



УДК 577.218

КЛОНИРОВАНИЕ кДНК ТРЕХ НОВЫХ ГОМЕОБОКССОДЕРЖАЩИХ ГЕНОВ КЛАССА *Anf* ИЗ ЧЕЛОВЕКА, КУРИЦЫ И ИСПАНСКОГО ТРИТОНА

© 1998 г. О. В. Казанская, Г. В. Ермакова, М. Паннезе*, Е. Бончинелли*, А. Г. Зарайский[#]Институт биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,
117871, Москва, ГСП-7, ул. Миклухо-Маклая, 16/10;

*Институт при Госпитале Сан Рафаэле, Милан

Поступила в редакцию 22.05.97 г. Принята к печати 25.11.97 г.

На основе анализа кДНК трех новых гомеобоксодержащих генов человека, курицы и испанского тритона, гомологичных ранее клонированному нами гену *Xanf-1* шпорцевой лягушки, охарактеризован новый класс гомеобоксных генов *Anf*. Основная функция этих генов, по-видимому, связана с определением судьбы клеток в головном отделе эмбриона. Гомеодомены белков, кодируемых этими генами, сильно отличаются по первичной структуре от гомеодоменов всех ранее описанных классов гомеобоксных генов. Высокая вариабельность гомеодоменных последовательностей данного класса белков свидетельствует об их быстрой эволюции.

Ключевые слова: гомеобоксодержащие гены, ПЦР, клонирование, передний мозг.

Гомеобоксные гены играют важную роль в определении судьбы клеток у многоклеточных организмов. Продукты этих генов содержат в своем составе консервативную последовательность аминокислот – гомеодомен, с помощью которого они способны связываться со специфическими участками ДНК в регуляторных областях других генов и влиять на эффективность их транскрипции (обзор см. в [1–3]). Показано, что специфичность узнавания и связывания последовательностей ДНК продуктами гомеобоксных генов во многом определяется именно первичной структурой гомеодомена [4]. В настоящее время известно более 30 классов гомеобоксных генов, для каждого из которых характерен свой консенсус последовательности гомеодомена [5].

Цель настоящей работы – характеристика нового класса гомеобоксных генов, сильно отличающихся по консенсусу гомеодоменной последовательности от всех ранее известных классов. кДНК первого представителя данного класса, гена *Xanf-1*, была клонирована нами у шпорцевой лягушки *Xenopus laevis* [6]. Нами же впервые было установлено, что экспрессия этого гена в эмбриогенезе имеет строго ограниченную локализацию – клетки переднеголовой нейроэктодермы и мезоэнтодермы [6, 7].

В дальнейшем у мыши был идентифицирован гомолог *Xanf-1*, названный *Hesx1* [8]. В последст-

вии этот ген был также описан под названием *Rpx* [9]. Недавно у шпорцевой лягушки была клонирована кДНК гена *Xanf-2*, очень близкая по первичной структуре последовательности *Xanf-1* [10]. В настоящей работе мы описываем кДНК трех новых гомологов гена *Xanf-1*: из испанского тритона (*Pleurodeles waltlii*) – ген *Panf*, курицы (*Gallus gallus*) – ген *Ganf* и человека (*Homo sapiens*) – ген *Hanf*.

Особенностью всех этих генов является их специфическая экспрессия на очень ранних стадиях развития в клетках переднеголовой мезоэнтодермы и нейроэктодермы, причем в клетках передней нейроэктодермы (у зародыша шпорцевой лягушки эта зона соответствует переднему нервному валику) наиболее интенсивная. Учитывая эту особенность экспрессии, а также название первого известного представителя данного класса (*Xanf-1*), мы предлагаем называть их генами *Anf* (от английского *anterior neural fold* – передний нервный валик). Полученные нами данные указывают на важную роль генов класса *Anf* в развитии переднеголового отдела зародыша позвоночных, в том числе переднего мозга.

Для клонирования гомеобоксных фрагментов кДНК гомологов гена *Xanf-1* у тритона и курицы мы использовали стратегию, основанную на амплификации эмбриональной суммарной кДНК путем ПЦР с вырожденными праймерами. Образцы суммарной кДНК получали из индивидуальных кусочков переднеголовой нейроэктодермы с помощью разработанного ранее метода супресси-

[#]Автор для переписки (e-mail: zar@humgen.siohc.ras.ru).

онной ПЦР [11] (см. “Экспериментальную часть”). Данную эмбриональную ткань мы выбрали, предположив, что экспрессия *Anf*-генов у представителей других видов позвоночных животных должна иметь такую же тканевую локализацию, что и у шпорцевой лягушки [6, 7].

Очевидно, что подобный подход для получения тотальной кДНК имеет существенное преимущество перед традиционно используемыми методиками, так как позволяет заранее обогатить материал последовательностями искомого гена благодаря прицельному взятию ткани непосредственно из области локализации экспрессии этих генов. Так, в нашем случае объем вырезаемой ткани был примерно в 20 раз меньше объема всего зародыша. Если учесть, что на данной стадии развития практически вся мРНК *Anf*-генов сосредоточена в клетках вырезаемой области, то следует ожидать примерно 20-кратного обогащения матрицами *Anf* получаемых образцов кДНК по сравнению с образцами, выделяемыми из целых зародышей.

Дополнительное преимущество данного подхода заключается в использовании ПЦР для приготовления исходных образцов кДНК, что исключает необходимость выделения больших количеств эмбриональной poly(A)-РНК для синтеза первой цепи кДНК.

Для амплификации фрагментов гомеобокса *Anf*-генов мы использовали набор из четырех выродженных праймеров, соответствующих двум консервативным участкам гомеобоксной последовательности и двум консервативным фрагментам кДНК, фланкирующим гомеобокс соответственно с 5'- и 3'-конца (рис. 1, см. также “Экспериментальную часть”). Консервативность этих участков была предварительно выявлена в результате сравнения последовательностей кДНК двух ранее клонированных *Anf*-генов: *Xanf-1* (шпорцевая лягушка) и *Hesx1/Rpx* (мышь).

Комбинируя эти праймеры, мы получили возможность подобрать для каждого из анализируемых организмов наиболее эффективно работающую пару. Кроме того, данный набор праймеров позволял быстро тестировать специфичность продуктов ПЦР с помощью перехода на один из внутренних праймеров, а в дальнейшем отбирать таким же способом специфические бактериальные клоны, получаемые в результате клонирования этих продуктов.

Для каждого из видов было таким образом идентифицировано и затем секвенировано по несколько позитивных клонов с фрагментами *Anf*-кДНК. Оказалось, что в каждом из случаев (у каждого вида) эти клоны содержали вставки с одинаковыми последовательностями. В то же время последовательности вставок у разных видов различались. Подобный результат указывает

на то, что, по-видимому, на данной стадии развития и в данной области эмбриона у каждого из проанализированных видов позвоночных экспрессируется только один ген *Anf*-класса (см. ниже).

Чтобы клонировать недостающие 5'- и 3'-участки кДНК *Anf*-генов, мы использовали RACE-технику (rapid amplification of cDNA ends), также основанную на методе супрессионной ПЦР [12]. Это дало возможность быстро изолировать искомые фрагменты (см. “Экспериментальную часть”).

В результате для кДНК гена *Panf* (из тритона) были клонированы перекрывающиеся фрагменты, соответствующие последовательности длиной 929 п. о. (номер депонирования в GenBank – U65434). При этом самая длинная открытая рамка считывания (ОРС) этой последовательности, начинающаяся с кодона метионина, соответствовала 185 а. о. (рис. 1).

Для кДНК куриного гена *Ganf* были выделены фрагменты, представляющие последовательность длиной 865 п. о. (номер депонирования – U65436), с ОРС, соответствующей 184 а. о. (рис. 1).

Для выделения человеческого гомолога в качестве матрицы использовали библиотеку кДНК недифференцированной тератокарциномы. В результате были выделены перекрывающиеся фрагменты, соответствующие кДНК гена *Hanf* длиной 754 п. о. (номер депонирования – U65437), кодирующие полипептид длиной 185 а. о., начинающийся с метионина.

Главной особенностью первичной структуры *Anf*-белков, позволяющей объединить их в отдельный класс, является гораздо более высокая взаимная гомология гомеодоменов (более 70%), чем гомология, обнаруживаемая при сравнении любого гомеодомена класса *Anf* с гомеодоменами любых других известных классов (не более 55%). При этом у всех известных *Anf*-гомеодоменов аминокислотные последовательности трех областей, непосредственно вовлеченных во взаимодействие с ДНК и, следовательно, определяющих специфичность гомеодомена (N-концевая часть, область в начале α -спирали 2 и α -спираль 3/4), практически идентичны (рис. 2), но при этом достаточно сильно отличаются от консенсуса последовательностей тех же участков гомеодоменов других классов.

Основную роль при взаимодействии с ДНК играет α -спираль 3/4, так называемая ДНК-узнающая спираль (обзор см. в [4]). В ДНК-гомеодоменном комплексе эта спираль располагается в большой бороздке двойной спирали ДНК, что позволяет аминокислотным остаткам войти в тесный контакт с определенной последовательностью нуклеотидов и образовать с ней многочисленные межмолекулярные связи. При этом первичная структура α -спирали 3/4 определяет

<i>Panf</i>	MFRHALQ--TSAEGPKPSGCS	FTIESILGL	DK-KEAA--AVLK	PHRPW	43
<i>Ganf</i>	ASQNLKRVSGFVENKTTQCS	FSIESILGL	EQKKD---GAAVK	PHRPW	44
<i>Hanf</i>	MSPS-LQEGAQLGENKPSTCS	FSIERILGL	DQKKD--CVPLMK	PHRPW	45
<i>Xanf-1</i>	M-SPALQKGSLLMENRSPSS	FSIEHILGL	DKKTDVASSPIIK	HHRPW	47
<i>Xanf-2</i>	M-STGLQKGSRLMENRSPSS	FSIEHILGL	DKKMDVASSPIIK	HHRPW	47
<i>Hesx1</i> (<i>Rpx</i>)	M-SPSLREGAQLRESKPAPCS	FSIESILGL	DQKKDCTTS--VR	PHRPW	45
<i>Panf</i>	LDPCTNGEGSGSPYIRLFEVSSSESPM-Q-SGGRTEPTQRNAGEE--MQSPAT				91
<i>Ganf</i>	MDGCMH--QLGDPHLQIPVVSYLENSLFH-ANSNLMQEEKVLNCEK-YFSVTV				91
<i>Hanf</i>	ADTSSSGKDGNLCLHVPNPPSGIS-FPSVVDHPMPEERASKYEN-YFSASE				95
<i>Xanf-1</i>	IE-CSSKGVVNGTCWQIPVIACDLPIQVHAVHRSEEEETKIRLEKCFG--DE				97
<i>Xanf-2</i>	IE-CSSK--VDGTFWQIPLISYDLPVQVDAMRRSAEEETKIRLDR--G--EE				93
<i>Hesx1</i> (<i>Rpx</i>)	TDTCGNSEKDGNPPLHAPDLPSETS-FPCPVDHPRPEERAPKYEN-YFSASE				95
<i>Panf</i>	E-RESC	RRDLSWYR	GRRPRTAFSRTQIETLESIFRLNSYPGIDIREELAE		139
<i>Ganf</i>	--RLSF	KRELSWYR	GRRPRTAFTRNQIEVLENVFKMNSYPGIDIREELAR		142
<i>Hanf</i>	--RLSL	KRELSWYR	GRRPRTAFTQNQIEVLENVFRVNCYPGIDIREDLAQ		143
<i>Xanf-1</i>	D-RLTY	KRELSWYR	GRRPRTAFTRSQIEILENVFRVNSYPGIDIVREELAS		145
<i>Xanf-2</i>	D-RLTY	KREQSWYR	GRRPRTAFTRGQIEILENVFRVNSYPGIDIVREELAS		141
<i>Hesx1</i> (<i>Rpx</i>)	T-R-SL	KRELSWYR	GRRPRTAFTQNQVEVLENVFRVNCYPGIDIREDLAQ		143
<i>Panf</i>	KLDLDEDRIQIWFQNRRAKLRSH	RESQFLMVK	NTFSTELQELDHS		185
<i>Ganf</i>	KLDLEEDRIQIWFQNRRAKLRSH	RESQFLMVK	NNFTSSLLE		184
<i>Hanf</i>	KLNLEEDRIQIWFQNRRAKLRSH	RESQFLMAK	KNFNNTLLE		185
<i>Xanf-1</i>	KLALDEDRIQIWFQNRRAKLRSH	RESQFLIVK	DSLSSKIQE		187
<i>Xanf-2</i>	KLALDEDRIQIWFQNRRAKLRSH	RESQFLIVK	DSLSSKIEE		183
<i>Hesx1</i> (<i>Rpx</i>)	KLNLEEDRIQIWFQNRRAKMKRSR	RESQFLMAR	KPFNPDLLK		185

Рис. 1. Сравнение предполагаемых белковых продуктов генов класса *Anf*: *Panf* (*P. waltlii*), *Ganf* (*G. domesticus*), *Hanf* (*H. sapiens*), *Xanf-1* и *Xanf-2* (*X. laevis*) и *Hesx1/Rpx* (*M. musculus*). Гомеодомены заключены в рамку. Другие консервативные области (детали см. в тексте) обведены штриховой линией. Консервативные аминокислотные последовательности, в соответствии с которыми были синтезированы вырожденные праймеры, подчеркнуты пунктирной линией (см. структуру праймеров в "Экспериментальной части").

специфичность связывания гомеодомена. Предполагается [2, 4], что главная роль в узнавании специфической последовательности ДНК принадлежит аминокислотным остаткам, демонстрирующим повышенную эволюционную изменчивость (позиции 1, 2, 5, 6 и 9 спирали 3/4). Экспериментально показано, что наиболее критична в этом отношении позиция 9 узнающей спирали. Так, замена в этой позиции аминокислотного остатка серина, характерного для гомеодомена класса *paired*, на лизин, специфичный для гомеодомена *bicoïd*, приводила к тому, что измененный таким образом продукт гена *paired* полностью потерял способность связываться с последовательностями ДНК, в норме являющимися его мишенями, и, наоборот, приобрел способность связываться с мишенями *bicoïd* [13]. Аналогичный эффект наблюдался и при замене серина-9 на глутамин, характерный для гомеодоменов I–III классов (ге-

ны *HOX*-комплекса, *engrailed*, *msh* и некоторые другие) по классификации Скотта [2]. Гомеодомены всех известных *Anf*-белков также имеют остаток глутамина в 9-й позиции α -спирали 3/4 и, следовательно, теоретически не должны связываться с последовательностями ДНК, специфичными для узнавания продуктами таких гомеобоксных генов, как *PAX*, *Otx* и *goosecoid*. Однако они могут обладать сродством к тем же мотивам ДНК, что и гомеодомены I–III классов.

В отличие от гомеодоменосодержащих белков других классов *Anf*-гомеодомены характеризуются собственным специфическим набором других эволюционно-вариабельных аминокислотных остатков узнающей спирали 3/4. Считается, что хотя эти аминокислоты и имеют меньшее значение для связывания с ДНК, чем аминокислотный остаток в позиции 9 этой спирали, однако они также могут влиять на степень аффинности

	1	10	20	30	40	50	60	
<i>Xanf-1</i>	GRRPRTAFTRSQIEILENVFRVNSYPGIDVREELASKLALDEDRIQIWFQNRRAKLERSH							лягушка
<i>Panf</i>	-----S-T---T--SI--L-----I-----E--D-----							тритон
<i>Xanf-2</i>	-----G-----							лягушка
<i>Ganf</i>	-----N--V-----KM-----I-----R--D-E-----							курица
<i>Hesx1</i>	-----QN-V-V-----C-----I--D--Q--N-E-----M--R							мышь
<i>Hanf</i>	-----QN--V-----C-----I--D--Q--N-E-----							человек
Консенсус	GRRPRTAF...Q...LE..F..N.YPGID.RE.LA.KL.L.EDRIQIWFQNRRAK.KRS.							
		спираль 1		спираль 2		спираль 3/4		
	N-концевая часть		петля					

Рис. 2. Сравнение последовательностей гомеодоменов генов класса *Anf* с последовательностью гомеодомена *Xanf-1* (вверху). Дефисы обозначают аминокислотные остатки, идентичные аминокислотным остаткам последовательности *Xanf-1*. Консенсус последовательностей для гомеодоменов класса *Anf* приведен внизу.

гомеодомена к специфическим последовательностям нуклеотидов [2, 4, 14]. В случае гомеодоменов *Anf*-класса особенно примечателен в этом отношении консервативный остаток аспарагиновой кислоты в позиции 2 узнающей спирали. Насколько нам известно, из более чем 300 описанных гомеодоменов животных такая аминокислота в данной позиции имеется всего у двух: у *scr-10* [15] и *Ost-4* [16]. Примечательно, что эти гомеодомены принадлежат к разным классам; таким образом, остаток аспарагиновой кислоты в позиции 2 представляет собой вариацию, характерную только для двух представителей этих классов. Между тем остаток аспарагиновой кислоты занимает позицию 2 у всех без исключения гомеодоменов класса *Anf*. Важность типа аминокислотного остатка, находящегося именно в этой позиции, для установления эффективного контакта гомеодомена с ДНК была продемонстрирована в экспериментах по комплементации на трансгенных линиях *Drosophila* на примере продукта гомеобоксного гена *ftz* [17].

В экспериментах с продуктами различных гомеобоксных генов было показано, что аминокислоты N-концевой части гомеодомена (аминокислотные остатки в позициях 1–7 гомеодомена) образуют связи с нуклеотидами в малой бороздке ДНК и также влияют на его функциональную специфичность [18, 19]. У всех известных белков *Anf*-класса эти аминокислоты абсолютно консервативны, и их последовательность, GRRPRTA, наиболее близка соответствующим участкам гомеодоменов *engrailed*, *even-skipped* и *paired*, отличающаяся от первых двух по 1-му и 2-му, а от последней – по 1, 2 и 4-му аминокислотным остаткам.

Аминокислоты, расположенные у *Anf*-белков в третьем, контактирующем с ДНК, участке го-

меодомена (между α-спиралями 1 и 2), также образуют полностью консервативную последовательность, PGID (рис. 2), характерную только для этого класса белков. Согласно данным экспериментов с гомеодоменом *ftz*, из числа эволюционно-вариабельных аминокислотных остатков этой области наибольшее влияние на специфичность связывания оказывает остаток в позиции 28 [17]. У продуктов *Anf*-генов это остаток изолейцина. Эта аминокислота достаточно редко встречается в данной позиции у других гомеодоменов и характерна в основном для некоторых гомеодоменов из класса *paired*.

Характерной особенностью всех белков *Anf*-класса является фланкированность их гомеодоменов как с N-, так и с C-конца консервативными специфическими последовательностями длиной в 9 аминокислот: K/RRE/DL/QSWYR и RESQFLM/IV/AK/R соответственно (рис. 1). Эти последовательности не имеют аналогов среди описанных ранее фланкирующих гомеодоменов последовательностей, специфичных для других классов гомеобоксных белков.

Кроме вышеуказанных гомологий все известные белки *Anf* имеют два достаточно консервативных специфических мотива из 6 (P/HHRPW) и 10 а. о. (FS/TIEN/SILGL) около N-конца гомеодомена. Последний мотив имеет отдаленное сходство с известным октапептидом, характерным для PAX-белков: HSIAGILG [28]. В то же время, поскольку известные *Anf*-белки лишены другого важного признака PAX-белков – домена *paired*, а также не имеют какой-либо заметной гомологии с этими белками в области гомеодомена, можно предположить, что данная гомология с октапептидом скорее обусловлена конвергенцией на основе какой-то общей особенности функционирования белков *Anf* и PAX (например, взаимодейст-

Anf					Otx-2			
лягушка (Xanf-1)	лягушка (Xanf-2)	курица	мышь	человек	лягушка	курица	мышь	человек
9	9	8	15	19	1	0	0	0
	1	7	12	9		1	1	1
		8	12	9			0	0
			9	7				0
				3				

Рис. 3. Число аминокислотных различий между гомеодоменами классов Anf и Otx-2 у разных организмов.

вием с родственными белковыми факторами при связывании с регуляторными последовательностями генов-мишеней), нежели их тесным филогенетическим родством.

Вне рассмотренных областей гомологии белки Anf сильно отличаются друг от друга. Исключение составляют продукты двух генов шпорцевой лягушки, *Xanf-1* и *Xanf-2*, имеющие высокую степень гомологии на всем своем протяжении (в среднем около 90%). В то же время одинаковый процент сходства этих генов и их белковых продуктов свидетельствует о том, что дупликация, приведшая к их образованию, произошла в эволюции сравнительно недавно. Поэтому вероятно, что наличие данной пары генов Anf характерно только для шпорцевой лягушки и, возможно, для близких видов амфибий. Дело в том, что у генов, дублировавшихся достаточно давно и вследствие этого присутствующих у разных классов позвоночных (например, гомеобоксные гены *otx-1* и *otx-2* или *en-1* и *en-2*), степень гомологии белковых последовательностей обычно заметно выше, чем нуклеотидных. Можно предположить, что *Xanf-1* и *Xanf-2* являются псевдоаллельными вариантами одного и того же хромосомного локуса, удвоившегося в результате общей дупликации генома, произошедшей в эволюционной линии, ведущей к шпорцевой лягушке [21]. Косвенным подтверждением уникальности дупликации Anf-генов у шпорцевой лягушки может служить также тот факт, что у трех представителей других классов позвоночных были найдены только единичные гомологи *Xanf*, имеющие сходный пространственно-временной характер экспрессии в эмбриогенезе (см. ниже). При этом все они оказались примерно в равной степени дивергировавшими как по отношению друг к другу, так и по отношению к *Xanf-1* и *Xanf-2* [22].

Поскольку гомеодомены Anf-белков у разных видов значительно отличаются друг от друга (рис. 3), может возникнуть предположение, что кодирующие их гены не являются прямыми (ортологическими) гомологами друг друга и в геноме каждого из рассматриваемых видов могут существовать какие-то неизвестные Anf-гены, ортологичные вышеописанным. Оценим вероятность такой ситуации. Для этого предположим, что эти три гена не являются ортологическими гомологами и, значит, у каждого из трех рассматриваемых видов организмов должны существовать по крайней мере еще по два Anf-гена, экспрессирующихся одновременно и в одних и тех же тканях. Для простоты примем, что уровень экспрессии всех этих генов примерно одинаков. В таком случае вероятность обнаружить при клонировании только неортологичные гены составит

$$P = p_1 p_2 p_3,$$

где p_1 – вероятность обнаружить только один тип последовательности у первого вида, p_2 – вероятность обнаружить другой, но также единственный тип у второго вида, p_3 – вероятность обнаружить третий и единственный тип последовательности у третьего вида.

Если, например, у каждого из видов были бы проанализированы только два независимых клон (в действительности мы анализировали по 3–5 клонов), то $p_1 = 3/3^2$; $p_2 = 2/3^2$; $p_3 = 1/3^2$. Тогда $P = 3!/3^6$, что приблизительно равняется 8×10^{-3} . Для трех клонов эта величина составит около 3×10^{-4} . Очевидно, что такая вероятность слишком низка, чтобы рассматривать предположение об абсолютной неортологичности известных Anf-генов как реалистичное.

Можно, однако, допустить, что по крайней мере два из найденных генов являются ортологическими гомологами, а третий им не ортологичен.

В этом случае взаимная гомология ортологических гомеодоменов должна была бы оказаться существенно выше, чем гомология, обнаруживаемая при сравнении каждого из этих двух гомеодоменов с неортологическим. Между тем гомеодомены всех трех генов отличаются друг от друга примерно в одинаковой степени (рис. 3). Таким образом, с большой вероятностью можно заключить, что все три найденных нами гена являются прямыми гомологами.

Если принять предположение об ортологичности известных *Anf*-генов, то обращает на себя внимание необычайно высокая степень различия между их гомеодоменами. Действительно, гомеодомены генов-ортологов всех других известных классов различаются между собой менее чем на 5%, даже если сравниваются гены таких филогенетически удаленных друг от друга таксонов, как рыбы и млекопитающие (исключение составляет класс NOT2, в пределах которого различия между гомеодоменами достигают 10%) [23]. В случае же *Anf*-генов 5%-ная степень различия между гомеодоменами достигается уже в пределах одного класса млекопитающих (человек и мышь) и приближается к 20% при сравнении млекопитающих и амфибий.

Интересно, что все замены в гомеодоменах *Anf* происходят в районах, которые, по-видимому, не находятся в тесном взаимодействии с ДНК (см. выше). В связи с этим можно высказать предположение, что все известные гомеодомены данного класса в меньшей степени вовлечены в белок-белковые кофакторные взаимодействия, чем гомеодомены других классов. В результате их участия, не включенные во взаимодействия с ДНК, получают возможность выйти из-под контроля стабилизирующего отбора и сравнительно быстро эволюционировать.

Отличительной особенностью известных *Anf*-генов является то, что их максимальная экспрессия осуществляется в клетках самой передней части оси тела в период гаструляции и нейруляции (подробнее экспрессия *Anf*-генов описана в [22]). При этом у всех исследованных организмов в первую половину гаструляции данные гены имеют максимум экспрессии в мезодермальном зародышевом листке, в клетках прехордальной пластинки, а начиная с середины гаструляции и далее – в эктодермальном листке, в клетках переднеголовой нейроектодермы.

Поскольку экспрессия всех исследованных *Anf*-генов оказалась локализованной в передней области зародыша в ранний период формирования комплекса осевых органов, причем в производных всех трех зародышевых листков, было бы логично предположить, что аналогично гомеобоксным генам *HOX*-комплекса их основная функция может быть связана с процессом интер-

претации клетками зародыша позиционной информации вдоль главной эмбриональной оси тела [3]. Однако в отличие от *HOX*-генов гены *Anf*-класса, очевидно, вовлечены в спецификацию не туловищного, а переднеголового региона зародыша.

Это предположение подтверждается рядом экспериментальных данных, полученных на шпорцевой лягушке. Так, нами было найдено, что эктопическая экспрессия гена *Xanf-1* в клетках вентральной губы бластопора ранней гаструлы шпорцевой лягушки индуцирует у этих клеток свойства, характерные для клеток переднеголовой области зачатка дорсальной энтомезодермы [7].

На важную роль *Anf*-генов в спецификации судьбы клеток в anteriорной части нейрального зачатка указывают данные по повышенной экспрессии гена *Xanf-1* в клетках нейроэктодермы *X. laevis* (Зарайский, неопубликованные данные). Как было показано, подобное усиление экспрессии вызывает в ряде случаев дубликацию или увеличение некоторых структур переднего мозга.

Итак, *Anf*-гены образуют четко отграниченный по структуре класс гомеобоксных генов, основная функция которых в развитии позвоночных, по-видимому, связана с определением судьбы клеток в переднеголовном отделе зародыша. Очевидно, что, осуществляя эту функцию, гены класса *Anf* должны взаимодействовать с другими регуляторными генами, экспрессирующимися в период гаструляции-нейруляции в переднеголовой области. К числу последних принадлежат, в частности, недавно описанные гомеобоксные гены из семейств *Dlx*, *Emx*, *Nkx-2*, *Otx* и *Pax* [24–27]. Изучение взаимодействия *Anf*-генов с представителями этих и других семейств регуляторных генов поможет приблизиться к пониманию механизмов раннего развития головного отдела тела у позвоночных.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Получение образцов нейроэктодермы. Зародыши испанского тритона (*P. waltlii*) были получены в результате естественного нереста. Стадии развития определяли по методике, описанной в работе [28]. Оплодотворенные яйца кур породы Белый леггорн инкубировали при 38°C, стадии развития определяли по методике [29]. Кусочки ткани вырезали из зародышей в подходящих для каждого вида физиологических растворах с помощью глазного микроскальпеля и оплавленного стеклянного капилляра.

Приготовление образцов амплифицированной κДНК. Тотальную РНК выделяли, используя гомогенизацию ткани в гуанидинизотиоцианате с последующей экстракцией смесью фенол-хлороформ [30]. Синтез первой цепи κДНК и получе-

ние амплифицированной кДНК, включая наращивание oligo-(dA)-хвоста на первую цепь кДНК и последующую ПЦР с Т-праймером (5') CGCCAGTCGACCG(T)₁₃, проводили по [11]. Очистку первой цепи кДНК и амплифицированной кДНК от невключившихся нуклеотидов и избытка праймера осуществляли с помощью Wizard PCR Prep DNA Purification System (Promega).

Амплификация фрагментов гомеобокса. Для амплификации фрагментов ДНК, кодирующих гомеобокс *Anf*, использовали следующие вырожденные праймеры, специфичные к его консервативным участкам (см. рис. 1): (5') AGAGAGARCT-NAGYTGGA для мотива RELSWY; (5') AAYT-CATAUCCNGGTATWGAT для мотива NS/CYPGID; (5') TATYAGRAAYTGKGYTC для мотива ESQFLM/L; (5') CCGNCGRTTYTGRAACCA для мотива WFQNR. ПЦР проводили с использованием предварительно амплифицированной кДНК в качестве матрицы и смеси термостабильных ДНК-полимераз KlenTaq (AB Peptides, USA) и Pfu (Stratagene) в соотношении 150 : 1. Условия: 2 цикла при 94°C (1 мин), 46°C (30 с), 72°C (2 мин), затем 25–28 циклов при 94°C (30 с), 46°C (30 с), 72°C (2 мин).

При поиске человеческого гомолога матрицей для ПЦР служила фаговая ДНК, полученная из библиотеки кДНК недифференцированной тератокарциномы NT2/D1 (Telethon Institute of Genetics and Medicine). В этом случае проводили дополнительный раунд ПЦР с заменой одного из вырожденных праймеров на новый, внутренний, с 1000-кратным разведением продукта ПЦР, полученного в первом раунде. Очищенный продукт ПЦР клонировали в плазмиду pBluescript KS (Stratagene, USA) и 20–30 полученных колоний тестировали ПЦР с внутренним вырожденным праймером, не использовавшимся в исходной ПЦР, и праймером, специфичным для вектора KS. Все положительные клоны секвенировали с использованием fmol sequencing kit (Promega).

Амплификация 3'- и 5'-концевых участков *Anf*-кДНК с помощью метода быстрой амплификации концов кДНК (RACE). Для изоляции недостающих 3'- и 5'-концевых участков кДНК *Anf* мы использовали методику, основанную на технологии супрессионной ПЦР [11, 12]. В отличие от упомянутых методов инвертированные концевые повторы вводили в амплифицированные образцы кДНК не путем лигирования, а с использованием 10 циклов ПЦР с удлинненным Т-праймером: (5') AGCACTCTCCAGCCTCTCACCGCAGTCGACCG(T)₁₃. По окончании этого раунда ПЦР реакционную смесь разводили в 1000 раз и проводили два новых раунда ПЦР по 25 циклов каждый. Первый осуществляли с праймером, специфичным для последовательности гомеобокса, и праймером, представляющим собой часть удли-

ненного Т-праймера: (5') GCACTCTCCAGCCTCTCACCGCA. Во втором раунде использовали заглубленный специфичный праймер и Т-праймер.

Для амплификации 3'-концевых участков кДНК использовали праймеры следующей структуры (для каждого вида первым приведен праймер, использовавшийся в первом раунде ПЦР): (5') AGCTGGTACCGAGGTCCGGA и (5') TCTAGAGGCCGAGGACAGCCTTCAG для тритона; (5') GGGGTAGAAGACCGAGAАСТ и (5') TCTAGACTGCTTTCACCTAGAAACCAG для курицы. Для амплификации 5'-концевых участков кДНК использовали следующие пары специфических праймеров: (5') CTCTCGAATGTCAATGCCG и (5') AGAAGCTTAATGTCAATGCCGGGATAC для тритона; (5') GGGATCTCTCAGTTTTGC и (5') AAGCTTCGGTTCTGGAACCAGAT для курицы.

Фрагменты наибольшей длины, полученные в результате ПЦР, извлекали из агарозного геля и клонировали в плазмиду pBluescript KS для последующего секвенирования.

В случае *Hanf* 5'- и 3'-концевые фрагменты кДНК были получены из библиотеки кДНК NT2/D1 с помощью ПЦР с комбинированием праймеров, специфичных к фаговой ДНК, и пары заглубленных праймеров следующей структуры: (5') CAAGAACTGCTTTTACTCAAAA и (5') GTGTTAGAAAATGTCTTTAG для 3'-концевого фрагмента, (5') CCTTTTCAGTTTTGCACC и (5') CTGGATTCTRTCTTCTCCTCTAG для 5'-концевого фрагмента.

кДНК *Anf* были затем снова амплифицированы с первой цепи кДНК с использованием смеси KlenTaq (AB Peptides, USA) и Pfu (Stratagene) и праймеров, разработанных на основе нуклеотидной последовательности фрагментов, полученных на предыдущем этапе работы: (5') GCTCCGC-CACGCGATC и (5') ACCATTAGAAAATGTTTT-TATTC для тритона; (5') GGGTACCAATCCATCAGCA и (5') GGAAAAGCTTCACTTTCTCCAC для курицы; (5') GCTCTGTGCAGACCACGAGA и (5') TCTGTGTCTAGTACCCTGGT для человека. Затем полученные фрагменты клонировали и секвенировали как описано выше.

Авторы выражают глубокую благодарность О.Л. Васильеву (ИБХ РАН) за ценные методические советы и техническую помощь.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 94-04-11320А), Государственной программы "Геном человека" и Программы INTAS. Работа О.В. Казанской была поддержана грантом ЕМВО для Восточной Европы.

СПИСОК ЛИТЕРАТУРЫ

1. Gehring W. // Science. 1987. V. 236. P. 1245–1252.

2. Scott M.T., Tamkun J.W., Hartzell G.W. // *Biochim. Biophys. Acta.* 1989. V. 989. P. 25–48.
3. McGinnis W., Krumlauf R. // *Cell.* 1992. V. 68. P. 283–302.
4. Gehring W.J., Qian Y.Q., Billeter M., Furukubo-Tokunaga K., Schier A.F., Resendez-Perez D., Affolter M., Otting G., Wuthrich K. // *Cell.* 1994. V. 78. P. 211–223.
5. Kappen C., Schughart K., Ruddle F.H. // *Genomics.* 1993. V. 18. P. 54–70.
6. Zaratisky A.G., Lukyanov S.A., Vasiliev O.L., Smirnov Y.V., Belyavsky A.V., Kazanskaya O.V. // *Dev. Biol.* 1992. V. 152. P. 373–382.
7. Zaratisky A.G., Ecochard V., Kazanskaya O.V., Lukyanov S.A., Fesenko I.V., Duprat A.-M. // *Development.* 1995. V. 121. P. 3839–3847.
8. Thomas P.Q., Johnson B.V., Rathjen J., Rathjen P.D. // *J. Biol. Chem.* 1995. V. 270. P. 3869–3875.
9. Hermesz E., Mackem S., Mahon K.A. // *Development.* 1996. V. 122. P. 41–52.
10. Mathers P., Miller A., Doniach T., Dirksen M.-L., Jamrich M. // *Dev. Biol.* 1995. V. 171. P. 641–654.
11. Lukyanov K.A., Launer G.A., Tarabykin V.S., Zaratisky A.G., Lukyanov S.A. // *Anal. Biochem.* 1995. V. 229. P. 198–202.
12. Chenchik A., Diachenko L., Tarabykin V., Lukyanov S., Siebert D. // *Biotechniques.* 1996. V. 21. P. 526–534.
13. Hanes S.D., Brent R. // *Cell.* 1989. V. 57. P. 1275–1283.
14. Desplan C., Theis J., O'Farrell P.H. // *Cell.* 1988. V. 54. P. 1081–1090.
15. Hawkins N.C., McGhee J.D. // *Nucleic Acids Res.* 1990. V. 18. P. 6101–6106.
16. Scholer H.R., Ruppert S., Suzuki N., Chowdhury K., Gruss P. // *Nature.* 1990. V. 344. P. 435–439.
17. Furukubo-Tokunaga K., Muller M., Affolter M., Pick L., Kloter U., Gehring W.J. // *Genes Dev.* 1992. V. 6. P. 1082–1096.
18. Gibson G., Schier A., LeMotte P., Gehring W.J. // *Cell.* 1990. V. 62. P. 1087–1103.
19. Lin L., McGinnis W. // *Genes Dev.* 1992. V. 6. P. 1071–1081.
20. Burri M., Tromvoukis Y., Bopp D., Frigerio G., Noll M. // *EMBO J.* 1989. V. 8. P. 1183–1190.
21. Richter K., Good P.J., Dawid I.B. // *New Biol.* 1990. V. 2. P. 556–565.
22. Kazanskaya O., Severtzova E., Barth K., Ermakova G., Lukyanov S., Benyumov A., Pannese M., Boncinelli E., Wilson S., Zaratisky A. // *Gene.* 1997. V. 200. P. 25–34.
23. Stein S., Fritsch R., Lemaire L., Kessel M. // *Mech. Dev.* 1996. V. 55. P. 91–108.
24. Papalopulo N., Kintner C. // *Development.* 1993. V. 117. P. 961–975.
25. Gulisano M., Broccoli V., Pardini C., Boncinelli E. // *Eur. J. Neurosci.* 1996. V. 8. P. 1037–1050.
26. Simeone A., Acampora D., Gulisano M., Stornaiuolo A. // *Nature.* 1992. V. 358. P. 687–690.
27. Puchel A., Gruss P., Westerfield M. // *Development.* 1992. V. 114. P. 643–651.
28. Gallien L., Durocher M. // *Bull. biol. France et Belg.* 1957. V. 91. P. 97–114.
29. Hamburger V., Hamilton H. // *J. Morphol.* 1951. V. 88. P. 49–92.
30. Chomczynski P., Sacchi N. // *Anal. Biochem.* 1987. V. 162. P. 156–159.

cDNA Cloning of Three New Homeobox-Containing Genes of the *Anf* Class from Human, Chicken, and Newt

O. V. Kazanskaya*, G. V. Ermakova*, M. Pannese**,
E. Boncinelli**, and A. G. Zaratisky*

*Shemyakin--Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
ul. Miklukho-Maklaya 16/10, GSP-7 Moscow, 117871 Russia

**Istituto Scientifico H. S. Raffaele, Milano, Italy

On the basis of the analysis of cDNA of three new homeobox-containing genes from human, chicken, and newt, a new class of homeobox genes *Anf* is characterized homologous to the *Xanf-1* gene from *Xenopus laevis*, earlier cloned by us. These genes may be largely involved in the specification of embryonic subdivisions in the forebrain region of the embryo. The homeodomains of the proteins encoded by these genes differ greatly in the primary structure from all previously described homeobox genes. The high variability of the homeodomain sequences of the proteins of this class imply their rapid evolution.

Key words: homeobox-containing genes, PCR, cloning, forebrain