



УДК 577.21

ПЕРВИЧНАЯ СТРУКТУРА ОБЛАСТИ ИНИЦИАЦИИ  
ТРАНСКРИПЦИИ ГЕНА ЭСТЕРАЗЫ *S DROSOPHILA VIRILIS*Ениколопов Г. Н., Хечумян Р. К., Кастильо Х.,  
Георгиев Г. И.Институт молекулярной биологии Академии наук СССР,  
Москва

Определена область инициации транскрипции гена органоспецифичной эстеразы *S Drosophila virilis*. Локализация точек инициации транскрипции достигнута при помощи  $S_1$ -анализа. Показано существование множественных точек инициации, две из которых являются мажорными. Их расположение, но не активность использования, не зависит от возраста, пола и линии дрозофил. Определена последовательность нуклеотидов участка вокруг точек инициации. Показано существование в этой области последовательностей, характерных для эукариотических промоторов.

Ген эстеразы *S Drosophila virilis* активируется только в эякуляторной луковице самцов в генетически детерминированный период индивидуального развития. Как показали исследования клонированного гена, интенсивность транскрипции гена быстро нарастает и на 7–8-е сут органоспецифичная эстераза становится одним из мажорных белков эякуляторной луковицы, а соответствующая мРНК составляет 2–3% от всей мРНК этого органа [1–4]. Подобное резкое увеличение количества транскриптов после периода практически полного молчания гена привело к предположению о наличии в нем тканеспецифичного активирующего элемента. Настоящее сообщение посвящено определению точки инициации транскрипции клонированного гена эстеразы *S D. virilis* [5] и анализу последовательности нуклеотидов вокруг нее.

Точку инициации транскрипции гена эстеразы определяли путем гибридизации различных препаратов РНК с фрагментами клонированного гена и обработки гибридов нуклеазой  $S_1$ , специфически расщепляющей однонитевые участки нуклеиновых кислот ( $S_1$ -анализ) [6, 7]. Предварительный поиск 5'-концевой границы транскрибируемой зоны выявил ее в середине фрагмента *EcoRI-F* [5], и мы использовали различные зонды, перекрывающие эту область, для точного определения участка начала считывания гена. Физическая карта гена, карта фрагмента *EcoRI-F*, использованные зонды и результаты анализа с применением различных зондов приведены на рис. 1 и 2. Эксперименты по  $S_1$ -анализу с применением всех указанных зондов дают идентичные результаты и указывают на область, находящуюся на расстоянии около 650 п.о. от левой границы фрагмента *EcoRI-F*, как на область инициации транскрипции гена эстеразы *S*. Альтернативным объяснением могло бы быть расположение в этой зоне 3'-донорного участка сплайсинга РНК (тем самым предполагается, что весь клонированный участок перед этой зоной является интроном, а предыдущий экзон (экзоны) лежит за пределами клонированного гена). Против этого предположения говорят результаты иных способов определения точки инициации транскрипции — элонгации праймера при помощи обратной транскриптазы, транскрипции *in vitro*\*, а также первичная структура этой области, не содержащей консенсусных последовательностей границ экзон/интрон [8], и, напротив, содержащей участки, характерные для областей инициации транскрипции эукариотических генов (см. ниже).

Для более точной локализации точки (точек) инициации транскрип-

\* Статья Г. Н. Ениколопова и др. подготовлена к печати.

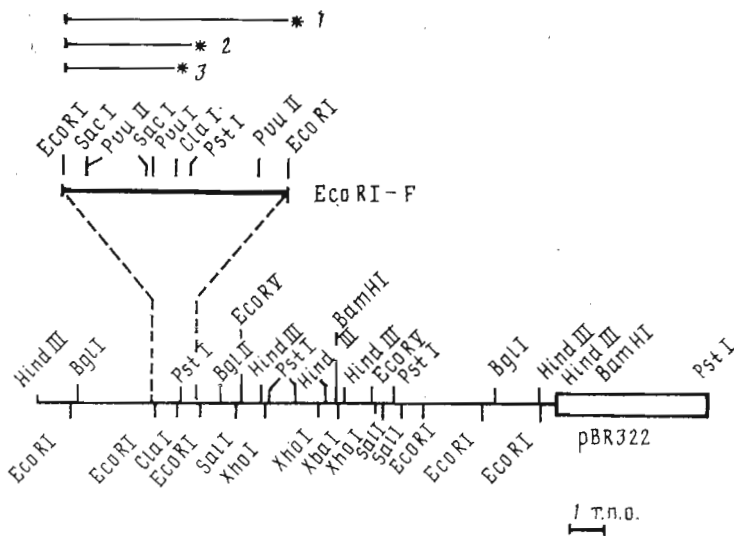


Рис. 1. Физическая карта плазмиды pVE9, содержащей ген эстеразы *S. D. virilis*. В увеличенном виде показана физическая карта субфрагмента *EcoRI-F* и радиоактивные зонды (1)–(3), использованные для  $S_1$ -анализа

дции мы воспользовались меченым по 5'-концу зондом (3) (рис. 1) длиной около 600 п.о., полученным с помощью рестриктазы *ClaI*. После гибридизации зонда с РНК зрелых 10-дневных самцов гибриды обрабатывали нуклеазой  $S_1$  при различных ее концентрациях и продукты разделяли высокоразрешающим электрофорезом в ПААГ. Применение различных концентраций нуклеазы  $S_1$ , перекрывающих 40-кратный интервал, увеличивает достоверность полученных результатов и сводит к минимальной возможности того, что некоторые полосы образованы за счет недостаточной активности нуклеазы, не до конца гидролизующей концевой участок гибрида с какой-либо резистентной к ее действию конформацией, или же, наоборот, за счет избытка нуклеазы, приводящего к разрывам меченой цепи зонда внутри гибрида. Результат эксперимента, приведенный на рис. 3а, показывает наличие двух мажорных точек инициации транскрипции, разделенных пятью парами нуклеотидов и расположенными на расстоянии 187 и 192 п.о. от участка расщепления рестриктазой *ClaI*; наблюдаемое видимое удвоение полос часто отмечается в опытах с применением нуклеазы  $S_1$  [6, 7] и, по-видимому, объясняется интерферирующим действием кэп-группировки на 5'-конце зрелых молекул мРНК. Помимо двух мажорных полос анализ выявляет дополнительные минорные полосы, расположенные как ниже, так и выше основных полос и отражающие существование альтернативных точек инициации транскрипции\*.

Представлялось важным узнать, насколько использование точек старта транскрипции постоянно на различных стадиях развития дрозофилы и в различных линиях. Результаты  $S_1$ -анализа набора препаратов РНК, приведенные на рис. 3б, демонстрируют, что расположение мажорных точек инициации транскрипции, во-первых, не меняется в процессе онтогенеза (результаты гибридизации РНК, выделенной из 5-, 12- и 16-дневных мух), во-вторых, не зависит от пола мух (хотя транскрипция гена намного сильнее у самцов *D. virilis*), и, в-третьих, не различается у линий *D. virilis*, различающихся по времени начала синтеза экстеразы (линии СЕ и 160) [4].

Для характеристики области инициации транскрипции мы определили первичную структуру участка, покрывающего область перед и после (по ходу транскрипции) мажорных точек инициации транскрипции. Для

\* Г. Н. Ениколопов и др. // Генетика. 1988. В печати.

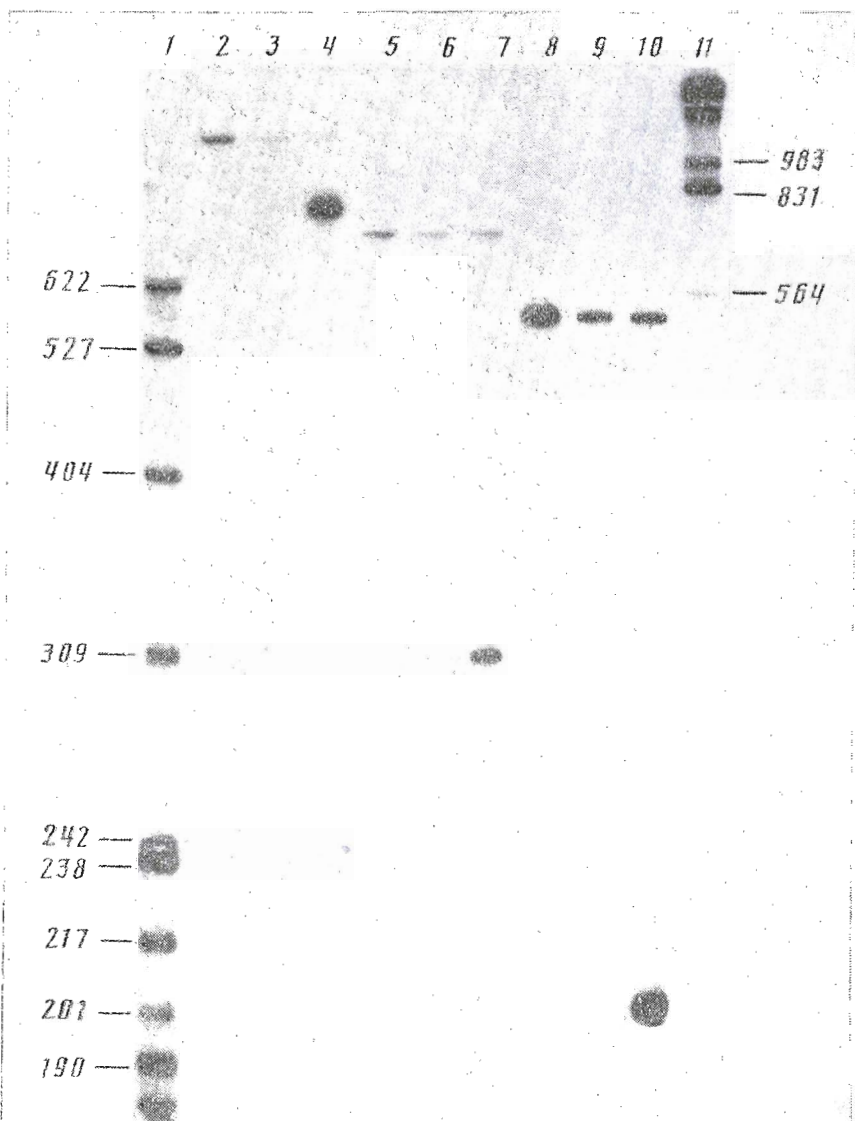


Рис. 2. Картирование области инициации транскрипции гена эстеразы S при помощи нуклеазы S<sub>1</sub> с применением различных зондов. Дорожки 2, 5, 8 – препараты меченых зондов (1), (2), (3) (см. рис. 1); 3, 6, 9 – контрольная гибридизация зондов (1), (2) и (3) с РНК мышц с последующей обработкой нуклеазой S<sub>1</sub>; 4, 7, 10 – то же, но при гибридизации с тотальной РНК зрелых 10-дневных самцов. Дорожки I и II – меченые маркерные фрагменты А и В соответственно; цифрами по краям геля указаны их размеры (п.о.)

секвенирования был выбран субклон, содержащий плазмиду pRF1, полученную при помощи гидродинамического фрагментирования ДНК, разделения смеси фрагментов в агарозном геле, клонирования в плазмидном векторе pUC19 и отборе гибридизацией. Последовательность нуклеотидов определяли методом Максама – Гилберта в стандартном варианте [9] и в его твердофазной модификации [10] и частично методом Сэнгера [11] так, что каждая пара нуклеотидов была прочитана несколько раз по обеим цепям с различных точек. Последовательность нуклеотидов приведена на рис. 4.

Характерные особенности последовательности сводятся к следующему:

1) за 28–32 нуклеотида от места старта транскрипции располагается последовательность, сходная с ТАТА-боксами, характерными для подавляющего большинства эукариотических генов, транскрибируемых РНК-

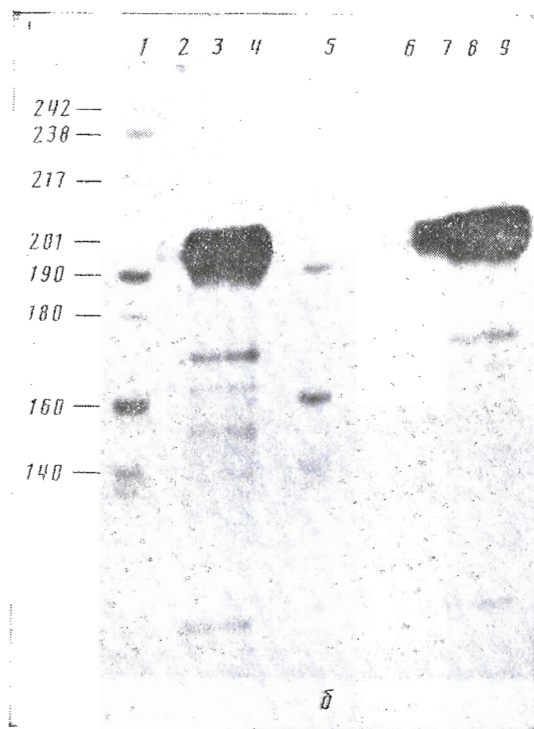
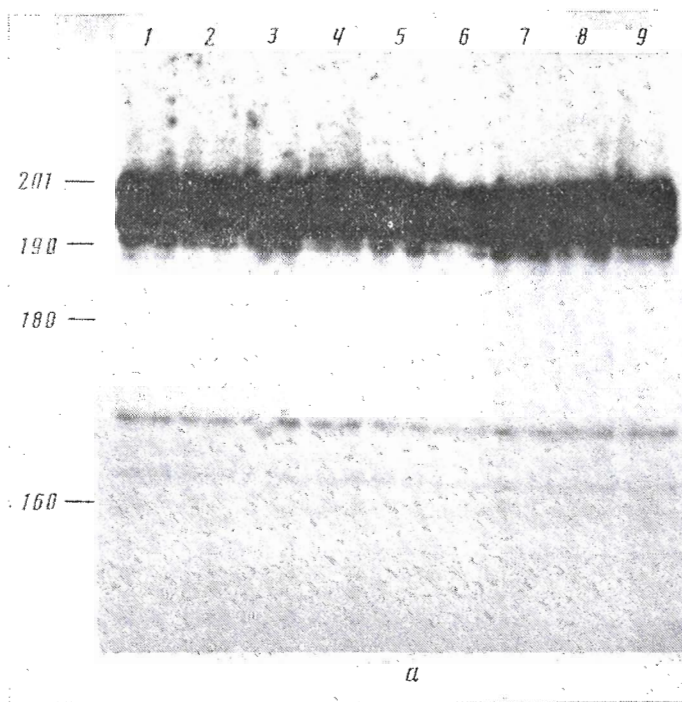


Рис. 3. Анализ транскрипции гена эстеразы S при помощи пуклеазы S<sub>1</sub> (гибридизация с меченым зондом (3)). а - гибриды обрабатывали при 18°С нуклеазой S<sub>1</sub> в количестве 10 (1), 20 (2), 30 (3), 40 (4), 70 (5), 100 (6), 150 (7), 250 (8), 400 ед. акт. (9). б - использование различных препаратов РНК *D. virilis*: РНК тела зрелых самцов после удаления семяносящих луковец (2), РНК 12-дневных самцов линии СЕ (3), РНК 12-дневных самцов линии 160 (4, 8), 5-дневных (7), 16-дневных (9), РНК 12-дневных самок (6). Дорожки 1, 5 - меченые маркерные фрагменты А; цифрами по краям геля указаны их размеры (п.о.)

CGTTCACITTT TACGAAAGTC AGCGCGTGAT TCATAATTTT ACAGTTCCTA 50  
 AGAGCGGAAC AACTTAGAAA TGGGGAATCA TCTATTAAG ACCCCATAAA 100  
 STAAGCGGTG AACCCAACTT CCAAACCTCTG TGGCCTCTAT AAAACAAATG 150  
 GCAGAGTCTT AAATTTTCAA ACGCAGITTA ACTGCCAATG CCAAGGGGCC 200  
 AGCCATCGGA ACTAATCGAA TAGCTGTCTT CTTTTCTCAC GTGAACAAA 250  
 CTTAGACTCA AGTGACCATA AAAAGGGAGT TGGCTGTCTC CGTTTCGAAC 500  
 ATTC AATTC AAGTT TGAGCTCCGT CCTGTGCACA CCATGACTCA AAIAC TGTTC 550  
 MetThrGln IleLeuLeu  
 CCGATGGCGT TGCTCTGTCT GTTTCAGCA TCAACCTCA GCAATGCC 400  
 ProIleAlaLeu LeuCysLeu PheAlaAla SerThrLeuSer AsnProLeu  
 GCTGTGGAG TTGCCAAATG CAGAATTGCC GGGACGGAC AATGGATTCT 450  
 LeuValGlu LeuProAsnGly GluLeuArg GlyArgAsp AsnGlyPhe  
 AITACAGCTA CGAGTCAATA CCCTATGCCG AACCCCAAT CGATGATCTG 500  
 TyrTyrSerTyr GluSerIle ProTyrAlaGlu ProProIle AspAspLeu  
 TCGTTGGAAC AACCTCGTCC STATACCGAA  
 CysLeuGluGlu ProArgPro TyrThrGln

Рис. 4. Первичная структура области инициации транскрипции гена эстеразы *S. D. virilis*. Приведена часть аминокислотной последовательности. Мажорный участок инициации транскрипции взят в рамку. Подчеркнуты участки гомологии с транскрипционно-активными элементами (направление стрелки соответствует цепи гена) — TGTCG-боксом (сплошная линия), длинными концевыми повторами ретровирусов (волнистая линия), длинными концевыми повторами мобильных элементов дрозофилы (штриховая линия)

полимеразой II (например, 12 нуклеотидов указанной последовательности полностью гомологичны TATA-боксам генов  $\beta$ -глобина кролика, андрогенстимулируемого белка предстательной железы крысы и др.) [12];

2) ген не несет в положении  $-70$  —  $-80$  последовательности, сходной с CAAT-боксом [12], характерным для многих, хотя и не всех, эукариотических генов;

3) строение участков инициации транскрипции, выявляемое по результатам  $S_1$ -анализа, совпадает с усредненными (весьма вырожденными) последовательностями участков инициации синтеза и копирования мРНК [12, 13];

4) в секвенированном фрагменте не встречаются усредненные последовательности границ соединений экзон/интрон и интрон/экзон [8];

5) ген несет относительно короткую (около 35 нуклеотидов) 5'-нетранслируемую часть мРНК. Область, непосредственно примыкающая к первому иницирующему кодону AUG (CACCAUG), соответствует усредненной последовательности области инициации трансляции генов животных (CANCAUG) [14] и уточненной последовательности для генов дрозофилы ( $^c/_A$  AA  $^A/_C$  AUG) [15];

6) при анализе последовательности секвенированного участка были выявлены значимые гомологии последовательностей нуклеотидов обеих цепей участка с энхансерами онкогенных вирусов, длинными концевыми повторами ретровирусов, длинными концевыми повторами мобильных элементов дрозофилы, регуляторными участками структурных генов эукариот.

Как известно, сильнейшие транскрипционно-активные элементы эукариотических и вирусных геномов, являющиеся энхансерами или обладающие энхансернодобными свойствами, не содержат протяженных участков гомологии между собой, а скорее представляют набор отдельных относительно коротких блоков, мозаично соединенных друг с другом, и, часто за счет синергизма действия, определяющих резкое повышение эффективности энхансера как транскрипционного усилителя [16—18].



Результаты поисков гомологий между энхансерами чаще всего сводятся к выявлению отдельных коротких блоков, характерных для большинства энхансеров, но расположенных друг относительно друга без видимой закономерности. На рис. 4 отмечены некоторые из обнаруженных участков, сходных с блоками энхансеров: так называемый TGTGG-блок, характерный практически для всех известных энхансеров; область, сходная с активизирующими элементами длинных концевых повторов онкогенных ретровирусов; области, сходные с последовательностями из длинных концевых повторов мобильных генетических элементов дрозофилы *mdr1* и *mdr3*. Обращает на себя внимание сходство с энхансерами на обеих цепях гена эстеразы *S*, что согласуется с представлениями о двунаправленном действии энхансеров. Подробный компьютерный анализ выявил множество подобных участков, развернутые данные исследования будут представлены позже.

Итак, в настоящей работе определена область инициации транскрипции гена органоспецифической эстеразы *S D. virilis*, установлена первичная структура этой области и показано наличие в ней нуклеотидных последовательностей, гомологичных энхансерам и энхансероподобным элементам. Дальнейшее изучение этого участка гена связано с доказательством его роли в активации транскрипции гена эстеразы в функциональных тестах *in vivo* и *in vitro* и поиске факторов, опосредующих его действие в ядре клетки.

### Экспериментальная часть

Использовали рестрикционные эндонуклеазы НПО «Фермент» (Вильнюс), ДНК-полимеразу *E. coli*, фрагмент Кленова ДНК-полимеразы *E. coli*, Т4-полинуклеотидкиназу (P-L Pharmacia, США), материалы для электрофореза в полиакриламидном геле (Serva, ФРГ). ДНК исходной плазмиды pVE9 [5], плазмиды pRF1 и плазмиды из других субклонов выделяли щелочным методом [19].

Первичную структуру фрагментов ДНК определяли методом Максама — Гилберта (после введения метки по 5'- или 3'-концам) в обычном [9] или твердофазном [10] вариантах. Часть последовательностей была определена также методом Сэнгера [11].

РНК выделяли модифицированным методом Чиргвина и др. [20], лизируя порошок трани имаго, растертых в жидком азоте, в растворе роданистого гуанидина, депротенинизируя лизат смесью фенол — хлороформ и удаляя ДНК многократными пересаживаниями из высокосолевого раствора. РНК хранили при  $-20^{\circ}\text{C}$  в виде спиртовой суспензии.

Для анализа транскриптов при помощи нуклеазы  $S_1$  [7] радиоактивные зонды (1) — (3), меченные по 5'-концу с помощью [ $\gamma$ - $^{32}\text{P}$ ] АТФ, выделяли элюцией с DEAE-бумаги после разделения фрагментов ДНК в агарозном геле и электрофоретического переноса их на DEAE-бумагу. Фрагменты, полученные под действием рестриктазы *Pst*I, метили по 5'-концу после денатурации и меченые цепи зонда разделяли в ПААГ. Меченые фрагменты ДНК смешивали с препаратом РНК (10 мкг), осаждали спиртом и после растворения в буфере, содержащем 0,4 М NaCl, 0,1 М 1,4-пиперазин-бис(этансульфоновою) кислоту (рН 6,4), 5 мМ EDTA, 80% формамид, и денатурации отжигали в течение 8—14 ч при  $48^{\circ}\text{C}$ . Гибриды РНК-ДНК обрабатывали нуклеазой  $S_1$  в растворе, содержащем 0,2 М NaCl, 0,1 М  $\text{CH}_3\text{COONa}$  (рН 4,5), 1 мМ  $\text{ZnSO}_4$ , при  $18^{\circ}\text{C}$  в течение 1 ч. Продукты гидролиза пересаждали и разделяли в высокоразрешающем 5% ПААГ.

Маркеры А — меченые фрагменты ДНК плазмиды pBR 322, обработанной рестриктазой *Pvu*II.

Маркеры В — меченые фрагменты ДНК фага  $\lambda$ , обработанной рестриктазами *Eco*RI и *Hind*III.

### ЛИТЕРАТУРА

1. Korochkin L. I. // Isozymes. V. 4/Eds Rattazzi M. C., Scandalios J. G., Whitt G. S. N. Y. Alan R. Liss. Inc. P. 159—202.
2. Korochkin L. I. // Gene regulation in development. Berlin-Heidelberg-New York: Springer-Verlag, 1981.
3. Корочкин Л. И. Биология развития и управление наследственностью. М.: Наука, 1983. С. 267—284.
4. Korochkin L. I., Matveeva N. M., Kuzin B. A., Karasik G. I., Maximovsky L. F. // Biochem. Genet. 1978. V. 16. № 5. P. 709—726.
5. Yevnikolov G. N., Kuzin B. A., Evgen'ev M. B., Ludwig M. Z., Korochkin L. I., Georgiev G. P. // EMBO J. 1983. V. 2. № 1. P. 1—7.
6. Favalaro J., Treisman R., Kamen R. // Meth. Enzymol. 1980. V. 68. P. 718—749.
7. Weaver R. F., Weissmann C. // Nucl. Acids Res. 1979. V. 7. № 5. P. 1175—1193.
8. Mount S. M. // Nucl. Acids Res. 1982. V. 10. № 2. P. 459—472.

9. Maxam A. M., Gilbert W. // Proc. Nat. Acad. Sci. USA. 1977. V. 74. № 2. P. 560—564.
10. Chuvpilo S. A., Kravchenko V. V. // FEBS Lett. 1985. V. 179. № 1. P. 34—36.
11. Sanger F., Nicklen S., Coulson A. R. // Proc. Nat. Acad. Sci. USA. 1977. V. 74. № 12. P. 5463—5467.
12. Breathnach R., Chambon P. // Ann. Rev. Biochem. 1981. V. 50. P. 349—383.
13. Hentschel C., Irminger J.-C., Bucher P., Birnstiel M. L. // Nature. 1980. V. 285. № 5761. P. 147—151.
14. Kozak M. // Nucl. Acids Res. 1984. V. 12. № 2. P. 857—872.
15. Cavener D. R. // Nucl. Acids Res. 1987. V. 15. № 4. P. 1353—1361.
16. Boshart M., Weber F., Jahn G., Dorsch-Häslter K., Fleckenstein B., Schaffner W. // Cell. 1985. V. 41. № 2. P. 521—530.
17. Zenke M., Grundström T., Matthes H., Wintzerith M., Schatz C., Wildeman A., Chambon P. // EMBO J. 1986. V. 5. № 2. P. 387—397.
18. Schirm S., Jiricny J., Schaffner W. // Genes and Developm. 1987. V. 1. № 1. P. 65—74.
19. Marco M. A., Chipperfield R., Bornboim H. C. // Analyt. Biochem. 1982. V. 121. № 2. P. 382—387.
20. Chirgwin J. M., Przybyla A. E., MacDonald R. J., Rutter W. J. // Biochemistry. 1979. V. 18. № 24. P. 5294—5299.

Поступила в редакцию  
13.V.1988  
После доработки  
23.VI.1988

### THE SEQUENCE OF THE TRANSCRIPTION INITIATION REGION OF THE ESTERASE S GENE OF *DROSOPHILA VIRILIS*

YENIKOLOPOV G. N., KHECHUMIAN R. K., CASTILLO J., GEORGIEV G. P.

*Institute of Molecular Biology, Academy of Sciences of the USSR, Moscow*

The transcription initiation region of the gene for tissue-specific esterase *S* of *D. virilis* has been analysed. By means of high-resolution  $S_1$ -mapping we located a set of transcription initiation points, two of them being major ones. The position of the major start points, but not the efficiency of their usage, does not depend on age, sex or strain of the flies. The transcription initiation region is sequenced. The region contains various motifs characteristic for eukaryotic promoters.