

УДК 577.218

## ГОМОЛОГИ БЕЛКА Xmas-2, ОСНОВНОГО КОМПОНЕНТА КОМПЛЕКСА ЭКСПОРТА мРНК TREX-2

© 2020 г. Е. Н. Набирочкина<sup>1</sup>, Д. В. Копытова<sup>1,\*</sup>

Представлено академиком РАН Г.П. Георгиевым

Поступило 17.07.2020 г.

После доработки 10.08.2020 г.

Принято к публикации 10.08.2020 г.

Комплекс TREX-2 отвечает за общий экспорт мРНК из ядра в цитоплазму у эукариот. Основным белком комплекса TREX-2 *D. melanogaster* является белок Xmas-2. Его гомологами у дрожжей и человека являются белки Sac3 и GANP соответственно. Все три белка содержат высококонсервативный домен Sac3-GANP, необходимый для взаимодействия комплекса TREX-2 с мРНК и другим белком комплекса, PCID2. Используя для поиска последовательность домена семейства Sac3-GANP, мы нашли двух гомологов Xmas-2 у *D. melanogaster*. Эти белки имеют общий домен, отвечающий за взаимодействие с белком PCID2 и с РНК, и присутствуют у других эукариот. Функция этих белков неизвестна, но, исходя из их структурной организации, можно предположить, что они взаимодействуют с нуклеиновыми кислотами.

*Ключевые слова:* TREX-2, Xmas-2, экспорт мРНК, транскрипция

**DOI:** 10.31857/S2686738920060190

Комплекс TREX-2 является основным комплексом экспорта мРНК из ядра в цитоплазму. Этот комплекс был обнаружен у дрожжей [1] и позже найден у других организмов [2–6]. Основными белками комплекса TREX-2 *D. melanogaster* являются: PCID2, ENY2 и Xmas-2, который играет роль скелета, с которым ассоциированы другие белки комплекса. Гомологи Xmas-2 у дрожжей и человека называются Sac3 и GANP соответственно, и имеют общий домен, названный доменом семейства Sac3-GANP. При исследовании гомологов Xmas-2 у дрожжей, методами кристаллографии показано, что этот домен семейства Sac3-GANP отвечает за связывание с другим белком TREX-2 комплекса – PCID2 [4]. Вместе эти белки формируют общую поверхность, отвечающую за связывание комплекса с РНК. С-концевая последовательность Xmas-2 отвечает за взаимодействие с другими белками комплекса TREX-2 и комплексом ядерной поры (NPC) [5, 7]. Таким обра-

зом, Xmas-2 является важным белком комплекса TREX-2.

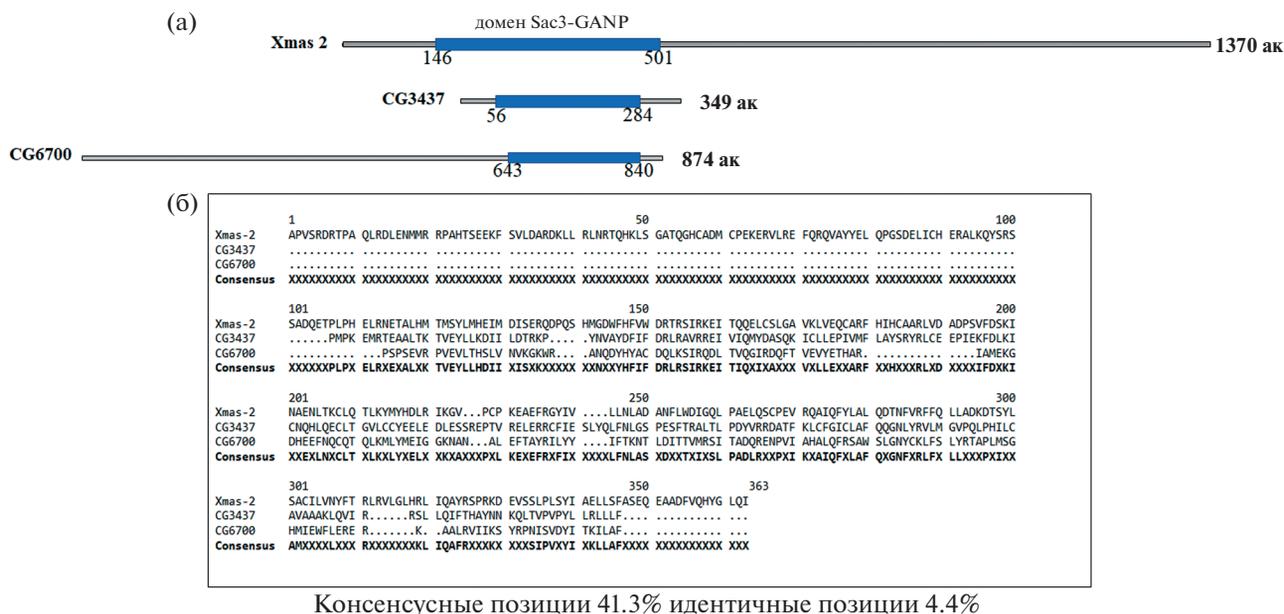
Ранее мы показали, что у *D. melanogaster* существует альтернативная изоформа Xmas-2, названный нами Xmas. Xmas имеет более длинную аминокислотную последовательность, чем белок Xmas-2 за счет дополнительной С-концевой последовательности. Интересно, что данная С-концевая последовательность также присутствует в клетках дрозофилы, как отдельная изоформа (имеет название Xmas-1) [8]. Белок Xmas, так же, как и его гомологи у других организмов, имеет домен семейства Sac3-GANP.

Нами далее был продолжен поиск гомологов Xmas-2, содержащих домен Sac3-GANP. Был проведен поиск белков, содержащих последовательности, гомологичные домену семейства Sac3-GANP в базе данных Flybase (flybase.org). У *D. melanogaster* были обнаружены два паралога Xmas-2, белки CG3437 и CG6700, имеющие общий домен семейства Sac3-GANP. Консенсус между доменами Sac3-GANP для всех трех белков 41.3%, идентичность – 4.4% (рис. 1).

Один из этих паралогов, CG3437, состоит практически только из домена Sac3-GANP и короткой N-концевой последовательности. Другой

<sup>1</sup> Институт биологии гена Российской академии наук, Москва, Россия

\*e-mail: d\_dmitrieva@mail.ru



**Рис. 1.** Гомологи Xmas-2 у *Drosophila melanogaster*. (а) Последовательности гомологов Xmas-2. Синим цветом выделен домен Sac3-GANP. (б) Результат анализа гомологичности последовательностей CG3437 и CG6700 (консенсус – 41.3%, идентичность – 4.4%).

паралог, CG6700, кроме домена Sac3-GANP, имеет длинную N-концевую последовательность.

Каждый из трех белков дрозофилы, имеющий Sac3-GANP домен имеет гомологичные белки в других организмах. Ортологи Xmas-2 *D. melanogaster* существуют у других эукариот (рис. 2). Все они имеют общий домен Sac3-GANP и длинную C-концевую область. Гомолог Xmas-2 человека – GANP также имеет протяженную N-концевую последовательность. Интересно, что C-концевая область белков значительно различается у различных видов и, возможно, определяет разный состав комплекса TREX-2 у различных организмов. Xmas-2 единственный из гомологов семейства, функции которого известны у *D. melanogaster*. Xmas-2 экспрессируется на высоком уровне на ранней эмбриональной стадии и в семенниках взрослых самцов мух. Все ортологи Xmas-2 участвуют в транспорте мРНК из ядра в цитоплазму.

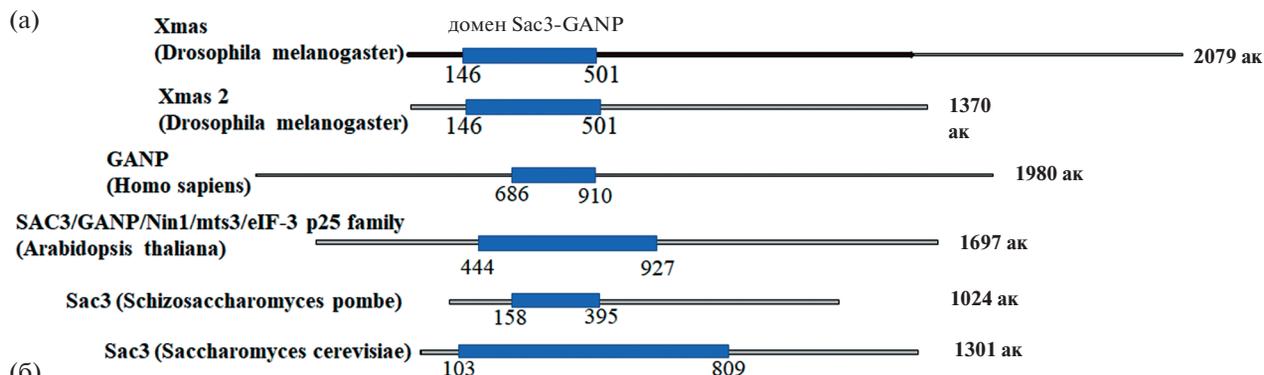
Гомологи CG3437 дрозофилы также присутствуют у других эукариот (рис. 3). Все они содержат домен семейства Sac3 – GANP и непродолжительные N- и C-концевые последовательности. Показано, что гомолог CG3437 человека – Sac3 domain-containing protein 1, является прогностическим маркером для ряда онкологических заболеваний [9–11]. Однако, функции его не изучены. Этот белок также экспрессируется на высо-

ком уровне на ранней эмбриональной стадии развития мухи и в яичниках взрослых самок мух.

Гомологи CG6700 дрозофилы, также были обнаружены у других эукариот (рис. 4), функции этих белков еще не исследованы. Однако показано, что SAC3A, гомолог CG6700 у *Arabidopsis thaliana*, входит в состав TREX-2 вместе с гомологом Xmas-2 – SAC3B [3]. Функция этого белка в составе комплекса пока неизвестна. Этот белок экспрессируется на всех стадиях развития мухи. Наибольший уровень экспрессии этого белка выявляется в мозге взрослых мух.

Таким образом, белки семейства Xmas-2 и его гомологи присутствуют у различных видов эукариот (дрожжи, растения, человек) [1–6]. Для всех белков семейства характерно наличие домена Sac3-GANP. У Xmas-2 этот домен отвечает за связывание с РНК и белками комплекса TREX-2. Возможно, этот домен у гомологов также отвечает за связывание с РНК/ДНК и белок-белковые взаимодействия. Изучение функций этих гомологов – предмет дальнейших исследований.

Гомологи CG3437 у всех организмов практически полностью состоят из домена семейства и не имеют длинной C-концевой последовательности Xmas-2, которая отвечает за связывание с ENY2 и комплексом ядерной поры. Таким образом, можно предположить, что гомологи CG3437 не способны связывать ядерную пору. У гомологов CG6700 имеют протяженные N-концевые после-

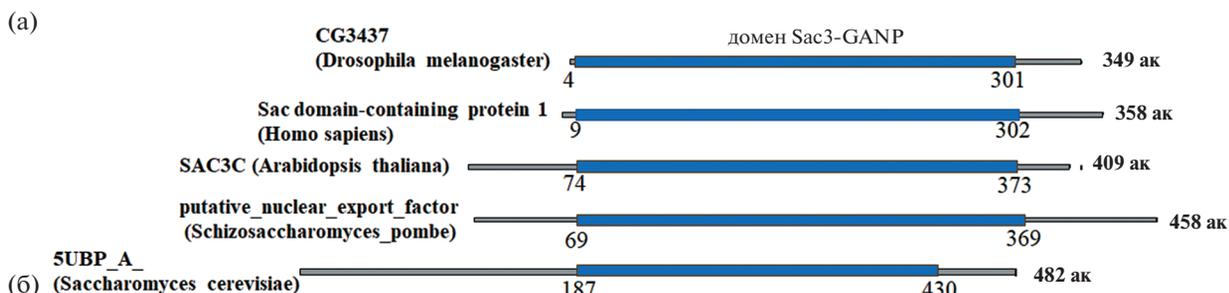


(б)

Xmas-2_(Drosophila_melanogaster)	1	PLPHELRLNET ALHMTMSYLM HEIMDISERQ DPQSHYGWDF HFVMDTRTSI RKEITQQELC SLGAVKLVQEQ CARFHIIHCAA RLVDADPS... VFDKIN	50
GANP_(Homo_sapiens)	1	PLPHELRLPL VLSRTHDYLL TQIMDQKE... GS.LRDWY DFVMDTRTSI RKDITQQELC DPLTVSLIEK CTRFHIIHCAH FMCEEPSM... SFDKIN	50
SAC3/GANP/Nin1/mts3/eIF-3_p25_(Arabidopsis_thaliana)	1	REAILIRPMP ILQNTMEYLL SLLDRPYN... EN.FLGMV NFLMDMRRAI RMDLRMQHIF NQEAITLLEQ MIRLHIIAMH ELCEYTKGEG FSEGFDAHLN	50
SAC3_(Schizosaccharomyces_pombe)	1	PLPSDVRPPP VLKKSLOYLV DKIVCGDDP... L.ENTH FFVDRTRTSI RQDFTLQNCR DLDAVACNER IARYHILLCIH QLCEKQ... FSAQQE	50
SAC3_(Saccharomyces_cerevisiae)	1	PLPSDVRPPH ILVKTLDYIV DNLLTTL... ESEG .FLMDMRRSI RQDFTYQNS GPEAVDCNER IVRHILLLHM IMKSNV... FSLQQE	50
Consensus	1	PLPKDLRPPX ILXKTMDYLV XXIXXXEXX XXXXXXXXX XFVMDTRTSI RDXITXQLX XEAVXLER IXRFHILCIH XLCEXXXXX XXXXFDAXIN	50
Xmas-2_(Drosophila_melanogaster)	101	AENLTK.CLQ TLKYMHDLR IKGVPCPEA EFRGYVLLN LADAN... FLMDIQQ LPAELQSCPE VRQAIQFVLA LQDTN... FVDFKIN	150
GANP_(Homo_sapiens)	101	NEIMTK.CLQ SLKEMEQDLR NKGWFCASEA EFGQYVLLS LNKGD... ILREVQQ FHPAVRNSSE VKFAVQFAA LNSNN... SFDKIN	150
SAC3/GANP/Nin1/mts3/eIF-3_p25_(Arabidopsis_thaliana)	101	IEQMK.TSV ELFQYDDHR KKGITVPEK EFRGYVLLK LDKHPGVKVE PSELSLDLAN MTEPIRQTSV VLFARNVAAA CRTGN... FVDFKIN	150
SAC3_(Schizosaccharomyces_pombe)	101	VEQLRKGILQ SLCEFYDDLK KVKIRCPNEP EFRSAYIITH LRDPD... VVRSQSI LPIEIFDDQR VQLALRLSAL AQKNRERVGH ILRPNTAACP	150
SAC3_(Saccharomyces_cerevisiae)	101	LEQLHK.SLI TLSEIYDDVR SSGGTCPEA EFRAYALLSK TRDPQ... YDENIQR LPKHIFQDKL VQMLCFRRV ISMSAYTERG... FVKTEHCL	150
Consensus	101	IEQLXXXXLQ SLXEHYDDLK XKGIXCPXEA EFRGYVLLX LDXXXXXXX XXXILXDIQQ LPXKIXQXXE VXXALQXXXL LQXXXXXXX XXXXXXXXX	150
Xmas-2_(Drosophila_melanogaster)	201	.FVRFQQL ADKDTSYLSA CILWYFTRL RVLGLHRLIQ AYRSPKDEV SSLPLSYIAE LLSFASEQEA ADFVQHYGLQ I	250
GANP_(Homo_sapiens)	201	.FVRFQQL QS..ASYLNA CLLHCYFSQI RKDALRALNF AYTVS.TQRS TIFPLDGVWR MLFR... FVDFKIN	250
SAC3/GANP/Nin1/mts3/eIF-3_p25_(Arabidopsis_thaliana)	201	.FIAFFRLA RK..ASYLQA CLMHAHFKL RTQALASLHS GLQIN... QGLPVSQMSN WIG... FVDFKIN	250
SAC3_(Schizosaccharomyces_pombe)	201	NLYTRFFKLV QSPAVTYLMA CLESHFMSI RKGALKAMRK AFMSAHAN... FPCGDLKR ILHF... FVDFKIN	250
SAC3_(Saccharomyces_cerevisiae)	201	NFYARFFQLM QSPSLPLLMG FFLQMLTLDI RPYALARLHS TLNKKHKP... IPIFYLEN MLF... FVDFKIN	250
Consensus	201	XXFVRFKLV QSXXXSYLXA CLLXXHFSXI RXXALRALXX AYXXXXXXX XLPLXLLXX LXXMLXFXX XXXXXXXXX X	250

Консенсусные позиции 53% идентичные позиции 11.4%

Рис. 2. Ортологи Xmas-2. (а) Последовательности гомологов Xmas-2 у разных видов эукариот. Синим цветом выделен домен Sac3-GANP. (б) Результат анализа гомологичности последовательностей ортологов Xmas-2. (консенсус – 53%, идентичность – 11.4%).

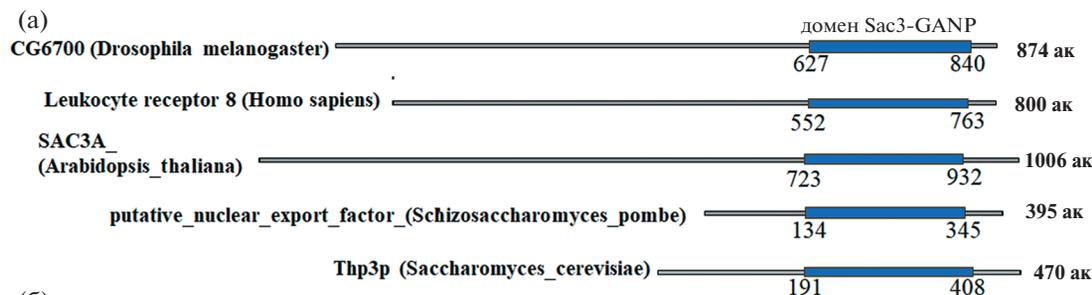


(б)

CG3437_(Drosophila_melanogaster)	101	.....	150	.....	200
Sac3_domain-containing_protein_1_(Homo_sapiens)	101	.....	150	.....	200
SAC3C_(Arabidopsis_thaliana)	101	.....	150	.....	200
putative_nuclear_export_factor_(Schizosaccharomyces_pombe)	101	.....	150	.....	200
SUBP_A_(Saccharomyces_cerevisiae)	101	.....	150	.....	200
Consensus	101	.....	150	.....	200
CG3437_(Drosophila_melanogaster)	201	.....	250	.....	300
Sac3_domain-containing_protein_1_(Homo_sapiens)	201	.....	250	.....	300
SAC3C_(Arabidopsis_thaliana)	201	.....	250	.....	300
putative_nuclear_export_factor_(Schizosaccharomyces_pombe)	201	.....	250	.....	300
SUBP_A_(Saccharomyces_cerevisiae)	201	.....	250	.....	300
Consensus	201	.....	250	.....	300
CG3437_(Drosophila_melanogaster)	301	.....	350	.....	400
Sac3_domain-containing_protein_1_(Homo_sapiens)	301	.....	350	.....	400
SAC3C_(Arabidopsis_thaliana)	301	.....	350	.....	400
putative_nuclear_export_factor_(Schizosaccharomyces_pombe)	301	.....	350	.....	400
SUBP_A_(Saccharomyces_cerevisiae)	301	.....	350	.....	400
Consensus	301	.....	350	.....	400

Консенсусные позиции 49.6% идентичные позиции 3.6%

Рис. 3. Гомологи CG3437. (а) Последовательности гомологов CG3437 у разных видов эукариот. Синим цветом выделен домен Sac3-GANP. (б) Результат анализа гомологичности последовательностей ортологов CG3437 (консенсус – 49.6%, идентичность – 3.6%).



	1	50	100
CG6700 (Drosophila melanogaster)	GTCRDLEKSF LRLTKAPSPS EVRPVEVLTH SLVNVKGMIR AN.QDYHYAC DQLKSRQDL TVQGIIRDQFT VEVYETHARI AMEKGDHEEF NQCQTQLKML		
Leukocyte_receptor_8 (Homo sapiens)	GTCRDITKHY LRLTCAPDPS TVRPVAVLKK SLCHVKCKMK EK.QDYAFAC EQMKSRQDL TVQGIIRTEFT VEVYETHARI ALEKGDHEEF NQCQTQLKSL		
SAC3A (Arabidopsis thaliana)	GTCQEIEKRY LRLTSAPDPA TVRPEDVLEK ALIMVQDSQK ...NYLFKC DQLKSRQDL TVQRINHLT AKVYETHARL ALEAGDPEY NQCLSQLKTL		
putative_nuclear_export_factor_ (Schizosaccharomyces pombe)	GTCQEIEKRY LRLTSAPDPA TVRPEDVLEK ALIMVQDSQK ...NYLFKC DQLKSRQDL TVQRINHLT AKVYETHARL ALEAGDPEY NQCLSQLKTL		
Thp3p (Saccharomyces cerevisiae)	GRQCILEKSY LRLTSEPNPD LIRPPNLLQK MYCLLMDKYQ SKTATYTYLC DQFKSMRQDL RVQMIENSFT IKVYQTHARI ALENGDLXEF NQCQRIZHAL		
Consensus	GTCQEIEKRY LRLTSAPDPA TVRPDXVLXK ALIMVQDSQK XXXXXYFXC DQLKSRQDL TVQXINXFT VKVYETHARI ALEKGDLEEF NQCQSQKTLT		
	101	150	200
CG6700 (Drosophila melanogaster)	YHEIGGKNAN ALEFTAYRIL YYIFTKNTL. DITTVMSRIT ADQR...ENP VIAHALQFRS ANSLGNVCKL FSLYRTAPLM SGMHIEWFLR RERKAALRVI		
Leukocyte_receptor_8 (Homo sapiens)	YAEN..LPGN VGEFTAYRIL YYIFTKNSG. DITTELAYLT RELK...ADP CVAHALALRT ANHALGNVHRF FRLYCHAPCM SGYLVDFAD RERKVALKAM		
SAC3A (Arabidopsis thaliana)	YAEQ..VEGC SLEFAAYSLL YITLHNNNR ELLSMSRSL EEDK...KDE AVRHALSVRA AVTSGNVYWF FRLYKTAPNM NSCLMDLVE KHRYKAVNFM		
putative_nuclear_export_factor_ (Schizosaccharomyces pombe)	YAEQ..VEGC SLEFAAYSLL YITLHNNNR ELLSMSRSL EEDK...KDE AVRHALSVRA AVTSGNVYWF FRLYKTAPNM NSCLMDLVE KHRYKAVNFM		
Thp3p (Saccharomyces cerevisiae)	FGNPITPKKS YSEFICYSVL YSHLTEDYPS ISHLKLLID DGSSELEDE HVKINFELSD MKLVGNVYWF MKNYLKLHF EKCLINSFLN LEKILFILTII		
Consensus	YAEXXXXXGK SLEFXAYSIL YXILTXNXXX DIXSXXSXLX EEXXXXXDE XVRHALSLRA AXTLGMVYWF FRLYKTAPXM XXCLIDXFLE KERXXALXXM		
	201	219	
CG6700 (Drosophila melanogaster)	IKSVRPNISV DYITKILAF		
Leukocyte_receptor_8 (Homo sapiens)	IKTRPALPV SYLQAELEF		
SAC3A (Arabidopsis thaliana)	SRSCRPTIPV SVIQVLGF		
putative_nuclear_export_factor_ (Schizosaccharomyces pombe)	SRSCRPTIPV SVIQVLGF		
Thp3p (Saccharomyces cerevisiae)	KCSYN.QVNL DFIKSEFNF		
Consensus	KXSVRXPXIPV SVIXXVLEF		

**Рис. 4.** Гомологи CG6700. (а) Последовательности гомологов CG6700 у разных видов эукариот. Синим цветом выделен домен Sac3-GANP. (б) Результат анализа гомологичности последовательностей гомологов CG6700 (консенсус – 76.3%, идентичность – 22.8%).

довательности, функция которых неясна. Эти последовательности содержат длинные гидрофобные области.

Интересно, что только CG6700 имеет достаточно высокий уровень экспрессии на различных стадиях развития мухи и во всех тканях взрослых мух, в то время как Xmas-2 и CG3437 на высоком уровне экспрессируются только на ранней эмбриональной стадии развития мухи и в гонадах взрослых мух (Xmas-2 в тканях семенников, CG3437 в тканях яичников).

Все эти отличия в последовательностях гомологов Xmas-2 и их распределении должны обеспечивать различные функции их содержащих комплексов, и, возможно, разную локализацию в клетке. Высокая эволюционная консервативность всех трех белков, содержащих Sac3-GANP домен, указывает на их важную функцию у эукариот.

## СПИСОК ЛИТЕРАТУРЫ

1. Fischer T., Strasser K., Racz A., Rodriguez-Navarro S., Oppizzi M., Ihrig P., Lechner J., and Hurt E. // The EMBO journal. 2002. V. 21. № 21. P. 5843–5852.
2. Kurshakova M.M., Krasnov A.N., Kopytova D. V., Shidlovskii Y. V., Nikolenko J. V., Nabirochkina E.N., Spohner D., Schultz P., Tora L., and Georgieva S.G. // The EMBO journal. 2007. V. 26. № 24. P. 4956–4965.
3. Lu Q., Tang X., Tian G., Wang F., Liu K., Nguyen V., Kothalmi S.E., Keller W.A., Tsang E.W.T., Harada J.J., Rothstein S.J., and Cui Y. // The Plant journal : for cell and molecular biology. 2010. V. 61. № 2. P. 259–270.
4. Ellisdon A.M., Dimitrova L., Hurt E., and Stewart M. // Nature structural & molecular biology. 2012. V. 19. № 3. P. 328–336.
5. Jani D., Lutz S., Marshall N.J., Fischer T., Kohler A., Ellisdon A.M., Hurt E., and Stewart M. // Molecular cell. 2009. V. 33. № 6. P. 727–737.
6. Wickramasinghe V.O., Stewart M., and Laskey R.A. // Nucleus (Austin, Tex.). 2010. V. 1. № 5. P. 393–396.
7. Jani D., Valkov E., and Stewart M. // Nucleic acids research. 2014. V. 42. № 10. P. 6686–6697.
8. Копытова Д.В., Ильин Ю.В., Набирочкина Е.Н. // ДАН. 2018. Т. 479. № 3. С. 340–342.
9. Liu A.-G., Zhong J.-C., Chen G., He R.-Q., He Y.-Q., Ma J., Yang L.-H., Wu X.-J., Huang J.-T., Li J.-J., Mo W.-J., and Qin X.-G. // International journal of oncology. 2020. V. 57. № 1. P. 122–138.
10. Wang X., Li G., Luo Q., and Gan C. // Oncology letters. 2018. V. 15. № 6. P. 8983–8990.
11. Fan J., Yan D., Teng M., Tang H., Zhou C., Wang X., Li D., Qiu G., and Peng Z. // Clinical cancer research : an official journal of the American Association for Cancer Research. 2011. V. 17. № 9. P. 2908–2918.

## THE Xmas-2 HOMOLOGS, THE MAIN COMPONENT OF THE TREX-2 mRNA EXPORT COMPLEX

E. N. Nabirochkina<sup>a</sup> and D. V. Kopytova<sup>a,#</sup>

<sup>a</sup> *Institute of gene biology Russian Academy of Science, Moscow, Russian Federation*

<sup>#</sup>*e-mail: d\_dmitrieva@mail.ru*

Presented by academician of the RAS G.P. Georgiev

TREX-2 complex is responsible for general mRNA export from nucleus to cytoplasm in eukaryote. The main protein of TREX-2 complex of *D. melanogaster* is protein Xmas-2. Its homologues in yeast and humans are Sac3 and GANP proteins respectively. All three proteins contain the highly conserved domain Sac3-GANP, which is essential for interaction of TREX-2 complex with mRNA and another protein of the complex, PCID2. We identified two Xmas-2 homologs in *D. melanogaster*, using the Sac3-GANP family domain sequence. These proteins have a common domain responsible for interaction with PCID2 protein and RNA, and are present in other eukaryotes. The function of these proteins is unknown, but based on their structural organization, we can assume that they interact with nucleic acids.

*Keywords:* TREX-2, Xmas-2, mRNA export, transcription