Zhang D., Hu T., Zhao H., Zhao X., Lou Z., He Y., Qin W., Xia J., Zhang X., Ye L.-C. (2019) KMT2A histone methyltransferase contributes to colorectal cancer development by promoting cathepsin Z transcriptional activation. *Cancer Med.*, **8**(7), 3544-3552. DOI: 10.1002/cam4.2226
25. Wensheng L., Bo Z., Qiangsheng H., Wenyan X., Shunrong J., Jin X., Quanxing N., Xianjun Y., Xiaowu X. (2019) MBD1 promotes the malignant behavior of gallbladder cancer cells and induces chemotherapeutic resistance to gemcitabine. *Cancer Cell Int.*, **19**, 232. DOI: 10.1186/s12935-019-0948-1
26. Jia G., Song Z., Xu Z., Tao Y., Wu Y., Wan X. (2021) Screening of gene markers related to the prognosis of metastatic skin cutaneous melanoma based on Logit regression and survival analysis. *BMC Med. Genomics*, **14**(1), 96. DOI: 10.1186/s12920-021-00923-0
27. Sun J., Long Y., Peng X., Xiao D., Zhou J., Tao Y., Liu S. (2019) The survival analysis and oncogenic effects of CFP1 and 14-3-3 expression on gastric cancer. *Cancer Cell Int.*, **19**, 225. DOI: 10.1186/s12935-019-0946-3
28. Li P., Ge D., Li P., Hu F., Chu J., Chen X., Song W., Wang A., Tian G., Gu X. (2020) CXXC finger protein 4 inhibits the CDK18-ERK1/2 axis to suppress the immune escape of gastric cancer cells with involvement of ELK1/MIR100HG pathway. *J. Cell Mol. Med.*, **24**(17), 10151-10165. DOI: 10.1111/jcmm.15625
29. Han M., Dai D., Yousafzai N.A., Wang F., Wang H., Zhou Q., Lu H., Xu W., Feng L., Jin H., Wang X. (2017) CXXC4 activates apoptosis through up-regulating GDF15 in gastric cancer. *Oncotarget*, **8**(61), 103557-103567. DOI: 10.18632/oncotarget.21581
30. Fu Y., Wang Z., Luo C., Wang Y., Wang Y., Zhong X., Zheng H. (2020) Downregulation of CXXC Finger protein 4 leads to a tamoxifen-resistant phenotype in breast cancer cells through activation of the Wnt/ $\beta$ -catenin pathway. *Transl. Oncol.*, **13**(2), 423-440. DOI: 10.1016/j.tranon.2019.12.005
31. Baylin S.B., Jones P.A. (2016) Epigenetic determinants of cancer. *Cold Spring Harb. Perspect. Biol.*, **8**(9), a019505. DOI: 10.1101/cshperspect.a019505
32. Fishilevich S., Nudel R., Rappaport N., Hadar R., Plaschkes I., Iny Stein T., Rosen N., Kohn A., Twik M., Safran M., Lancet D., Cohen D. (2017) GeneHancer: Genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)*, **2017**, bax028. DOI: 10.1093/database/bax028
33. GeneCards®: The Human Gene Database. <https://www.genecards.org/>
34. KEGG: Kyoto Encyclopedia of Genes and Genomes. <https://www.kegg.jp/>
35. CGC: Cancer Gene Census. <https://cancer.sanger.ac.uk/cosmic/census>
36. Wang F., Wang L., Xu Z., Liang G. (2013) Identification and analysis of multi-protein complexes in placenta. *PLoS One*, **8**(4), e62988. DOI: 10.1371/journal.pone.0062988
37. Will T., Helms V. (2019) Differential analysis of combinatorial protein complexes with ComplexXChange. *BMC Bioinformatics*, **20**(1), 300. DOI: 10.1186/s12859-019-2852-z
38. Alexeyenko A., Schmitt T., Tjärnberg A., Guala D., Frings O., Sonnhammer E.L.L. (2012) Comparative interactomics with Funcoup 2.0. *Nucleic Acids Res.*, **40**(Database issue), D821-D828. DOI: 10.1093/nar/gkr1062
39. Calderone A., Castagnoli L., Cesareni G. (2013) Mentha: A resource for browsing integrated protein-interaction networks. *Nat. Methods*, **10**(8), 690-691. DOI: 10.1038/nmeth.2561
40. Alonso-López D., Campos-Laborie F.J., Gutiérrez M.A., Lambourne L., Calderwood M.A., Vidal M., de las Rivas J. (2019) APID database: Redefining protein-protein interaction experimental evidences and binary interactomes. *Database (Oxford)*, **2019**, baz005. DOI: 10.1093/database/baz005
41. Licata L., Briganti L., Peluso D., Perfetto L., Iannuccelli M., Galeota E., Sacco F., Palma A., Nardoza A.P., Santonico E., Castagnoli L., Cesareni G. (2012) MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res.*, **40**(Database issue), D857-D861. DOI: 10.1093/nar/gkr930
42. Luck K., Kim D.-K., Lambourne L., Spirohn K., Begg B.E., Bian W., Brignall R., Cafarelli T., Campos-Laborie F.J., Charlotiaux B., Choi D., Coté A.G., Daley M., Deimling S., Desbuleux A., Dricot A., Gebbia M., Hardy M.F., Kishore N., Knapp J.J., Kovács I.A., Lemmens I., Mee M.W., Mellor J.C., Pollis C., Pons C., Richardson A.D., Schlabach S., Teeking B., Yadav A., Babor M., Balcha D., Basha O., Bowman-Colin C., Chin S.-F., Choi S.G., Colabella C., Coppin G., d'Amata C., de Ridder D., de Rouck S., Duran-Frigola M., Ennajaoui H., Goebels F., Goehring L., Gopal A., Haddad G., Hatchi E., Helmy M., Jacob Y., Kassa Y., Landini S., Li R., van Lieshout N., MacWilliams A., Markey D., Paulson J.N., Rangarajan S., Rasla J., Rayhan A., Rolland T., San-Miguel A., Shen Y., Sheykhkarimli D., Sheynkman G.M., Simonovsky E., Taşan M., Tejeda A., Tropepe V., Twizere J.-C., Wang Y., Weatheritt R.J., Weile J., Xia Y., Yang X., Yeager-Lotem E., Zhong Q., Aloy P., Bader G.D., de las Rivas J., Gaudet S., Hao T., Rak J., Tavernier J., Hill D.E., Vidal M., Roth F.P., Calderwood M.A. (2020) A reference map of the human binary protein interactome. *Nature*, **580**(7803), 402-408. DOI: 10.1038/s41586-020-2188-x

43. Kotlyar M., Pastrello C., Malik Z., Jurisica I. (2019) IID 2018 update: Context-specific physical protein-protein interactions in human, model organisms and domesticated species. *Nucleic Acids Res.*, **47**(D1), D581-D589. DOI: 10.1093/nar/gky1037
44. Schweppe D.K., Huttlin E.L., Harper J.W., Gygi S.P. (2018) BioPlex Display: An interactive suite for large-scale AP-MS protein-protein interaction data. *J. Proteome Res.*, **17**(1), 722-726. DOI: 10.1021/acs.jproteome.7b00572
45. Orii N., Ganapathiraju M.K. (2012) Wiki-pi: A web-server of annotated human protein-protein interactions to aid in discovery of protein function. *PLoS One*, **7**(11), e49029. DOI: 10.1371/journal.pone.0049029
46. Alanis-Lobato G., Andrade-Navarro M.A., Schaefer M.H. (2017) HIPPIE v2.0: Enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Res.*, **45**(D1), D408-D414. DOI: 10.1093/nar/gkw985
47. Das J., Yu H. (2012) HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC Syst. Biol.*, **6**, 92. DOI: 10.1186/1752-0509-6-92
48. Szklarczyk D., Gable A.L., Lyon D., Junge A., Wyder S., Huerta-Cepas J., Simonovic M., Doncheva N.T., Morris J.H., Bork P., Jensen L.J., von Mering C. (2019) STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*, **47**(D1), D607-D613. DOI: 10.1093/nar/gky1131
49. Liao Y., Wang J., Jaehnig E.J., Shi Z., Zhang B. (2019) WebGestalt 2019: Gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.*, **47**(W1), W199-W205. DOI: 10.1093/nar/gkz401
50. Giurgiu M., Reinhard J., Brauner B., Dunger-Kaltenbach I., Fobo G., Frishman G., Montrone C., Ruepp A. (2019) CORUM: The comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res.*, **47**(D1), D559-D563. DOI: 10.1093/nar/gky973
51. Szklarczyk D., Gable A.L., Nastou K.C., Lyon D., Kirsch R., Pyysalo S., Doncheva N.T., Legeay M., Fang T., Bork P., Jensen L.J., von Mering C. (2021) The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.*, **49**(D1), D605-D612. DOI: 10.1093/nar/gkaa1074
52. Park S.-J., Yoon B.-H., Kim S.-K., Kim S.-Y. (2019) GENT2: An updated gene expression database for normal and tumor tissues. *BMC Med. Genomics*, **12**(Suppl 5), 101. DOI: 10.1186/s12920-019-0514-7
53. Tang Z., Kang B., Li C., Chen T., Zhang Z. (2019) GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.*, **47**(W1), W556-W560. DOI: 10.1093/nar/gkz430
54. Cheray M., Nadaradjane A., Bonnet P., Routier S., Vallette F.M., Cartron P.-F. (2014) Specific inhibition of DNMT1/CFP1 reduces cancer phenotypes and enhances chemotherapy effectiveness. *Epigenomics*, **6**(3), 267-275. DOI: 10.2217/epi.14.18
55. Skalnik D.G. (2010) The epigenetic regulator Cfp1. *Biomol. Concepts*, **1**(5-6), 325-334. DOI: 10.1515/bmc.2010.031
56. Yang Y., Yang Y., Chan K., Couture J.-F. (2021) Analyzing the impact of CFP1 mutational landscape on epigenetic signaling. *FASEB J.*, **35**(8), e21790. DOI: 10.1096/fj.202100427R
57. Derks S., Bosch L.J.W., Niessen H.E.C., Moerkerk P.T.M., van den Bosch S.M., Carvalho B., Mongera S., Voncken J.W., Meijer G.A., de Bruijn A.P., Herman J.G., van Engeland M. (2009) Promoter CpG island hypermethylation- and H3K9me3 and H3K27me3-mediated epigenetic silencing targets the deleted in colon cancer (DCC) gene in colorectal carcinogenesis without affecting neighboring genes on chromosomal region 18q21. *Carcinogenesis*, **30**(6), 1041-1048. DOI: 10.1093/carcin/bgp073
58. Wang L., Collings C.K., Zhao Z., Cozzolino K.A., Ma Q., Liang K., Marshall S.A., Sze C.C., Hashizume R., Savas J.N., Shilatifard A. (2017) A cytoplasmic COMPASS is necessary for cell survival and triple-negative breast cancer pathogenesis by regulating metabolism. *Genes Dev.*, **31**(20), 2056-2066. DOI: 10.1101/gad.306092.117
59. van de Lagemaat L.N., Flenley M., Lynch M.D., Garrick D., Tomlinson S.R., Kranc K.R., Vernimmen D. (2018) CpG binding protein (CFP1) occupies open chromatin regions of active genes, including enhancers and non-CpG islands. *Epigenetics Chromatin*, **11**(1), 59. DOI: 10.1186/s13072-018-0230-0
60. Lin F., Meng X., Guo Y., Cao Y., Liu W., Xia Q., Hui Z., Chen J., Hong S., Zhang X., Wu C., Wang D., Wang J., Lu L., Qian W., Wei L., Wang L. (2019) Epigenetic initiation of the TH17 differentiation program is promoted by CXXC finger protein 1. *Sci. Adv.*, **5**(10), eaax1608. DOI: 10.1126/sciadv.aax1608
61. Young S.R.L., Skalnik D.G. (2007) CXXC finger protein 1 is required for normal proliferation and differentiation of the PLB-985 myeloid cell line. *DNA Cell Biol.*, **26**(2), 80-90. DOI: 10.1089/dna.2006.0535
62. Chun K.T., Li B., Dobrota E., Tate C., Lee J.-H., Khan S., Haneline L., HogenEsch H., Skalnik D.G. (2014) The epigenetic regulator CXXC finger protein 1 is essential for murine hematopoiesis. *PLoS One*, **9**(12), e113745. DOI: 10.1371/journal.pone.0113745
63. Jiang Y., Zhang H.-Y., Lin Z., Zhu Y.-Z., Yu C., Sha Q.-Q., Tong M.-H., Shen L., Fan H.-Y. (2020) CXXC finger protein 1-mediated histone H3 lysine-4 trimethylation is essential for proper meiotic crossover formation in mice. *Development*, **147**(6), dev183764. DOI: 10.1242/dev.183764
64. Sha Q.-Q., Dai X.-X., Jiang J.-C., Yu C., Jiang Y., Liu J., Ou X.-H., Zhang S.-Y., Fan H.-Y. (2018) CFP1 coordinates histone H3 lysine-4 trimethylation and meiotic cell cycle progression in mouse oocytes. *Nat. Commun.*, **9**(1), 3477. DOI: 10.1038/s41467-018-05930-x
65. Cao W., Guo J., Wen X., Miao L., Lin F., Xu G., Ma R., Yin S., Hui Z., Chen T., Guo S., Chen W., Huang Y., Liu Y., Wang J., Wei L., Wang L. (2016) CXXC finger protein 1 is critical for T-cell intrathymic development through regulating H3K4 trimethylation. *Nat. Commun.*, **7**, 11687. DOI: 10.1038/ncomms11687
66. Sha Q.-Q., Zhu Y.-Z., Xiang Y., Yu J.-L., Fan X.-Y., Li Y.-C., Wu Y.-W., Shen L., Fan H.-Y. (2021) Role of CxxC-finger protein 1 in establishing mouse oocyte epigenetic landscapes. *Nucleic Acids Res.*, **49**(5), 2569-2582. DOI: 10.1093/nar/gkab107
67. Еришов П.В., Мезенцев Ю.В., Яблоков Е.О., Калужский Л.А., Флоринская А.В., Гнеденко О.В., Згода В.Г., Вахрушев И.В., Раева О.С., Ярыгин К.Н., Гилеп А.А., Усанов С.А., Медведев А.Е., Иванов А.С. (2019) Прямой молекулярный фиджинг в лизате клеток HepG2 потенциальных партнёров белок-белковых взаимодействий для белков, кодируемых генами 18-й хромосомы человека. *Биоорганическая Химия*, **45**(2), 155-165. [Ershov P.V., Mezentshev Yu.V., Yablokov E.O., Kaluzhskiy L.A., Florinskaya A.V., Gnedenko O.V., Zgoda V.G., Vakhrushev I.V., Raeva O.S., Yarygin K.N., Gilep A.A., Usanov S.A., Medvedev A.E., Ivanov A.S. (2018) Direct molecular fishing of potential partners of protein-protein interactions in the HepG2 cells

- lysate that involve proteins encoded by genes of human chromosome 18. Russian Journal of Bioorganic Chemistry, **44**(6), 759-768. DOI: 10.1134/S0132342319010068
68. Havugimana P.C., Hart G.T., Nepusz T., Yang H., Turinsky A.L., Li Z., Wang P.I., Boutz D.R., Fong V., Phanse S., Babu M., Craig S.A., Hu P., Wan C., Vlasblom J., Dar V.-N., Bezginov A., Clark G.W., Wu G.C., Wodak S.J., Tillier E.R.M., Paccanaro A., Marcotte E.M., Emili A. (2012) A census of human soluble protein complexes. *Cell*, **150**(5), 1068-1081. DOI: 10.1016/j.cell.2012.08.011
69. Heusel M., Bludau I., Rosenberger G., Hafen R., Frank M., Banaei-Esfahani A., van Drogen A., Collins B.C., Gstaiger M., Aebersold R. (2019) Complex-centric proteome profiling by SEC-SWATH-MS. *Mol. Syst. Biol.*, **15**(1), e8438. DOI: 10.15252/msb.20188438
70. Heusel M., Frank M., Köhler M., Amon S., Frommelt F., Rosenberger G., Bludau I., Aulakh S., Linder M.I., Liu Y., Collins B.C., Gstaiger M., Kutay U., Aebersold R. (2020) A global screen for assembly state changes of the mitotic proteome by SEC-SWATH-MS. *Cell Syst.*, **10**(2), 133-155.e6. DOI: 10.1016/j.cels.2020.01.001
71. Larance M., Kirkwood K.J., Tinti M., Brenes Murillo A., Ferguson M.A.J., Lamond A.I. (2016) Global membrane protein interactome analysis using *in vivo* crosslinking and mass spectrometry-based protein correlation profiling. *Mol. Cell Proteomics*, **15**(7), 2476-2490. DOI: 10.1074/mcp.O115.055467
72. Mellacheruvu D., Wright Z., Couzens A.L., Lambert J.-P., St-Denis N.A., Li T., Miteva Y.V., Hauri S., Sardi M.E., Low T.Y., Halim V.A., Bagshaw R.D., Hubner N.C., Al-Hakim A., Bouchard A., Faubert D., Fermin D., Dunham W.H., Goudreau M., Lin Z.-Y., Badillo B.G., Pawson T., Durocher D., Coulombe B., Aebersold R., Superti-Furga G., Colinge J., Heck A.J.R., Choi H., Gstaiger M., Mohammed S., Cristea I.M., Bennett K.L., Washburn M.P., Raught B., Ewing R.M., Gingras A.-C., Nesvizhskii A.I. (2013) The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods*, **10**(8), 730-736. DOI: 10.1038/nmeth.2557
73. van Dam S., Craig T., de Magalhães J.P. (2015) GeneFriends: A human RNA-seq-based gene and transcript co-expression database. *Nucleic Acids Res.*, **43**(Database issue), D1124-D1132. DOI: 10.1093/nar/gku1042
74. Armaos A., Zacco E., Sanchez de Groot N., Tartaglia G.G. (2021) RNA-protein interactions: Central players in coordination of regulatory networks. *Bioessays*, **43**(2), e2000118. DOI: 10.1002/bies.202000118
75. Du X., Xiao R. (2020) An emerging role of chromatin-interacting RNA-binding proteins in transcription regulation. *Essays Biochem.*, **64**(6), 907-918. DOI: 10.1042/EBC20200004
76. Xiao R., Chen J.-Y., Liang Z., Luo D., Chen G., Lu Z.J., Chen Y., Zhou B., Li H., Du X., Yang Y., San M., Wei X., Liu W., Lécuyer E., Graveley B.R., Yeo G.W., Burge C.B., Zhang M.Q., Zhou Y., Fu X.-D. (2019) Pervasive chromatin-RNA binding protein interactions enable RNA-based regulation of transcription. *Cell*, **178**(1), 107-121.e18. DOI: 10.1016/j.cell.2019.06.001
77. Wang Q., Guo Y., Wang W., Liu B., Yang G., Xu Z., Li J., Liu Z. (2021) RNA binding protein DAZAP1 promotes HCC progression and regulates ferroptosis by interacting with SLC7A11 mRNA. *Exp. Cell Res.*, **399**(1), 112453. DOI: 10.1016/j.yexcr.2020.112453
78. Oughtred R., Rust J., Chang C., Breitkreutz B.-J., Stark C., Willems A., Boucher L., Leung G., Kolas N., Zhang F., Dolma S., Coulombe-Huntington J., Chatr-Aryamontri A., Dolinski K., Tyers M. (2021) The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.*, **30**(1), 187-200. DOI: 10.1002/pro.3978
79. Malovannaya A., Lanz R.B., Jung S.Y., Bulynko Y., Le N.T., Chan D.W., Ding C., Shi Y., Yucer N., Krenciute G., Kim B.-J., Li C., Chen R., Li W., Wang Y., O'Malley B.W., Qin J. (2011) Analysis of the human endogenous coregulator complexome. *Cell*, **145**(5), 787-799. DOI: 10.1016/j.cell.2011.05.006
80. Zhao W., Zhao J., Hou M., Wang Y., Zhang Y., Zhao X., Zhang C., Guo D. (2014) HuR and TIA1/TIAL1 are involved in regulation of alternative splicing of SIRT1 pre-mRNA. *Int. J. Mol. Sci.*, **15**(2), 2946-2958. DOI: 10.3390/ijms15022946
81. Cheng D., Côté J., Shaaban S., Bedford M.T. (2007) The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. *Mol. Cell*, **25**(1), 71-83. DOI: 10.1016/j.molcel.2006.11.019
82. Greulich F., Wierer M., Mechtidou A., Gonzalez-Garcia O., Uhlenhaut N.H. (2021) The glucocorticoid receptor recruits the COMPASS complex to regulate inflammatory transcription at macrophage enhancers. *Cell Rep.*, **34**(6), 108742. DOI: 10.1016/j.celrep.2021.108742
83. Daverkausen-Fischer L., Pröls F. (2021) Dual topology of co-chaperones at the membrane of the endoplasmic reticulum. *Cell Death Discov.*, **7**(1), 203. DOI: 10.1038/s41420-021-00594-x
84. Dong M., Bridges J.P., Apsley K., Xu Y., Weaver T.E. (2008) ERdj4 and ERdj5 are required for endoplasmic reticulum-associated protein degradation of misfolded surfactant protein C. *Mol. Biol. Cell*, **19**(6), 2620-2630. DOI: 10.1091/mbc.e07-07-0674
85. Lee S.F., Li L., Jalal N., Halperin S.A. (2021) Identification of a thiol-disulfide oxidoreductase (SdbA) catalyzing disulfide bond formation in the superantigen SpeA in *Streptococcus pyogenes*. *J. Bacteriol.*, **203**(17), e0015321. DOI: 10.1128/JB.00153-21
86. Zhang X., Zhou Y., Yu X., Huang Q., Fang W., Li J., Bonventre J.V., Sukhova G.K., Libby P., Shi G.-P. (2019) Differential roles of cysteinyl cathepsins in TGF- $\beta$  signaling and tissue fibrosis. *iScience*, **19**, 607-622. DOI: 10.1016/j.isci.2019.08.014
87. Gondi C.S., Rao J.S. (2013) Cathepsin B as a cancer target. *Expert Opin. Ther. Targets*, **17**(3), 281-291. DOI: 10.1517/14728222.2013.740461
88. Nettesheim A., Shim M.S., Dixon A., Raychaudhuri U., Gong H., Liton P.B. (2020) Cathepsin B localizes in the caveolae and participates in the proteolytic cascade in trabecular meshwork cells. Potential new drug target for the treatment of glaucoma. *J. Clin. Med.*, **10**(1), E78. DOI: 10.3390/jcm10010078
89. Colland F., Jacq X., Trouplin V., Mouglin C., Groizeleau C., Hamburger A., Meil A., Wojcik J., Legrain P., Gauthier J.-M. (2004) Functional proteomics mapping of a human signaling pathway. *Genome Res.*, **14**(7), 1324-1332. DOI: 10.1101/gr.2334104
90. Baas R., Sijm A., van Teeffelen H.A.A.M., van Es R., Vos H.R., Marc Timmers H.T. (2016) Quantitative proteomics of the SMAD (Suppressor of Mothers Against Decapentaplegic) transcription factor family identifies importin 5 as a bone morphogenic protein receptor SMAD-specific importin. *J. Biol. Chem.*, **291**(46), 24121-24132. DOI: 10.1074/jbc.M116.748582

Поступила в редакцию: 27. 09. 2022.  
После доработки: 24. 10. 2022.  
Принята к печати: 24. 10. 2022.

**INTERACTOMICS OF CXXC PROTEINS INVOLVED  
IN EPIGENETIC REGULATION OF GENE EXPRESSION**

*P.V. Ershov, E.O. Yablokov\*, Y.V. Mezentssev, A.S. Ivanov*

Institute of Biomedical Chemistry,  
10 bld. 8 Pogodinskaya str., Moscow, 119121 Russia; \*e-mail: evgeyablokov1988@mail.ru

Regulation of gene expression is an extremely complex and multicomponent biological phenomenon. Proteins containing the CXXC-domain “zinc fingers” (CXXC-proteins) are master regulators of expression of many genes and have conserved functions of methylation of DNA bases and histone proteins. CXXC proteins function as a part of multiprotein complexes, which indicates the fundamental importance of studying post-translational regulation through modulation of the protein-protein interaction spectrum (PPI) in both normal and pathological conditions. In this paper we discuss general aspects of the involvement of CXXC proteins and their protein partners in neoplastic processes, both from the literature data and our own studies. Special attention is paid to recent data on the particular interactomics of the CFP1 protein encoded by the *CXXCI* gene located on the human chromosome 18. CFP1 is devoid of enzymatic activity and implements epigenetic regulation of expression through binding to chromatin and a certain spectrum of PPIs.

**Key words:** interactomics; CXXC proteins; molecular fishing; SEC-fractionation; epigenetic regulation

**Funding.** The work was carried out within the framework of the Program of Fundamental Scientific Research in the Russian Federation for a long-term period (2021-2030) (No. 122030100168-2).

Received: 27.09.2022; revised: 24.10.2022; accepted: 24.10.2022.