RESEARCH PAPER

Translation 1:2, e27469; July-December 2013; © 2013 Landes Bioscience

In silico motif analysis suggests an interplay of transcriptional and translational control in mTOR response

Irina A Eliseeva[1,*], Ilya E Vorontsov[2], Kirill E Babeyev[3], Sofya M Buyanova[4], Maria A Sysoeva[5], Fyodor A Kondrashov[6,7,8] and Ivan V Kulakovskiy[2,9,*]

[1]Group of Protein Biosynthesis Regulation; Institute of Protein Research; Russian Academy of Sciences; Pushchino, Russia; [2]Department of Computational Systems Biology; Vavilov Institute of General Genetics; Russian Academy of Sciences; Moscow, Russia; [3]Gymnasium MBGO; Obninsk, Russia; [4]Faculty of Bioengineering and Bioinformatics; Lomonosov Moscow State University; Moscow, Russia; [5]Mining Faculty; National Mining Resources University; St. Petersburg, Russia; [6]Bioinformatics and Genomics Programme; Centre for Genomic Regulation (CRG); Barcelona, Spain; [7]Universitat Pompeu Fabra (UPF); Barcelona, Spain; [8]Institució Catalana de Recerca i Estudis Avançats (ICREA); Barcelona, Spain; [9]Engelhardt Institute of Molecular Biology; Russian Academy of Sciences; Moscow, Russia

Keywords: mTOR, TOP, OP, sequence motif, TSS, translation regulation, motif discovery

Abbreviations: TSS, transcription start site(s); OP, oligopyrimidine tract(s) (DNA motif); TOP, terminal oligopyrimidine tract(s) (mRNA motif); mTOR, mammalian target of rapamycin; UTR, untranslated region

The short 5'-terminal oligopyrimidine tract (TOP) of 5' UTRs is a well-known regulatory sequence motif of mRNAs that are subject to growth-dependent translation. Specifically, translation of TOP mRNAs is regulated by the mTOR signaling pathway that is involved in cell proliferation, cancer development and aging. High throughput data permit detailed study of specific features of the mRNA TOP motif and its DNA origins at transcription start sites (TSS). Recently, ribosome profiling was used to identify mRNA targets of the mTOR pathway in PC3 cells. A novel pyrimidine-rich translational element (PRTE) was reported to play a key role without positional preferences within the 5' UTRs, unlike 5' TOP, which are strictly located at the 5' ends. In this study, we couple recently reported ribosome profiling data on the mTOR mRNA targets with the annotation of TSS obtained by HeliScopeCAGE. We confirm the canonical TOP and strong positional preferences of respective oligopyrimidine tracts (OP) straddling the experimentally validated TSS regions at the DNA level. Such OP localization ensures that transcription from OP segments creates the 5'-terminal TOP in the corresponding mRNAs. We demonstrate that OP are not overrepresented in downstream regions of 5' UTRs of mTOR targets. Finally, we highlight several mTOR target genes with broad and multimodal TSS spanning dozens of nucleotides that are only partically covered with an OP. Therefore, in such cases only a fraction of all produced mRNAs carry a TOP regulatory motif and, thus, respond to mTOR via TOP mechanism. We hypothesize that the interplay between transcription and translation may play a crucial role in the regulation of the mTOR response.

Introduction

The mammalian target of rapamycin (mTORC1) is one of major regulators of cell growth and proliferation, including protein biosynthesis, in higher eukaryotes.[1,2] An important role of the mTOR pathway in cancer development pushes forward studies on mTOR translational regulation.[3,4] Sequence features of 5' UTRs of mRNAs with growth-dependent translation, including those regulated by mTOR, have been studied extensively.[5] Twenty years ago, it was discovered that the short 5'-terminal oligopyrimidine tract (TOP) of 5' UTRs plays an important role in translational regulation of ribosomal genes.[6-9] Many of corresponding mRNAs carry evolutionary conserved TOP motif,[10] characterized as cytidine followed by an uninterrupted oligopyrimidine sequence of 5–14 nt, which are located at the 5' ends of 5' UTRs.[5]

In parallel, the TCT-promoters were discovered as a specific class of promoter sequences. Normal transcription from a TCT-promoter depends on the TCT-motif localized at transcription start site (TSS).[11,12] This suggests a possible interplay between transcription and translation through a double-purpose pyrimidine-rich motif. Furthermore, commonly used gene annotations do not always correctly localize TSS, which must be taken into account when studying terminal 5' UTR sequences.[13]

High-throughput ribosome profiling experiments challenged the assumption that the mTOR mRNAs carry terminal TOP motif suggesting that some mRNAs may contain nearly terminal TOP-like motifs[14] or even downstream pyrimidine-rich translational elements (PRTE).[15] Databases of experimentally determined TSS, such as DBTSS (http://dbtss.hgc.jp/),[16] were

*Correspondence to: Ivan V Kulakovskiy and Irina A Eliseeva; Email: ivan.kulakovskiy@gmail.com and yeliseeva@vega.protres.ru
Submitted: 06/12/2013; Revised: 11/10/2013; Accepted: 12/09/2013; Published Online: 12/11/2013
Citation: Eliseeva IA, Vorontsov IE, Babeyev KE, Buyanova SM, Sysoeva MA, Kondrashov FA, Kulakovskiy IV. In silico motif analysis suggests an interplay of transcriptional and translational control in mTOR response. Translation 2013; 1:e27469; http://dx.doi.org/10.4161/trla.27469

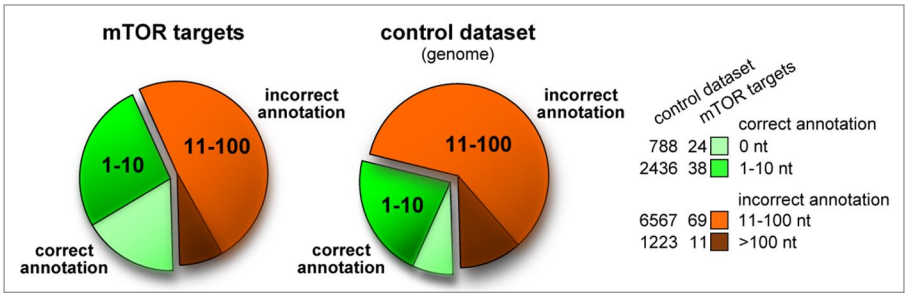

Figure 1. The number of genes with a given distance between annotated and verified TSS. Data for the mTOR mRNA targets and the control whole-genome data set are given. TSS of the hg18 genome assembly are compared with HeliScopeCAGE peak maxima. For each UCSC annotated transcript, HeliScopeCAGE values are extracted in the region from +100 nt upstream of the annotated TSS to the 3' end of the annotated 5' UTR. For a given gene, the resulting distance is selected as the best (minimal) value over all UCSC transcripts.

complicating the picture, having limited agreement with the annotated TSS while providing rather sparse data.

A new source of quantitative data on TSS appeared with HeliScopeCAGE, Cap Analysis of Gene Expression using Helicos single-molecule sequencing.[17] HeliScopeCAGE allowed high-throughput quantitative analysis of TSS at a single-nucleotide resolution. Thus, it is tempting to couple high-throughput translational (ribosome profiling) and transcriptional (HeliScopeCAGE) data to study the sequence motifs involved in the mTOR translational response.

To further elucidate the role of oligopyrimidine motifs in mTOR regulation we, first, confirm the TOP motif by de novo motif discovery in 5' UTRs of mTOR target genes and highlight oligopyrimidine tracts, OP, straddling respective TSS. Second, we demonstrate that a commonly used TSS annotation is not precise enough to study features of terminal sequences in 5' UTRs. Third, we apply a motif finding approach to the 5' UTRs of mTOR targets and demonstrate significant OP enrichment in the vicinity of TSS and no significant enrichment in the downstream segments of the 5' UTRs. Fourth, we describe different OP motif subtypes for broad and narrow TSS. A possible functional role of broad and multimodal TSS only partially covered by OP is discussed.

Results

Known TOP motif confirmed by de novo motif discovery

By de novo motif discovery we have found a CT-rich motif similar to the mRNA TOP using 250 5' UTRs extended upstream by additional 100 nt to take into account possible TSS misannotation. The sequences corresponded to different 5'UTRs of UCSC annotated transcripts for 142 human mTOR target genes that were recently described by Hsieh et al.[15]

The 14 nt motif (see **Figure 4a** below) was in agreement with the previously reported motifs of comparable lengths with major cytosine in the TCT context.[10,13] The motif occurrences were found in 214 of 250 5' UTR sequences (for 126 of 142 genes). It is worth noting that motif occurrences often contained at least one purine (184/108 5' UTRs/genes) and in many cases (84/58 5' UTRs/genes) the purine was found in one of the 8 middle positions, implying that the motif does not seem to conform to a strict "continuous" pyrimidine-only structure.

By analogy with the mRNA TOP signal, the Terminal OligoPyrimidine tract, we refer to the corresponding OligoPyrimidine tract in DNA as OP.

OP/TOP has strict positional preferences to TSS/5' end of 5' UTR

The existing TSS annotation of mTOR targets limits the study of OP/TOP

To assess positional preferences relative to the 5' end of 5' UTRs, a precise annotation of transcription start sites is required. Experimental data sets such as DBTSS[16] provide a notable improvement over common genome annotation. In this study, we use the recent HeliScopeCAGE data set that provides quantitative data on TSS at single base pair resolution.[17]

Compared with the control, i.e., the whole set of protein-coding genes excluding mTOR targets, the commonly used TSS annotation (as in UCSC Genome Browser[18]) is more accurate for the mTOR targets (**Fig. 1**). However, even for the mTOR target genes the existing annotation is poor: less than half of the mTOR targets have an annotated TSS more than 10 nt apart from the respective HeliScopeCAGE peak maximum (which denote the TSS of the major mRNA isoforms; see also the Supplementary file 1). Thus, usage of existing TSS annotation would not allow studying proper positioning of the TOP relative to the 5' ends of 5' UTRs and of the OP relative to TSS.

OP positional preferences

To study OP positional preferences we assessed the OP location relative to the verified TSS. We counted the number of 5' UTRs with OP hits at a given fixed position relative to the HeliScopeCAGE peak maximums (the precise TSS of major transcribed mRNA isoforms). We computed the statistical significance of the association between the mTOR targets and OP hits in 5' UTRs under the null hypothesis that the frequency and location of OP hits in the 5' UTRs of mTOR targets are the same as in the control data set of other protein-coding genes.

The enrichment of OP upstream and in TSS-overlapping locations was highly significant ($P << 0.05$) independently of the OP detection threshold (**Fig. 2**, Supplementary File 2). Enrichment for nearby downstream OP had lower significance

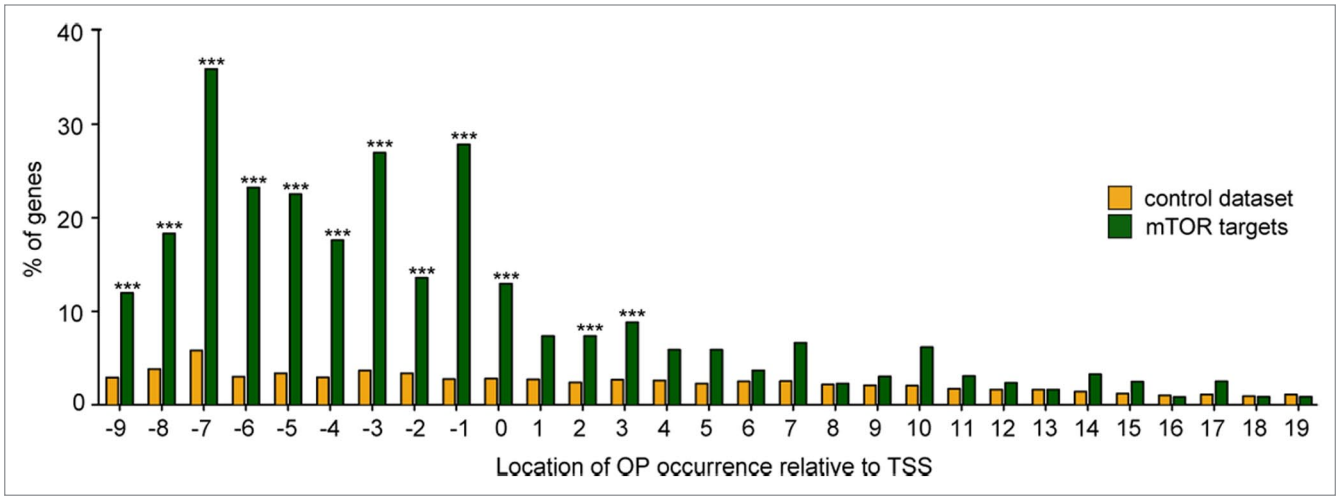

Figure 2. The fraction of genes (Y axis) having OP motif occurrence at a given relative distance (X axis) from the HeliScopeCAGE peak maximum. Motif occurrences are detected at 0.005 *P* value (data for other detection thresholds are shown in the Supplementary file 2). A significant difference between the mTOR targets and the control whole-genome data set is marked with *** (Fisher's exact test for 5' UTRs of sufficient length, $P < 0.05$, Holm adjustment for multiple testing).

with some OP in the +1 through +4 positions showing $P < 0.05$ significance level (depending on the OP detection threshold). Also, we inspected the general OP positional preferences by looking for OP hits not farther than +19 nt downstream from the verified TSS (heads of the 5' UTRs) and in the remainders of the 5' UTRs, i.e., more than +19 nt downstream of the TSS (tails of the 5' UTRs). Comparing to the control data set, OP hits were significantly overrepresented in heads ($P << 0.05$) but not in tails of the 5' UTRs of mTOR targets ($P >> 0.05$, see Supplementary File 2). De novo motif discovery in 5' UTR tails was also unable to detect any overrepresented pyrimidine-rich motifs.

The OP positional preference to TSS and the absence of other pyrimidine-rich motifs called into question the existence of the PRTE elements recently reported by Hsieh and colleagues,[15] which were claimed to be located in downstream regions of 5' UTRs.

OP are different for narrow and broad promoters

Translational genes are mostly transcribed from a special class of promoters, the TCT-promoters, which are characterized by the central TCT consensus and a TSS-straddling OP with a strict TSS positioning on the inner invariant C. The importance of the TCT consensus at the transcriptional level has been validated experimentally.[12] This suggests an interplay between transcriptional and translational regulatory mechanisms, since the transcriptional TCT-motif significantly influences the presence of TOP in the transcribed mRNAs.

At the same time, not all mTOR-responsive mRNAs are transcribed from narrow TSS, such as TCT-promoters. We utilized the quantitative nature of HeliScopeCAGE data to estimate typical TSS width of mTOR targets. First, we tested the fraction of mRNAs transcribed from a particular genomic position having the maximal HeliScopeCAGE peak signal (**Supplementary Figure 1**). We found that more than half of mTOR targets have less than 50% of mRNAs transcribed from a single major TSS position. Thus, the mTOR targets, even taking into account the TCT-promoters, poorly fit the "single-position TSS" assumption.

Thus, for each UCSC annotated transcript we estimated the TSS width as the minimal width of the genomic region such that at least 2/3 of the mRNA pool would be transcribed from that region (**Fig. 3**, Supplementary File 3). Even with the best (minimal) values taken into account for each gene, nearly 30% of mTOR targets had a TSS width of over 10 nt.

We performed de novo motif discovery separately for 167 sequences of extended 5' UTRs with narrow (width ≤ 10 nt) and 83 with broad (width > 10 nt) TSS. We found the motifs (Supplementary File 4) shown in **Figure 4** (**b**,**c**) in 5' UTRs of 138 and 63 transcripts with narrow and broad TSS, respectively. The motifs were similar, except for the major T in a motif found for broad TSS indirectly resembling the PRTE motif, suggesting that previous studies were possibly affected by these broad TSS.

OP and OP-free modes may exist in a single multimodal TSS

The OP motif straddles experimentally verified TSS regions whereby after transcription the corresponding mRNAs begin with the TOP sequence. The shape of the TSS peaks defines the fraction of the mRNA pool transcribed from each particular genomic position. For broad TSS, OP may cover only a fraction of the TSS region. An extreme example is given by multimodal TSS exhibiting several sharp peaks (modes) with not all of them carrying the start of the OP occurrence in permissive -9..+4 positions. We report several mTOR targets possibly having TSS with such properties (Supplementary File 5), supporting the hypothesis that mTOR translational response may be regulated on the transcriptional level. **Figure 5** shows annotation of the classical TOP gene PABPC1, UBA52 with a multimodal TSS, and YBX1 as a special case with a broad TSS and a weak OP occurrence.

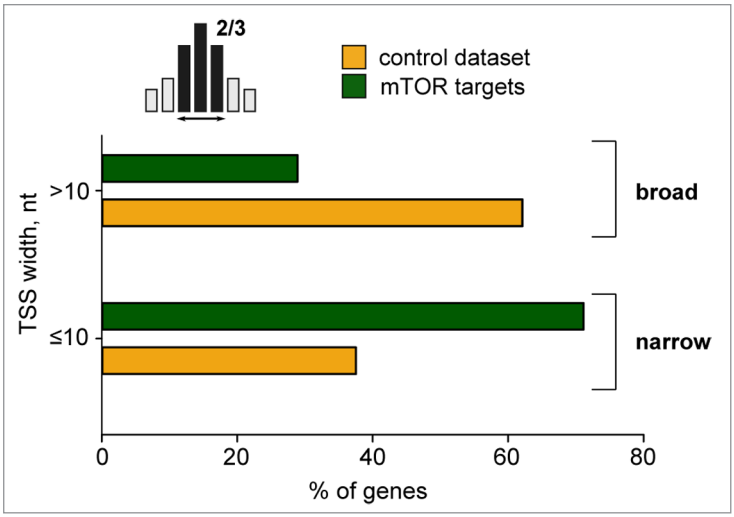

Figure 3. The percentage of genes (X axis) having a given TSS width (Y axis). TSS width is defined as the minimal length of the region aggregating at least 2/3 of the total HeliScopeCAGE signal. For each gene, the best (minimal) value is taken among all corresponding transcripts of the UCSC genome annotation.

Discussion

While our approach yielded several interesting observations on OP/TOP motifs, it is somewhat limited for several reasons. First of all, both the TSS width and the length of the OP region can vary. Thus, the gapless multiple local alignment and the positional weight matrix, typically used for motif modeling, may be not fully suitable to study the OP/TOP phenomena. Therefore, it may be worthwhile to look for a more flexible in silico model of the OP/TOP motifs. Another problem is connected with the TCT-promoter motif[12] having its own structure linked with corresponding transcriptional mechanisms, making it is difficult to clearly separate the transcriptional and translational sequence features.

Also, HeliScopeCAGE data were produced for THP-1 and HeLa cells,[17] while PC3 cells were used for ribosome profiling.[15] Many mTOR targets are housekeeping genes with strictly defined TSS, with good agreement between HeliScopeCAGE peaks for THP-1 and HeLa (**Supplementary Figure 2**). However, it is possible that some TSS with cell type-specific expression were missed or shifted from PC3-specific locations and falsely detected as OP-overlapping or OP-free.

We detected dozens of target genes with OP-free TSS, TSS with weak OP occurrences, and broad TSS with fuzzy pyrimidine tracks. One such example is YBX1 with a very broad TSS leading to rare mRNA isoforms, one of which was previously used to validate functionality of putative PRTE.[15] Recent experimental verification[19] showed that translation of the truncated TOP-free YBX1 mRNA, that lacks any CT-rich segments in its 5' UTR, is still regulated by mTOR (rabbit YBX1 5' UTR having 100% identity with the truncated human sequence). Therefore, it is likely that mTOR translational response of YBX1 mRNA is controlled through an alternative OP-independent mechanism, possibly involving other regulatory elements.

We were unable to confirm the presence or significance of pyrimidine-rich motifs, such as PRTE,[15] downstream of TSS. However, compared with the control genome-wide data set, the 5' UTRs of mTOR targets have special features, e.g., generally shorter lengths and pyrimidine-rich nucleotide composition (Supplementary file 6). Therefore, the question of basic sequence-level features of 5' UTRs and their correct statistical evaluation remains unresolved. The presence and significance of OP GC-tails as well as other downstream sequence features also require further investigation.[6,20]

The discovered multimodal OP/OP-free TSS would benefit from experimental study. TSS modes may be simultaneously active, or may switch on and off in a particular tissue or growth conditions. It was shown, that growth-dependent translation of ribosomal mRNAs can be different in different cell types.[21] Furthermore, in different tissues a particular gene can be transcribed from alternative TSS producing either TOP or TOP-free mRNAs with different base translational efficiency.[22] We suppose that new high-throughput experimental data on tissue-specific TSS activity would provide additional insights into the transcription-translation regulatory interplay.

Conclusion

Transcription from a TSS inside the OP results in the TOP motif being present in the transcribed mRNA. Our observations on multimodal OP/OP-free TSS suggest that fine tuning of mTOR-driven regulation may occur at the transcription-translation interface. OP motifs associated with TOP mRNAs may contain purine substitutions and can be slightly shifted downstream of the major TSS.

We would like to emphasize the importance of precise transcription start annotation for sequence analysis of 5' UTRs. Usage of improper annotation is unreliable leading to questionable discoveries, such as the downstream PRTE motif of mTOR mRNA targets, which we were unable to confirm neither by motif finding nor by de novo motif discovery.

Finally, we provide a comprehensive annotation of OP/TOP motifs in 5' UTRs of mTOR mRNA targets along with the corresponding HeliScopeCAGE peak profiles as TSS annotation (Supplementary file 1).

Materials and Methods

HeliScopeCAGE data preparation

HeliScopeCAGE peak data of THP-1 and HeLa cells were taken from http://fantom.gsc.riken.jp/5/suppl/Kanamori-Katayama_et_al_2011/.[17] For each of these two cell lines, values at each genomic position were averaged across replicates (rounding down to the nearest integer value) discarding the positive values if present only in a single replicate. The resulting data were then averaged across the two cell lines (rounding down to the nearest integer value as well). The HeliScopeCAGE signal

is proportional to the number of mRNAs transcribed from each particular genomic position.

Sequence data sets preparation

The set of protein-coding genes mapped to Entrez gene IDs was extracted from the HGNC database.[23] Corresponding transcripts and basic TSS annotations were taken according to the UCSC hg18 human genome annotation.[18] Intronic sequences were removed; 5' UTRs were additionally extended with +100 nt upstream segments to account for possibly misannotated TSS.

144 mTOR target genes identified in PC3 cells by ribosome profiling were taken from Hsieh et al.[15] 142 genes had UCSC-to-Entrez mapping and at least one UCSC transcript annotation with non-empty HeliScopeCAGE peak profile. We refer to these 142 genes and respective extended 250 5' UTRs (Supplementary file 7) based on UCSC transcript annotations as the mTOR targets.

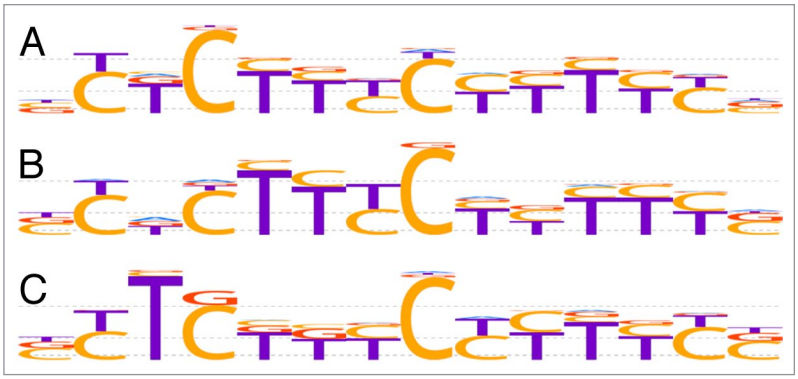

Figure 4. Motifs discovered in the whole set of 5' UTR sequences of mTOR targets (**A**), and subsets of narrow (**B**) and broad (**C**) TSS. Positional weight matrices are given in the Supplementary file 4.

The set of all remaining protein-coding genes curated in HGNC[23] with UCSC-to-Entrez mapping and non-empty HeliScopeCAGE peak profile for at least one UCSC annotated transcript was used as the control data set (17671 5' UTRs of 11027 genes in total).

De novo motif discovery

We used ChIPMunk motif discovery tool (http://autosome.ru/ChIPMunk/) for de novo motif discovery[24] in 5' UTRs of mTOR targets. MEME software[25] was additionally applied on the same data sets ensuring agreement with the ChIPMunk results. ChIPMunk was searching for the optimal gapless multiple local alignment accounting for local nucleotide composition. Motif discovery was performed in the extreme precision mode (with 100 times more random seeds than by default). The automatically determined motif length was 25 with 14 nt long OP sequence flanked by GC-rich tails. Thus, the fixed length of 14 nt was adopted to rediscover the OP motif. The presence and significance of GC-tails as well as GA-rich motifs downstream of TSS requires further investigation.

Motif finding in 5' UTRs

To search for OP motif occurrences, we estimated thresholds for the positional weight matrix (PWM) model representing the motif.[26] Fixed thresholds (low, medium, high) were selected by MACRO-APE software[27] (http://autosome.ru/macroape/) using the strategy of Touzet and Varre[28] according to the motif *P* value of 0.005, 0.0005 and 0.00005.

Assessing OP positional preferences

For the mTOR targets and control data sets we counted 5' UTRs containing an OP/TOP occurrence starting at a given position relative to the HeliScopeCAGE peak maximum (denoting the precise TSS of the major mRNA isoform). The distance of zero corresponded to the OP occurrence starting exactly at the HeliScopeCAGE peak maximum. For each gene all respective annotated transcripts were considered (i.e., a gene was considered to contain OP at a particular position if any of respective transcripts contained OP at that position in 5' UTR). Fisher's exact test was used to estimate *P* values. For exact OP positioning relative to TSS, *P* values were adjusted for multiple testing by Holm correction.[29]

Detection of OP/OP-free multimodal TSS

The UCSC annotated transcript was identified as a carrying multimodal TSS if there were at least two positions with HeliScopeCAGE peak heights of no less than 10% of the total signal for the whole region. From the list of sequences with multimodal TSS we selected those carrying OP (using the low detection threshold) for some but not all of its modes. The allowed OP location (the location of the starting OP nucleotide) relative to a given TSS position was adopted as (-9..+4), where the left boundary (-9) ensured that at least 5 nt (as the commonly accepted minimal TOP length) of the 14 nt OP were included in the transcribed mRNA. The permissible downstream location up to +4 was estimated in this study (see corresponding section in the Results).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The authors thank the School for Molecular and Theoretical Biology of the Dynasty Foundation and Anya Piotrovskaya for making this research possible. We thank Yulia Medvedeva, Dmitry Chudakov, Mikhail Gelfand and Dmitry Lyabin for fruitful discussions and Eugenia V. Serebrova for help in manuscript preparation. This work was supported, in part, by the Dynasty Foundation Fellowship to I.V.K., the Russian Foundation for Basic Research (#14-04-01838) and the Program on Molecular and Cell Biology from the Presidium of RAS.

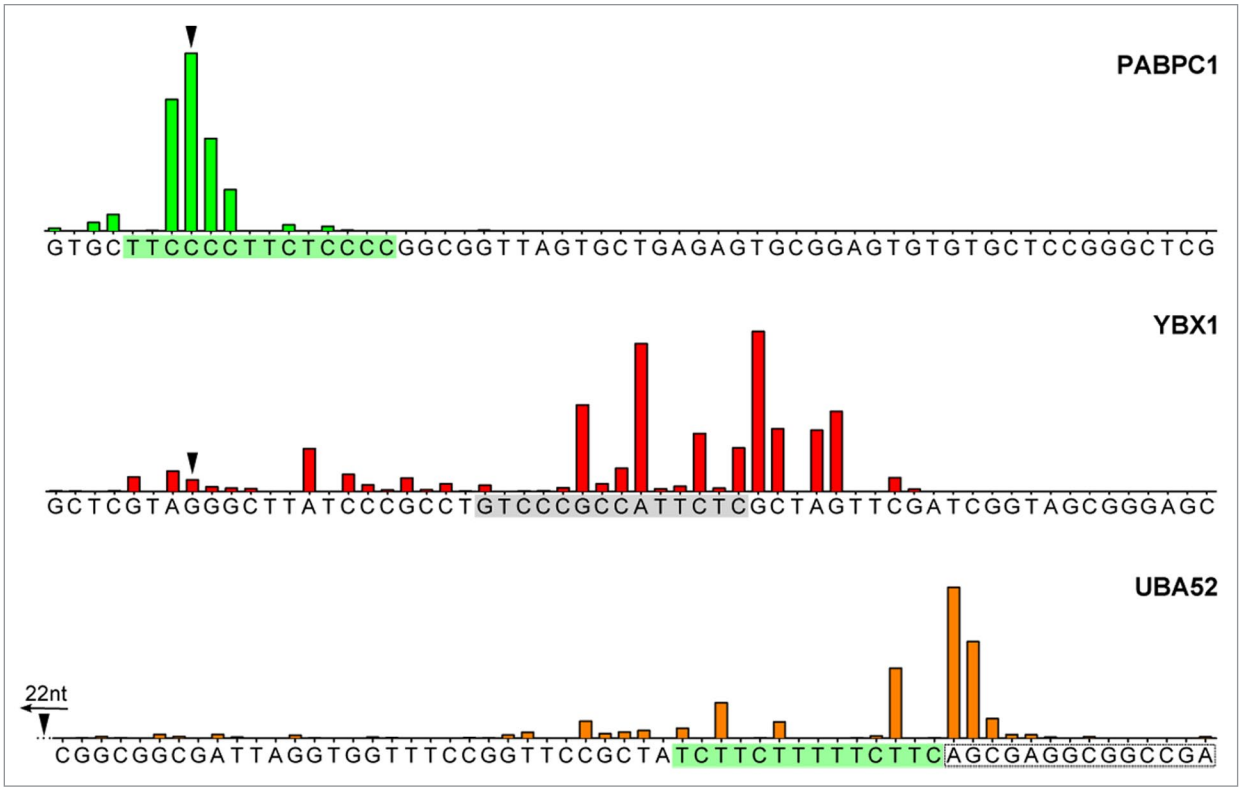

Figure 5. HeliScopeCAGE signal and OP motif occurrences in extended 5' UTRs of three mTOR targets: PABPC1 (top panel), YBX1 (middle panel), UBA52 (bottom panel). The UCSC annotated TSS (hg18 genome annotation) are denoted by black triangles. Best OP motif occurrences in the region are highlighted. UBA52 has a multimodal TSS only partially covered by OP.

References

1. Johnson SC, Rabinovitch PS, Kaeberlein M. mTOR is a key modulator of ageing and age-related disease. Nature 2013; 493:338-45; PMID:23325216; http://dx.doi.org/10.1038/nature11861.
2. Wang X, Proud CG. mTORC1 signaling: what we still don't know. J Mol Cell Biol 2011; 3:206-20; PMID:21138990; http://dx.doi.org/10.1093/jmcb/mjq038.
3. Zoncu R, Efeyan A, Sabatini DM. mTOR: from growth signal integration to cancer, diabetes and ageing. Nat Rev Mol Cell Biol 2011; 12:21-35; PMID:21157483; http://dx.doi.org/10.1038/nrm3025.
4. Topisirovic I, Sonenberg N. mRNA translation and energy metabolism in cancer: the role of the MAPK and mTORC1 pathways. Cold Spring Harb Symp Quant Biol 2011; 76:355-67; PMID:22123850; http://dx.doi.org/10.1101/sqb.2011.76.010785.
5. Meyuhas O. Synthesis of the translational apparatus is regulated at the translational level. European journal of biochemistry / FEBS 2000; 267:6321–30.
6. Avni D, Shama S, Loreni F, Meyuhas O. Vertebrate mRNAs with a 5'-terminal pyrimidine tract are candidates for translational repression in quiescent cells: characterization of the translational cis-regulatory element. Mol Cell Biol 1994; 14:3822-33; PMID:8196625.
7. Jefferies HB, Thomas G, Thomas G. Elongation factor-1 alpha mRNA is selectively translated following mitogenic stimulation. J Biol Chem 1994; 269:4367-72; PMID:8308006.
8. Terada N, Patel HR, Takase K, Kohno K, Nairn AC, Gelfand EW. Rapamycin selectively inhibits translation of mRNAs encoding elongation factors and ribosomal proteins. Proc Natl Acad Sci USA 1994; 91:11477-81; PMID:7972087; http://dx.doi.org/10.1073/pnas.91.24.11477.
9. Hornstein E, Git A, Braunstein I, Avni D, Meyuhas O. The expression of poly(A)-binding protein gene is translationally regulated in a growth-dependent fashion through a 5'-terminal oligopyrimidine tract motif. J Biol Chem 1999; 274:1708-14; PMID:9880551; http://dx.doi.org/10.1074/jbc.274.3.1708.
10. Perry RP. The architecture of mammalian ribosomal protein promoters. BMC Evol Biol 2005; 5:15; PMID:15707503; http://dx.doi.org/10.1186/1471-2148-5-15.
11. Rach EA, Winter DR, Benjamin AM, Corcoran DL, Ni T, Zhu J, Ohler U. Transcription initiation patterns indicate divergent strategies for gene regulation at the chromatin level. PLoS Genet 2011; 7:e1001274; PMID:21249180; http://dx.doi.org/10.1371/journal.pgen.1001274.
12. Parry TJ, Theisen JWM, Hsu J-Y, Wang Y-L, Corcoran DL, Eustice M, Ohler U, Kadonaga JT. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. Genes Dev 2010; 24:2013-8; PMID:20801935; http://dx.doi.org/10.1101/gad.1951110.
13. Yamashita R, Suzuki Y, Takeuchi N, Wakaguri H, Ueda T, Sugano S, Nakai K. Comprehensive detection of human terminal oligo-pyrimidine (TOP) genes and analysis of their characteristics. Nucleic Acids Res 2008; 36:3707-15; PMID:18480124; http://dx.doi.org/10.1093/nar/gkn248.
14. Thoreen CC, Chantranupong L, Keys HR, Wang T, Gray NS, Sabatini DM. A unifying model for mTORC1-mediated regulation of mRNA translation. Nature 2012; 485:109-13; PMID:22552098; http://dx.doi.org/10.1038/nature11083.
15. Hsieh AC, Liu Y, Edlind MP, Ingolia NT, Janes MR, Sher A, Shi EY, Stumpf CR, Christensen C, Bonham MJ, et al. The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature 2012; 485:55-61; PMID:22367541; http://dx.doi.org/10.1038/nature10912.
16. Yamashita R, Sugano S, Suzuki Y, Nakai K. DBTSS: DataBase of Transcriptional Start Sites progress report in 2012. Nucleic Acids Res 2012; 40:D150-4; PMID:22086958; http://dx.doi.org/10.1093/nar/gkr1005.
17. Kanamori-Katayama M, Itoh M, Kawaji H, Lassmann T, Katayama S, Kojima M, Bertin N, Kaiho A, Ninomiya N, Daub CO, et al. Unamplified cap analysis of gene expression on a single-molecule sequencer. Genome Res 2011; 21:1150-9; PMID:21596820; http://dx.doi.org/10.1101/gr.115469.110.
18. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, Sloan CA, Rosenbloom KR, Roe G, Rhead B, et al. The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res 2013; 41:D64-9; PMID:23155063; http://dx.doi.org/10.1093/nar/gks1048.
19. Lyabin DN, Eliseeva IA, Ovchinnikov LP. YB-1 synthesis is regulated by mTOR signaling pathway. PLoS ONE 2012; 7:e52527; PMID:23285076; http://dx.doi.org/10.1371/journal.pone.0052527.

20. Biberman Y, Meyuhas O. Substitution of just five nucleotides at and around the transcription start site of rat beta-actin promoter is sufficient to render the resulting transcript a subject for translational control. FEBS Lett 1997; 405:333-6; PMID:9108314; http://dx.doi.org/10.1016/S0014-5793(97)00234-2.
21. Avni D, Biberman Y, Meyuhas O. The 5' terminal oligopyrimidine tract confers translational control on TOP mRNAs in a cell type- and sequence context-dependent manner. Nucleic Acids Res 1997; 25:995-1001; PMID:9023110; http://dx.doi.org/10.1093/nar/25.5.995.
22. Kleene KC, Cataldo L, Mastrangelo M-A, Tagne J-B. Alternative patterns of transcription and translation of the ribosomal protein L32 mRNA in somatic and spermatogenic cells in mice. Exp Cell Res 2003; 291:101-10; PMID:14597412; http://dx.doi.org/10.1016/S0014-4827(03)00339-2.
23. Gray KA, Daugherty LC, Gordon SM, Seal RL, Wright MW, Bruford EA. Genenames.org: the HGNC resources in 2013. Nucleic Acids Res 2013; 41:D545-52; PMID:23161694; http://dx.doi.org/10.1093/nar/gks1066.
24. Kulakovskiy IV, Boeva VA, Favorov AV, Makeev VJ. Deep and wide digging for binding motifs in ChIP-Seq data. Bioinformatics 2010; 26:2622-3; PMID:20736340; http://dx.doi.org/10.1093/bioinformatics/btq488.
25. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 2009; 37:W202-8; PMID:19458158; http://dx.doi.org/10.1093/nar/gkp335.
26. Stormo GD. DNA binding sites: representation and discovery. Bioinformatics 2000; 16:16-23; PMID:10812473; http://dx.doi.org/10.1093/bioinformatics/16.1.16.
27. Vorontsov IE, Kulakovskiy IV, Makeev VJ. Jaccard index based similarity measure to compare transcription factor binding site models. Algorithms Mol Biol 2013; 8:23; PMID:24074225; http://dx.doi.org/10.1186/1748-7188-8-23.
28. Touzet H, Varré J-S. Efficient and accurate P-value computation for Position Weight Matrices. Algorithms Mol Biol 2007; 2:15; PMID:18072973; http://dx.doi.org/10.1186/1748-7188-2-15.
29. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scand J Stat 1979; 6:65-70