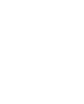
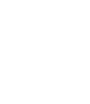
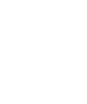
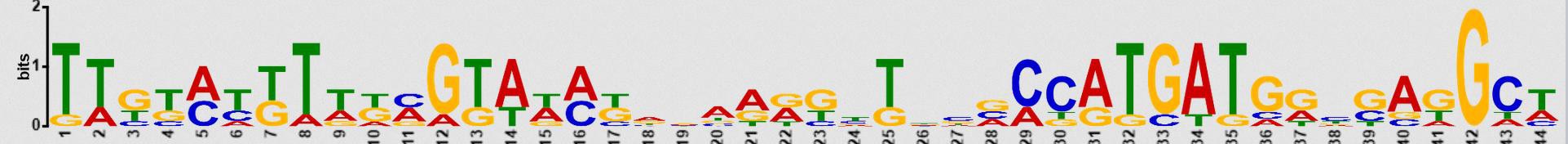
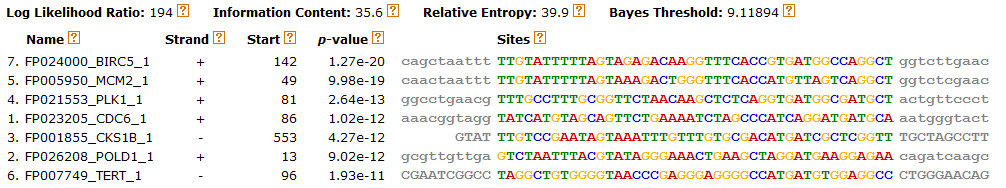


**Рис. S1.** Анализ обогащением (enrichment analysis) профилей узнавания транскрипционных факторов в семи опухолеспецифичных промоторах

относительно промоторов генов домашнего хозяйства с помощью программы CiiiDER (*Gearing L.J., Cumming H.E., Chapman R., Finkel A.M., Woodhouse I.B., Luu K., Gould J.A., Forster S.C., Hertzog P.J.* // PLoS One. 2019. V. 14. P. e0215495. https://doi.org/10.1371/journal.pone.0215495). Подписаны наиболее значимые факторы.



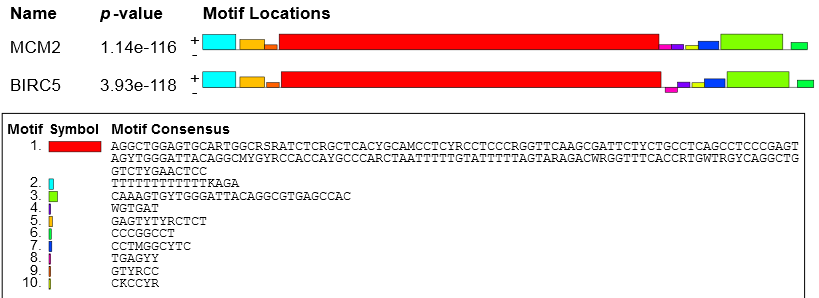
1a

1b

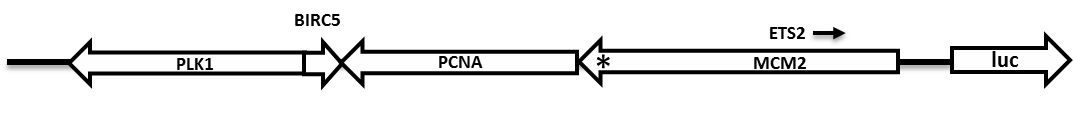
1c

**Рис. S2.** Мотив 1, отличающий опухолеспецифические промоторы от промоторов генов домашнего хозяйства, и субмотивы в его составе.

“Старт” указан от –499 нуклеотида промотора относительно точки начала транскрипции.



**Рис. S3.** Общий консенсусный мотив промоторов *MCM2* и *BIRC5*. Мотив 2 длиной 184 п.н. показан красным цветом.



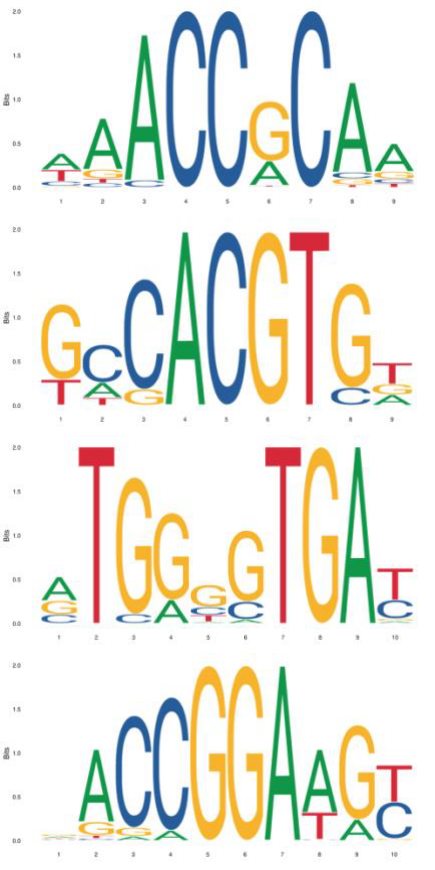
(*a*) Химерный промотор CH2

(*б*) Химерный промотор CH26

(*в*) Химерный промотор CH20

**Рис. S4.** Примеры химерных промоторов (CH2, CH26, CH20), состоящих из фрагментов нескольких промоторов в разной ориентации.

Отмечены точки начала транскрипции исходных промоторов (\*) и положение профилей узнавания некоторых факторов транскрипции.



**Таблица S2.** Профили узнавания факторов транскрипции, более частые в опухолеспецифических промоторах,

чем в промоторах генов домашнего хозяйства

Фактор транскрипции

Номенклатура

JASPAR

Профиль узнавания

JASPAR

Опухоли с дифференциальной экспрессией (GEPIA2, TCGA)\*

Прогностическое значение (THPA)\*\*

RUNX2

MA0511.2

ESCA, LAML, PAAD, STAD, THYM (+); TGCT (–)

Рак почки (–), уротелия (–), желудка (–)

Creb3l2

MA0608.1

DLBC, GBM, LGG, PAAD, SKCM, THYM (+);

CESC, OV, UCEC, UCS (–)

НП

SREBF2

MA0596.1

DLBC, THYM (+); GBM, KIRC (–)

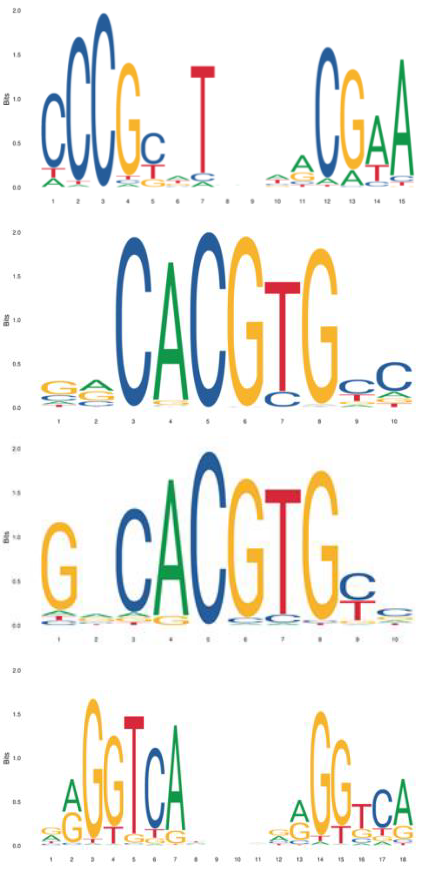
Рак почки (+)

ETS2

MA1484.1

COAD, PAAD, READ, STAD (+); BLCA, BRCA, DLBC, GBM, KICH, LUAD, LUSC, OV, PRAD, SKCM, THCA, UCEC, USC (–)

НП



CENPB

MA0637.1

DBLC, GBM, LGG, PAAD, THYM (+)

Рак печени (–)

HEY1

MA0823.1

GBM, KICH, KIRC, LGG, PAAD, SKCM, THYM, UCS (+);

BRCA, LAML, LUAD (–)

НП

HEY2

MA0649.1

GBM, KIRC, LGG, THCA, THYM (+); BRCA, CESC, TGCT (–)

НП

RARA::RXRG

MA1149.1

RARA—ESCA, GBM, LAML, LGG, PAAD, STAD (+);

CESC, DLBC, LUSC, TGCT, THYM (–);

RARG—SKCM, THCA (+);

BRCA, CESC, COAD, GBM, LUAD, LUSC, READ (–)

RARA – рак эндометрия (+); RXRG – рак

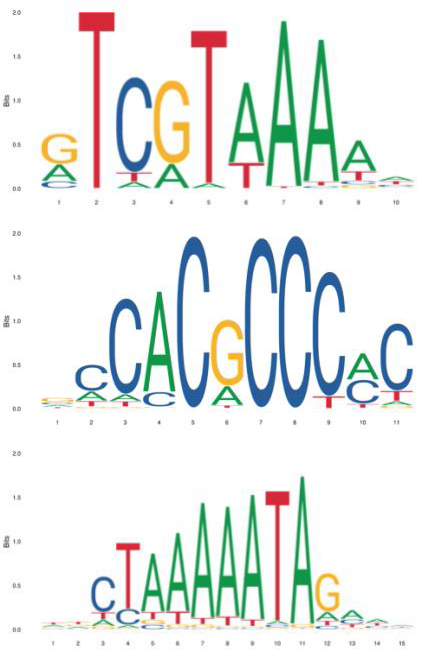
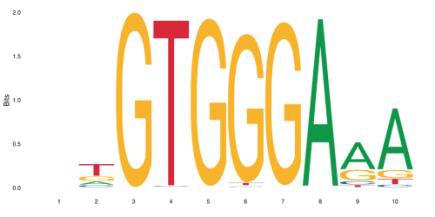
щитовидной железы (+)

Srebf1 (var.2)

MA0829.1

DLBC, KICH, PAAD, THYM (+); ACC, OV, TGCT, UCS (–)

Рак эндометрия (+), поджелудочной железы (+)



ZNF75D

MA1601.1

LAML, THYM (+); LUSC, OV, USC (–)

Рак уротелия (+),

почки (+)

Zfx

MA0146.2

GBM, LAML, LGG, THYM (+); UCEC, UCS (–)

НП

HOXD11

MA0908.1

ESCA, GBM, HNSC, LUSC (+); COAD, KICH, KIRP, PRAD, UCEC, UCS (–)

НП

KLF11

MA1512.1

BLCA, BRCA, SKCM, TGCT, UCEC, UCS (–)

НП

MEF2A

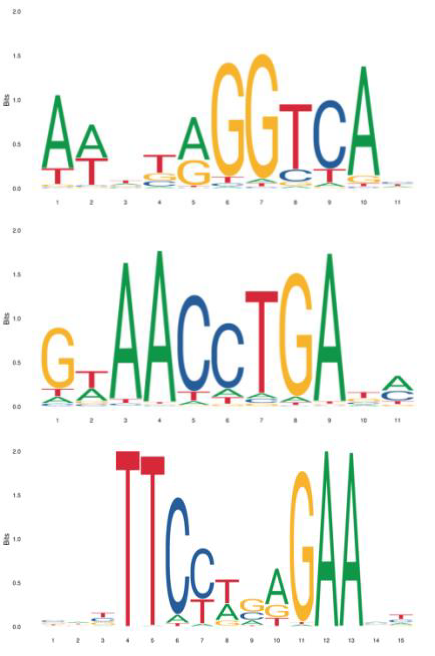
MA0052.4

PAAD (+);

BLCA, CESC, COAD, LUAD, LUSC, READ, SKCM, UCEC, UCS (–)

Рак почки (+), желудка

(–)



\* Номенклатура опухолей согласно базе данных The Cancer Genome Atlas: (+) – повышенная экспрессия, (–) – пониженная экспрессия.

\*\* Прогностическое значение согласно базе данных The Human Protein Atlas: (+) – благоприятный прогноз, (–) – неблагоприятный прогноз,

НП – нет связи с прогнозом.

RORB

MA1150.1

GBM, LGG (–)

НП

SIX1

MA1118.1

BLCA, BRCA, CESC, ESCA, GBM, KICH, LGG, LUAD, LUSC, OV, THYM, UCEC, UCS (+);

TGCT, THCA (-)

Рак эндометрия (–),

легких (+)

Stat6

MA0520.1

LAML, PAAD (+);

CESC, LGG, LUAD, LUSC, OV, PRAD, UCEC, UCS (–)

Рак уротелия (+), эндометрия (+), щитовидной железы (+)

**Таблица S3.** Встречаемость профилей узнавания факторов транскрипции в опухолеспецифических промоторах

Промотор

Профили факторов транскрипции

*MCM2*

Creb3l2, SREBF2, ETS2, CENPB, HEY1, HEY2, RARA::RXRG, SREBF1(var2), ZNF75D, Zfx, MEF2A, RORB, Stat6

*CKS1B*

RUNX2, SREBF2, ETS2, CENPB, RARA::RXRG, ZNF75D, Zfx, HOXD1, SIX1, Stat6

*PLK1*

RUNX2, Creb3l2, SREBF2, ETS2, HEY1, HEY2, SREBF1(var.2), ZNF75D, SIX1

*POLD1*

RUNX2, SREBF2, RARA::RXRG, ZNF75D, Zfx, HOXD11, KLF11, RORB

*TERT*

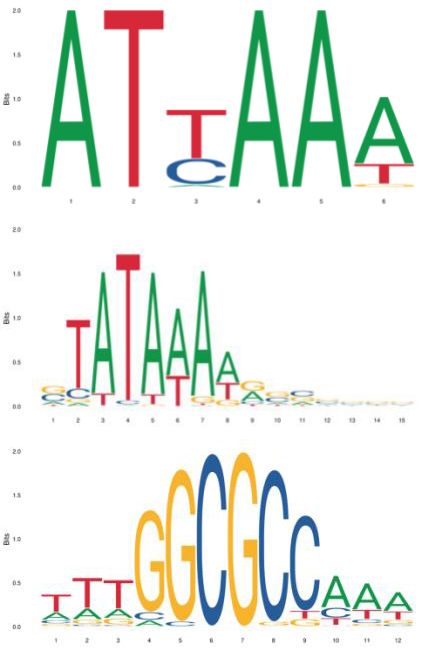
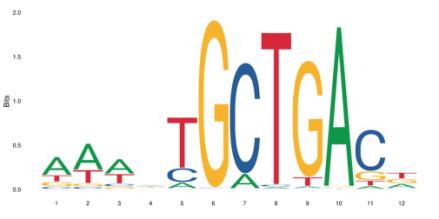
Creb3l2, SREBF2, ETS2, HEY1, HEY2, ZNF75D, Zfx

*CDC6*

RUNX2, Creb3l2, ETS2, CENPB, ZNF75D, Zfx

*BIRC5*

RUNX2, SREBF2, SREBF1(var.2), Zfx, KLF11, MEF2A



**Таблица S4.** Профили узнавания факторов транскрипции, более частых в промоторах генов домашнего хозяйства, чем в

опухолеспецифических промоторах

Фактор транскрипции

Номенклатура

JASPAR

Профиль узнавания

JASPAR

Опухоли с дифференциальной экспрессией (GEPIA2, TCGA)\*

Прогностическое значение

(THPA)\*\*

Mafb

MA0117.2

GBM, LGG, OV, PAAD, STAD, TGCT, UCS (+);

SKCM (–)

Рак эндометрия (–)

Arid3a

MA0151.1

COAD, READ, STAD (+); TGCT, THYM (–)

Рак поджелудочной железы

(+), почки (–)

TBP

MA0108.2

DLBC, THYM (+); LAML, TGCT (–)

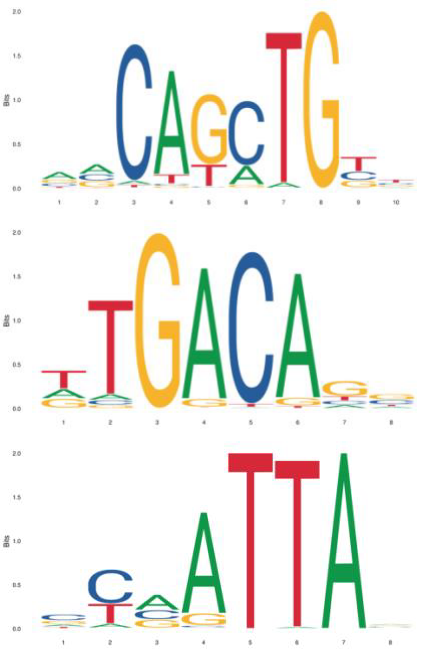
Рак печени (–)

E2F1

MA0024.3

ACC, BLCA, BRCA, CESC, COAD, DLBC, ESCA, GBM, HNSC, KICH, KIRC, KIRP, LIHC, LUAD, LUSC, OV, PAAD, READ, SKCM, STAD, THCA, THYM, UCEC, UCS (+)

Рак шейки матки (+), печени (–), эндометрия (–), почки (–), щитовидной железы (+), поджелудочной железы (–)



\* Номенклатура опухолей согласно базе данных The Cancer Genome Atlas: (+) – повышенная экспрессия; (–) – пониженная экспрессия.

\*\* Прогностическое значение согласно базе данных The Human Protein Atlas: (+) – благоприятный прогноз; (–) – неблагоприятный прогноз,

НП – нет связи с прогнозом.

BHLHA15(var.2)

MA1472.1

OV, PRAD (+); PAAD, STAD (–)

НП

MEIS3

MA0775.1

DLBC, HNSC, KIRP, PAAD (+); CESC, COAD, OV, READ, THCA, UCEC, UCS (–)

Рак почки (–)

BSX

MA0876.1

—

НП

**Таблица S5.** Свойства факторов транскрипции, более частых в промоторах генов домашнего хозяйства, чем в

опухолеспецифических промоторах

Фактор транскрипции

Номенклатура

JASPAR

Класс

Семейство

Опухолеспецифические промоторы, содержащие профиль

Число промоторов генов домашнего хозяйства, содержащих профиль

Mafb

MA0117.2

Basic leucine zipper factors (bZIP)

Maf-related factors

*MCM2*, *POLD1*, *BIRC5*

22

Arid3a

MA0151.1

ARID domain factors

ARID-related factors

*POLD1*, *BIRC5*

20

TBP

MA0108.2

TATA-binding proteins

TBP-related factors

–

14

E2F1

MA0024.3

Fork head / winged helix factors

E2F-related factors

–

12

BHLHA15(var.2)

MA1472.1

Basic helix-loop-helix factors (bHLH)

Tal-related factors

*POLD1*, *TERT*

18

MEIS3

MA0775.1

Homeo domain factors

TALE-type homeo domain factors

*MCM2*, *POLD1*, *PLK1*,

*CDC6*

22

BSX

MA0876.1

Homeo domain factors

NK-related factors

*BIRC5*

15

**Таблица S6.** Профили некоторых известных факторов транскрипции, входящие в Мотив 1

Субмотив

Профили факторов транскрипции

1a

MEF2A, MEF2B, MEF2C, MEF2D, FOXE1, Nr2e1, PHOX2B, Pou2f2, Pou2f3, POU3F1, POU3F2, Sp100 (первичный)

1b

Rxra (вторичный), Dmrt1, Irf3 (вторичный)

1c

ATF6, CREB3, CREB3L1, Creb3l2, Pou6f1, ZNF135

**Таблица S7.** Профили узнавания условно опухолеспецифических факторов транскрипции в химерных промоторах

Примечание: (+) – присутствует; (–) – отсутствует.

\* Положение этих профилей в промоторах показано на рис. S4.

Фактор транскрипции

Номенклатура

JASPAR

Промоторы

CH2

CH8

CH9

CH12

CH16

CH20

CH26

CH10

CMV

*PCNA*

RUNX2

MA0511.2

–

–

–

–

+

–

–

–

+

–

Creb3l2\*

MA0608.1

*+*

–

–

–

–

*+*

–

–

–

–

SREBF2

MA0596.1

–

+

–

–

+

–

–

+

–

–

ETS2\*

MA1484.1

*+*

–

–

–

–

*+*

*+*

–

–

–

CENPB

MA0637.1

+

+

+

+

–

+

+

+

–

–

HEY1\*

MA0823.1

+

–

–

–

–

+

–

–

–

–

HEY2\*

MA0649.1

*+*

–

–

–

–

*+*

–

–

–

–

RARA::RXRG

MA1149.1

–

+

+

+

+

–

–

+

+

–

Srebf1(var.2)\*

MA0829.1

*+*

–

–

–

–

*+*

–

–

–

–

ZNF75D

MA1601.1

–

–

–

+

+

+

–

+

+

+

Zfx

MA0146.2

–

–

–

–

+

+

–

+

–

+

HOXD11

MA0908.1

–

–

–

–

–

–

–

–

–

+

KLF11

MA1512.1

–

–

–

–

+

–

–

–

–

–

MEF2A

MA0052.4

–

–

–

–

–

–

–

–

–

–

RORB

MA1150.1

–

–

–

–

+

–

–

+

–

–

SIX1

MA1118.1

–

–

–

–

–

–

–

–

–

–

Stat6

MA0520.1

–

–

–

–

–

–

–

–

–

–

**Таблица S8.** Профили узнавания условно неспецифических факторов транскрипции в химерных промоторах

Примечание: (+) – присутствует; (–) – отсутствует.

Фактор транскрипции

Номенклатура

JASPAR

Промоторы

CH2

CH8

CH9

CH12

CH16

CH20

CH26

CH10

CMV

*PCNA*

Mafb

MA0117.2

–

–

–

–

+

–

–

+

+

+

Arid3a

MA0151.1

–

+

–

–

+

–

–

–

+

+

TBP

MA0108.2

–

–

–

–

–

–

–

–

+

+

E2F1

MA0024.3

–

–

–

–

–

–

–

–

–

–

BHLHA15(var.2)

MA1472.1

–

–

–

–

–

–

–

–

–

+

MEIS3

MA0775.1

–

–

–

–

+

–

–

–

–

+

BSX

MA0876.1

–

–

–

–

+

–

–

–

+

+