Dicle Tıp Dergisi / Dicle Med J (2018) 45 (4) : 415-429

www.diclemedj.org

Özgün Araştırma / Original Article

In silico analysis of biomarker potentials of miRNA-mediated ceRNAs in prostate cancer

Sercan Ergün[1]

1 Ordu University, Ulubey Vocational Higher School, 52850, Ulubey, Ordu, Turkey, ORCID: 0000-0002-6733-9848

Received: 22.05.2018; Revised: 23.07.2018; Accepted: 27.07.2018

Abstract

Objective: The objective of this study is to define novel biomarkers for Prostate Cancer (PCa) via in silico analysis that takes PCa-specific miRNAs, finds their combinatorial target genes (potential ceRNAs), selects ones containing Transcribed Ultra Conserved Region (T-UCR) among them and potentiates their relevance with PCa.

Methods: Thirty-four miRNAs of which clinical relevances with PCa were proved experimentally were exported via miRWalk database.Using the ComiR database, 859 genes targeted by these 34 miRNAs simultaneously were identified. Genes with ComiR score above 0.911 were taken into account. Genes containing T-UCR and showing potential ceRNA activity were extracted. Among PCa-associated ceRNAs including T-UCR, we identified genes with significant expression differences between PCa and normal prostate tissue using the GEPIA database. The statistical evaluation of the association of NFAT5 and PTBP2 genes with PCa was performed by Spearman correlation test in GEPIA database.

Results: PCa-associated ceRNAs cross-matching with genes including T-UCR in their exonic regions were NFAT5, CLK3, PTBP2, CPEB4, MIPOL1 and TCF4. We identified genes with significant expression differences between PCa and normal prostate tissues among PCa-associated ceRNAs including T-UCR. According to this analysis, NFAT5 and PTBP2 genes were significantly less expressed in PCa than in normal prostate tissue while the others didn't show any significant differential expression pattern. NFAT5 and PTBP2 genes were found to be significantly associated with PCa ($p=0.000012$; $R=0.72$).

Conclusion: All in all, this is the study associating NFAT5 and PTBP2 genes with PCa and giving them tumor suppressive potential for PCa. Still, larger and more comprehensive studies are needed on this issue.

Keywords: Prostate cancer, miRNA, ceRNA, T-UCR, In silico analysis.

DOI: 10.5798/dicletip.497900

Yazışma Adresi / Correspondence:, *Sercan Ergün, Ordu University, Ulubey Vocational Higher School, 52850, Ulubey, Ordu, Turkey*
e-mail: sercanergun@msn.com

Prostat kanserinde miRNA aracılıklı ceRNA'ların biyobelirteç potansiyellerinin in siliko analizi

Öz

Amaç: Bu çalışmanın amacı, PK'ye özgü miRNA'ları tespit edip, onların kombinatoryal olarak hedefledikleri genleri (potansiyel ceRNA'lar) bulup, aralarından T-UCR içerenleri seçip, bunların istatistiksel korelasyon yöntemleri ile PK ile olan ilişkilerini değerlendiren in siliko analiz yoluyla PK için yeni biyobelirteçler tanımlamaktır.

Yöntemler: Klinik olarak PK ile ilişkisi deneysel olarak ispatlanmış 34 miRNA miRWalk veri tabanı kullanılarak tespit edildi. ComiR veri tabanı kullanılarak, bu 34 miRNA tarafından eş zamanlı olarak hedeflenen 859 gen tanımlandı. ComiR skoru 0.911'in üzerinde olan genler dikkate alındı. T-UCR içeren ve ceRNA aktivitesi gösteren genler bulunmuştur. T-UCR içeren PK ile ilişkili ceRNA'lar arasında, GEPIA veritabanı kullanılarak PK ve normal prostat dokusu arasındaki belirgin ekspresyon farklılıklarına sahip olan genler tanımlandı. NFAT5 ve PTBP2 genlerinin PK ile ilişkisinin istatistiksel değerlendirmesi, GEPIA veri tabanında Pearson korelasyon testi ile gerçekleştirildi.

Bulgular: Eksonik bölgelerinde T-UCR içeren genler PK-ilişkili ceRNA'lar NFAT5, CLK3, PTBP2, CPEB4, MIPOL1 ve TCF4 genleri olarak tespit edildi. T-UCR içeren PK ile ilişkili ceRNA'lar arasında PCa ve normal prostat dokuları arasında belirgin ekspresyon farklılıklarına sahip genleri tanımladık. Bu analize göre, NFAT5 ve PTBP2 genleri, PK'de normal prostat dokusundan çok daha az eksprese edilirken, diğerleri ifade düzeyi açısından anlamlı bir farklılık göstermemiştir. NFAT5 ve PTBP2 gen çiftinin PK ile anlamlı derecede ilişkili olduğunu bulduk ($p = 0.000012$; $R = 0.72$).

Sonuç: Sonuç olarak, bu çalışma PK ile NFAT5 ve PTBP2 genlerini ilişkilendiren ve bu genlere PK için tümör baskılayıcı fonksiyon öngören ilk çalışmadır. Yine de, bu konuda daha geniş ve daha kapsamlı çalışmalara ihtiyaç vardır.

Anahtar kelimeler: Prostat kanseri, miRNA, ceRNA, T-UCR, In siliko analiz.

INTRODUCTION

PCa is now considered one of the most important health problems that the male population is exposed to. In Europe, prostate cancer is seen in 214 out of 1000 men and it is the most frequent type of solid cancer in men. Prostate cancer is followed by lung and colorectal cancers, respectively. PCa is also the second most widespread reason for cancer deaths in men[1].

miRNAs are RNAs that have a length of 18-22 nucleotides, do not encode proteins and are naturally produced by cells. One of the most current and known RNA-induced silencing mechanisms is silencing by miRNAs. Regulation of gene expression through miRNAs is a new research topic that is widely found in today's science community. Until today, many miRNA studies have been realized. In these studies, even though the mechanisms of regulation of genes targeted by miRNAs have been explored or miRNA expression levels have been examined in certain diseases, information about the regulation of miRNA expression is still insufficient[2].

ceRNAs are RNA transcripts that carry common miRNA target regions by themselves and can communicate with each other by pulling miRNAs onto themselves. Reductions or deletions of transcription levels of genes carrying a common miRNA target region will cause miRNAs targeting these regions to be released and to seek new targets. These miRNAs will suppress transcriptionally their activities by selecting the ceRNAs bearing the same miRNA binding region as their new target. The increase in transcription levels of these mRNAs, which exhibit an opposite effect with ceRNA activity, will automatically reduce the effect of miRNAs on their previous targets by drawing common miRNAs on themselves. With this mechanism in mind, specific databases can be used to detect genes that may

exhibit possible ceRNA activity, as well as experimental activities[3].

Ultra Conserved Regions (UCRs), a class of genetic elements with almost-precise evolutionary conservation among various mammalian genomes, are primarily defined by comparing the human, rat, and mouse genomes. Almost 93% of UCRs can be transcribed in many normal human tissues and RNA transcribed from UCRs is regarded as T-UCR. T-UCRs function as a type of special long non-coding RNAs (lncRNAs) but have exceptional features of lncRNAs. T-UCRs may take a crucial function in development of diseases, like cancer[4].

The onset of total prostate specific antigen (total PSA) testing in blood has transformed the detection and handling of men with PCa. PSA is a powerful prognostic indicator for long-term risk of medically relevant cancer. Yet, there is a requirement for novel biomarkers that help clinical decision management about biopsy and primary medication[5]. So, the aim of this study is to define novel biomarkers for PCa via in silico analysis that takes PCa-specific miRNAs, finds their combinatorial target genes (potential ceRNAs), selects ones containing T-UCR among them and potentiates their relevance with PCa by statistical correlation methods.

METHODS

Selection of miRNAs taking role in PCa pathogenesis

Thirty-four miRNAs of which clinical relevances with PCa were proved experimentally were exported via miRWalk database. miRWalk database presents predicted and validated data on miRNA-target interaction. That type of data source empowers scientists to validate novel targets of miRNAs both on 3′-UTR and on the other parts of all known genes. The 'Validated Target module' used in this study is updated every month[6].

Analysis of PCa-specific miRNA-mediated ceRNAs

Using the ComiR database, 859 genes targeted by these 34 miRNAs simultaneously were identified. We took into account genes with ComiR score above 0.911.

ComiR is an online application for combinatorial microRNA (miRNA) target prediction. Upon uploading of a messenger RNA (mRNA) in human, mouse, fly or worm genomes, ComiR determines the potency of being targeted by a group of miRNAs. In computing the regulative potency of an mRNA from a group of miRNAs, ComiR utilizes user-provided miRNA expression levels in a combinatorial manner with suitable machine learning methods and thermodynamic modeling to perform more precise predictions. ComiR gives the possibility of being a functional target of a group of miRNAs, depending on the corresponding miRNA expression levels, for each gene[7].

We anticipate that RNA transcripts of these genes show potential ceRNA activity for these miRNAs and that they can regulate their regulation through miRNA-sponging mechanism.

Matching of PCa-associated ceRNA with genes including T-UCR

Bejerano et al. detected the UCRs in the human genome. Genes containing these regions have been classified as upstream, within (exonic) and downstream according to where it is located in the gene[8]. We also identified genes containing T-UCR in their exonic regions and extracted ones showing potential ceRNA activity in our previous analysis among them.

Analysis of PCa-associated ceRNAs including T-UCR with respect to differential gene expression between PCa and normal prostate tissues

Among PCa-associated ceRNAs including T-UCR, we identified genes with significant expression differences between PCa and

normal prostate tissue using the GEPIA database[9].

Correlation analysis of NFAT5 and PTBP2 genes in PCa

The statistical evaluation of the association of NFAT5 and PTBP2 genes with PCa was performed by Pearson correlation test in GEPIA database.

RESULTS

A list of 34 miRNAs experimentally associated with PCa using miRWalk database is given in Table I.

Table I: List of miRNAs taking role in PCa pathogenesis

sa-let-7a-5p	hsa-miR-145-5p	hsa-miR-29b-3p
sa-miR-101-3p	hsa-miR-146a-5p	hsa-miR-32-5p
sa-miR-106b-5p	hsa-miR-16-5p	hsa-miR-330-3p
sa-miR-107	hsa-miR-17-3p	hsa-miR-331-3p
sa-miR-122-5p	hsa-miR-185-5p	hsa-miR-34a-5p
sa-miR-125b-5p	hsa-miR-200c-3p	hsa-miR-377-3p
sa-miR-126-3p	hsa-miR-205-5p	hsa-miR-449a
sa-miR-126-5p	hsa-miR-211-5p	hsa-miR-521
sa-miR-1296-5p	hsa-miR-21-5p	hsa-miR-616-3p
sa-miR-138-5p	hsa-miR-217	hsa-miR-616-5p
sa-miR-141-3p	hsa-miR-221-3p	
sa-miR-143-3p	hsa-miR-26a-5p	

List of 859 genes targeted by these 34 miRNAs simultaneously was given in Supplementary 1. Genes having ComiR equal abundance score above 0.911 were listed in a decreasing order.

From the list of genes containing T-UCR according to the study of Bejerano et al., we identified genes containing T-UCR in their exonic regions (Supplementary 2)[8]. Then, we extracted ones showing potential ceRNA activity in our previous analysis among them (Table II).

Table II: List of PCa-associated ceRNAs cross-matching with genes including T-UCR in their exonic regions

NFAT5
CLK3
PTBP2
CPEB4
MIPOL1
TCF4

We identified genes with significant expression differences between PCa and normal prostate tissues among PCa-associated ceRNAs including T-UCR. According to this analysis, NFAT5 and PTBP2 genes were significantly less expressed in PCa than in normal prostate tissue while the others didn't show any significant differential expression pattern (Table III).

Table III: Expression values of PCa-associated ceRNAs including T-UCR between PCa and normal prostate tissues.

Gene ID	PCa	Normal prostate
TCF4	6.16	11.35
*NFAT5**	3.48	7.04
CLK3	47.29	71.43
*PTBP2**	6.94	14.08
CPEB4	6.58	7.71
MIPOL1	5.78	4.96

*shows significantly differential expression pattern between PCa and normal prostate tissues

A statistical evaluation of the association of NFAT5 and PTBP2 genes with PCa was performed via the GEPIA database. NFAT5 and PTBP2 gene pair were found to be significantly associated with PCa according to the Spearman correlation analysis (Figure 1). ($p=0.000012$; $R=0.72$).

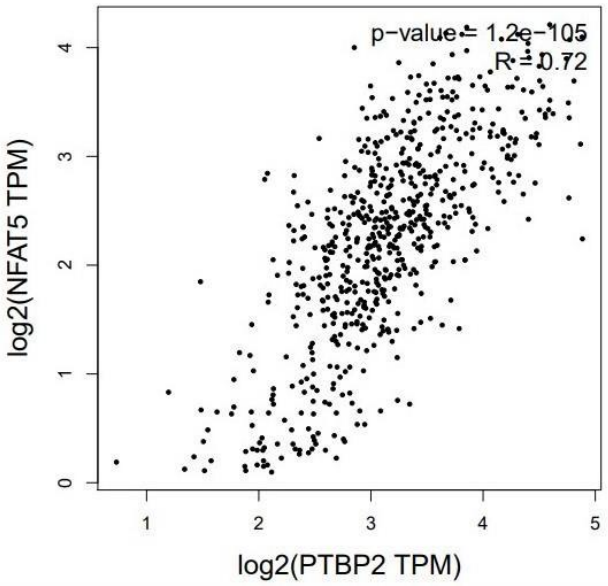

Figure 1: Spearman correlation analysis of NFAT5 and PTBP2 genes with PCa

DISCUSSION

PSA is one of the most widely used tumor markers and strongly correlates with the risk of harboring PCa. However, there is a need for novel biomarkers that aid clinical decision making about biopsy and initial treatment[5,10]. Therefore, the purpose of this study is to present novel biomarkers for PCa via in silico analysis that takes PCa-specific miRNAs, finds their combinatorial target genes (potential ceRNAs), selects ones containing T-UCR among them and potentiates their relevance with PCa by statistical correlation methods.

In this study, 34 miRNAs experimentally associated with PCa was extracted via miRWalk database (Table I). Among 859 genes targeted by these 34 miRNAs simultaneously, genes having ComiR equal abundance score above 0.911 were listed in a decreasing order. From the list of genes containing T-UCR according to the study of Bejerano et al., genes containing T-UCR in their exonic regions were identified[8]. Then, we extracted ones showing potential ceRNA activity in our previous analysis among them (Table II). Then, we selected genes with significant expression differences between PCa and normal prostate tissues among PCa-associated ceRNAs including T-UCR. According to this analysis, NFAT5 and PTBP2 genes were significantly less expressed in PCa than in normal prostate tissue while the others didn't show any significant differential expression pattern. Also, NFAT5 and PTBP2 gene pair were found to be significantly associated with PCa according to the Spearman correlation analysis.

None of NFAT5 and PTBP2 genes have been experimentally associated with prostate cancer before. Our study is the unique study to link these two genes with prostatic cancers. If we examine the role of these two genes in other types of cancer, there are contradictory results for both of them.

NFAT5 is a member of NFAT protein family having a DNA binding domain with structural similarity to the Rel-homology-region of NF-κB. Apart from NFAT1-4 proteins are moderated by calcineurin, NFAT5 is modulated by osmotic pressure at nuclear localization, transcriptional and expression levels. Upon activation, NFAT5 triggers target genes' transcription by binding to tonicity enhancer elements (TonE) in regulatory domains, like sodium-myoinositol transporter 1, aldose reductase, betaine GABA transporter, the neuropathy target esterase and taurine transporter which provide cells to stimulate cell survival in hypertonic conditions[11]. NFAT5 shows its oncogenic role in via different pathways in renal cell carcinoma, breast cancer, lung adenocarcinoma and colon cancer. NFAT5-mediated expression of S100A4 stimulates migration and proliferation of renal carcinoma cells[12]. Also, NFAT5/STAT3 interaction moderates synergism of high salt with IL-17 towards induction of VEGF-A expression in breast cancer cells[13]. Moreover, NFAT5 stimulates migration and proliferation of lung adenocarcinoma cells in part via modulating AQP5 expression[14]. Furthermore, Src kinase pathway is included in NFAT5-mediated S100A4 induction by hyperosmotic stress in colon cancer cells[15]. However, NFAT5 is tumor suppressor by inhibiting invasion and

inducing apoptosis in hepatocellular carcinoma according to the literature[16].

The protein encoded by PTBP2 gene binds to intronic polypyrimidine bundles in pre-mRNA molecules and functions in moderating the other splicing-regulatory proteins' assembly. This protein is very similar to the polypyrimidine tract binding protein (PTB) but many of its isoforms are expressed firstly in the brain[17]. Studies have shown that the PTBP2 is highly expressed in cancer cells and can promote the growth of cancer cells[18]. Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex[19]. Also, splicing factors PTBP1 and PTBP2 stimulate migration and proliferation of glioma cell lines[20]. Moreover, BCR-ABL mediated repression of miR-223 results in the activation of MEF2C and PTBP2 in chronic myeloid leukemia[21]. On the contrary, PTBP2 have some tumor suppressive roles. For example, oncogenic miR 132 sustains proliferation and self renewal potential by inhibition of polypyrimidine tract binding protein 2 in glioblastoma cells[22].

In the present study, NFAT5 and PTBP2 genes were associated with PCa as unique in the literature and our in silico analysis results foresee that they may potentially have tumor suppressive role in PCa. The fact that there are contradictory results on their roles in different cancer types may make our study results preliminary for the next in vitro and in vivo studies realized to find out exact roles NFAT5 and PTBP2 genes in PCa progression.

CONCLUSION

All in all, this is the study associating NFAT5 and PTBP2 genes with PCa for the fisrt time and giving them tumor suppressive potential for PCa. Still, larger and more comprehensive studies are needed on this issue.

Declaration of Conflicting Interests: The authors declare that they have no conflict of interest.

Financial Disclosure: No financial support was received.

REFERENCES

1. Yikilmaz TN, Öztürk E. Yüksek Riskli Prostat Kanserinde Radikal Prostatektomi/Radical Prostatectomy In High-Risk Prostate Cancer. Dicle Med J. 2016; 43 :419.
2. Saydam F, Değirmenci İ, Güneş HV. MicroRNAs and cancer. Dicle Med J. 2011; 38 :113-20.
3. Ergun S, Oztuzcu S. Oncocers: ceRNA-mediated cross-talk by sponging miRNAs in oncogenic pathways. Tumor Biol. 2015; 36 :3129-36.
4. Zhou J, Wang R, Zhang J, et al. Conserved expression of ultra-conserved noncoding RNA in mammalian nervous system. BBA-Gene Regul Mech. 2017; 1860 :1159-68.
5. Bodakçi MN, Bozkurt Y, Atar M, et al. The results of transrectal prostate biopsy in patients with low levels of prostate specific antigen. Dicle Med J. 2012; 39.
6. Dweep H, Gretz N. miRWalk2. 0: a comprehensive atlas of microRNA-target interactions. Nat Methods. 2015; 12 :697-.
7. Coronnello C, Benos PV. ComiR: combinatorial microRNA target prediction tool. Nucleic Acids Res. 2013; 41(W1): W159-W64.
8. Bejerano G, Pheasant M, Makunin I, et al. Ultraconserved elements in the human genome. Science. 2004; 304(5675): 1321-5.
9. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017; 45(W1): W98-W102.
10. Ediz C, İhvan AN, Hayit H, et al. Positive Predictive Values in Diagnosis of Incidental Prostate Cancer. Dicle Med J. 2016; 43.
11. Cheung CY, Ko BC. NFAT5 in cellular adaptation to hypertonic stress–regulations and functional significance. J Mol Signal. 2013; 8 :5.
12. Küper C, Beck F-X, Neuhofer W. NFAT5-mediated expression of S100A4 contributes to proliferation and migration of renal carcinoma cells. Front Physiol. 2014; 5: 293.
13. Amara S, Alotaibi D, Tiriveedhi V. NFAT5/STAT3 interaction mediates synergism of high salt with IL-17 towards induction of VEGF-A expression in breast cancer cells. Oncol Lett. 2016; 12 :933-43.

14. Guo K, Jin F. NFAT5 promotes proliferation and migration of lung adenocarcinoma cells in part through regulating AQP5 expression. Biochem Bioph Res Co. 2015; 465 :644-9.

15. Chen M, Sastry SK, O'Connor KL. Src kinase pathway is involved in NFAT5-mediated S100A4 induction by hyperosmotic stress in colon cancer cells. Am J Physiol-Cell Ph. 2011; 300 :C1155-C63.

16. Qin X, Wang Y, Li J, et al. NFAT5 inhibits invasion and promotes apoptosis in hepatocellular carcinoma associated with osmolality. Neoplasma. 2017; 64 :502-10.

17. Licatalosi DD, Yano M, Fak JJ, et al. Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. Gene Dev. 2012; 26 :1626-42.

18. He X, Pool M, Darcy K, et al. Knockdown of polypyrimidine tract-binding protein suppresses ovarian tumor cell growth and invasiveness in vitro. Oncogene. 2007; 26 :4961.

19. Ji Q, Zhang L, Liu X, et al. Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex. Brit J Cancer. 2014; 111 :736.

20. Cheung HC, Hai T, Zhu W, et al. Splicing factors PTBP1 and PTBP2 promote proliferation and migration of glioma cell lines. Brain. 2009; 132 :2277-88.

21. Agatheeswaran S, Singh S, Biswas S, et al. BCR-ABL mediated repression of miR-223 results in the activation of MEF2C and PTBP2 in chronic myeloid leukemia. Leukemia. 2013; 27 :1578.

22. Lou S, Ji J, Cheng X, et al. Oncogenic miR‐132 sustains proliferation and self‐renewal potential by inhibition of polypyrimidine tract‐binding protein 2 in glioblastoma cells. Mol Med Rep. 2017; 16 :7221-8.

Supplementary 1: List of genes targeted by these 34 PCa-associated miRNAs simultaneously

Gene ID	ComiR equal abundance score
SMAD2	0.9241
RPS6KA5	0.9241
FLRT2	0.9241
ATXN3	0.9241
GNAI3	0.9241
GRIN2B	0.9241
ABI2	0.924
TMED3	0.924
AAK1	0.924
SRGAP1	0.924
KCNC4	0.924
ARIH1	0.924
MGAT4C	0.924
CNKSR3	0.924
HEMK1	0.9239
DNAJC10	0.9239
SLC35E3	0.9239
PEX26	0.9239
ZBTB37	0.9239
PCDH9	0.9239
AGO3	0.9239
NTRK3	0.9239
ZC3H14	0.9238
TSPAN14	0.9238
PEAK1	0.9238
SLITRK5	0.9238
CCDC85C	0.9238
KMT2C	0.9238
RBM28	0.9237
DGKH	0.9237
ZNF26	0.9237
XKR4	0.9237
KLRD1	0.9237
PLEKHA1	0.9237
ONECUT2	0.9237
USP8	0.9236
SOD2	0.9236
UBN2	0.9236
ZBTB8B	0.9236
KCNJ6	0.9236
ZNF207	0.9236
KYNU	0.9235
PCNXL4	0.9235
SYNE3	0.9235
FAM83F	0.9235
MRPL42	0.9235
NCKAP1	0.9235
GAN	0.9235
POU2F1	0.9235
ZNF704	0.9235
GUCY1A2	0.9234
INTS6	0.9234
ZNF431	0.9234
SUGT1	0.9234
HIF1AN	0.9234
TNRC6B	0.9234
SLC8A1	0.9234
YIPF4	0.9233
USP15	0.9233
SDHC	0.9233
LPP	0.9233

RPAP2	0.9233	*HIPK2*	0.9228	*INO80D*	0.9223
GPATCH2L	0.9233	*ABL2*	0.9228	*FRK*	0.9223
KLC1	0.9232	*REV3L*	0.9228	*GTDC1*	0.9222
NAP1L1	0.9232	*PTCHD1*	0.9228	*REL*	0.9222
ST8SIA5	0.9232	*INF2*	0.9227	*FGFR1OP*	0.9222
RASAL2	0.9232	*KCTD16*	0.9227	*CDK6*	0.9222
ZNF891	0.9232	*SLC16A10*	0.9227	*MACC1*	0.9222
NFAT5	0.9232	*TMEM178B*	0.9227	*RAD51D*	0.9221
EPN1	0.9231	*PLEKHA3*	0.9226	*CACNA1E*	0.9221
RAP1B	0.9231	*HELB*	0.9226	*LRRK1*	0.9221
CPSF2	0.9231	*SLC16A7*	0.9226	*CBX5*	0.9221
ZNF562	0.9231	*C15orf40*	0.9226	*RORA*	0.9221
TTC7B	0.9231	*SDR42E1*	0.9226	*ARL10*	0.9221
CNNM2	0.9231	*B3GALT5*	0.9226	*TNRC6A*	0.9221
HOOK3	0.9231	*FAM204A*	0.9226	*PTCH1*	0.9221
GOLGB1	0.9231	*KSR2*	0.9226	*MDM2*	0.922
TTL	0.923	*ZNF268*	0.9225	*NDUFA9*	0.922
KCNN3	0.923	*NEGR1*	0.9225	*INTU*	0.922
PPP1R12B	0.923	*PPP2R2B*	0.9225	*CD84*	0.922
FUT9	0.923	*TMOD2*	0.9225	*PTPN14*	0.922
STX7	0.923	*SAMD12*	0.9225	*PRTG*	0.922
DTWD1	0.9229	*WDR7*	0.9224	*LONRF2*	0.922
SPRYD4	0.9229	*AP1M1*	0.9224	*ZNF605*	0.9219
ACVR2B	0.9229	*ILDR2*	0.9224	*VWC2*	0.9219
CLK3	0.9229	*NABP1*	0.9224	*CSRNP3*	0.9219
MAVS	0.9229	*FMNL3*	0.9224	*SFT2D2*	0.9218
SYT16	0.9229	*PURA*	0.9224	*CHST9*	0.9218
PPP2R2D	0.9229	*NRDE2*	0.9224	*SYT14*	0.9218
WTIP	0.9229	*ZBTB25*	0.9224	*ITM2B*	0.9218
AGO1	0.9229	*MBNL3*	0.9224	*FAM126A*	0.9218
BNC2	0.9229	*KLF12*	0.9224	*PLEKHG4B*	0.9218
RAB21	0.9228	*CFLAR*	0.9223	*CAMK4*	0.9218
TRIM44	0.9228	*NA*	0.9223	*ENTPD1*	0.9218
PCDHA10	0.9228	*RGMA*	0.9223	*HEBP2*	0.9218
MDM4	0.9228	*C16orf72*	0.9223	*C17orf51*	0.9218
GREM1	0.9228	*FAM227A*	0.9223	*CLSTN2*	0.9218
PPM1L	0.9228	*CDS2*	0.9223	*CYP20A1*	0.9217

ELK4	0.9217	*GALR1*	0.9211	*NQO2*	0.9206
ASXL2	0.9217	*EMC10*	0.9211	*SLC5A3*	0.9206
WNT2B	0.9216	*IYD*	0.9211	*PLXDC2*	0.9206
DDHD1	0.9216	*TBC1D16*	0.9211	*KIAA1958*	0.9206
CDKL5	0.9216	*PAQR3*	0.9211	*MAP3K9*	0.9206
MMP16	0.9216	*G3BP1*	0.9211	*RNF24*	0.9205
ORAI2	0.9216	*NHLRC2*	0.9211	*PURB*	0.9205
MDGA1	0.9216	*SEMA5A*	0.9211	*NA*	0.9205
FAXC	0.9216	*TRHDE*	0.921	*CPEB3*	0.9205
ENAH	0.9215	*FGF14*	0.921	*PHACTR2*	0.9205
AGAP1	0.9215	*MGAT4A*	0.921	*SNTB2*	0.9205
WDFY2	0.9215	*PPARGC1B*	0.921	*NFIA*	0.9205
DCTN5	0.9215	*RAB3B*	0.921	*EIF4EBP2*	0.9205
NPFFR1	0.9215	*ITSN1*	0.921	*CASK*	0.9205
STK24	0.9215	*CCDC50*	0.921	*DLGAP2*	0.9205
DIS3	0.9215	*MED28*	0.9209	*LCOR*	0.9204
SOGA1	0.9215	*FBXL20*	0.9209	*H6PD*	0.9204
TAOK1	0.9215	*OTUD7A*	0.9209	*HOMER2*	0.9203
CLMN	0.9215	*FBXO25*	0.9209	*TMEM192*	0.9203
SH3PXD2A	0.9215	*GRIN2A*	0.9209	*CIITA*	0.9203
MAP3K2	0.9215	*RGS17*	0.9209	*SSH1*	0.9203
LDLRAD4	0.9215	*MAPK1*	0.9209	*SV2C*	0.9203
FAM179A	0.9215	*C1orf21*	0.9209	*SKP1*	0.9203
TMOD3	0.9214	*GABRA4*	0.9209	*PHC3*	0.9203
CADM1	0.9214	*FOXK1*	0.9209	*EIF2AK2*	0.9203
PGBD5	0.9214	*ERBB4*	0.9208	*N4BP2*	0.9203
DENND1B	0.9214	*TTC39B*	0.9208	*PPARA*	0.9203
FZD3	0.9212	*ZNF8*	0.9208	*EIF4E*	0.9203
PDK1	0.9212	*KIAA1244*	0.9208	*AFF2*	0.9203
DBT	0.9212	*VANGL1*	0.9208	*ST8SIA3*	0.9203
TMEM106B	0.9212	*ESPL1*	0.9207	*ZFYVE26*	0.9203
VPS53	0.9212	*GPRIN3*	0.9207	*MKLN1*	0.9203
SLC1A2	0.9212	*STXBP4*	0.9207	*LANCL3*	0.9203
TMEM154	0.9212	*ADAM10*	0.9206	*NUDT3*	0.9203
RNF217	0.9212	*ZNF264*	0.9206	*ZNF765*	0.9202
AJAP1	0.9212	*EIF4E3*	0.9206	*CALN1*	0.9202
SCO1	0.9211	*IKZF3*	0.9206	*HMGA2*	0.9202

DCP2	0.9202	*LPHN3*	0.9196	*LIPG*	0.9189
EPT1	0.9201	*SSR1*	0.9196	*USP6NL*	0.9189
CACUL1	0.92	*SNX1*	0.9196	*CDH7*	0.9189
PPIP5K2	0.92	*CCDC127*	0.9196	*ALG14*	0.9189
ITGB8	0.9199	*FIGN*	0.9196	*PCDHA4*	0.9189
SLC4A8	0.9199	*TET2*	0.9195	*CEP250*	0.9189
LRRC27	0.9199	*CD34*	0.9195	*NUFIP2*	0.9189
LYST	0.9199	*TSC22D2*	0.9195	*PANK3*	0.9189
MON2	0.9199	*ARHGAP26*	0.9195	*TFCP2L1*	0.9189
PTPLAD2	0.9199	*AGPAT4*	0.9195	*SLC24A2*	0.9189
XYLT1	0.9199	*DSC2*	0.9194	*MPRIP*	0.9189
RAB30	0.9199	*NDUFS1*	0.9194	*SV2B*	0.9189
PGR	0.9199	*GABRG3*	0.9194	*IRGQ*	0.9189
POLR1A	0.9199	*PDZD8*	0.9194	*GNB5*	0.9188
CLOCK	0.9199	*PRDM11*	0.9194	*JRK*	0.9188
PTAR1	0.9199	*GCC2*	0.9194	*OTULIN*	0.9188
DDI2	0.9199	*GJC1*	0.9194	*SSBP2*	0.9188
IPO9	0.9199	*USP49*	0.9194	*HELZ*	0.9188
MTO1	0.9199	*KCNB1*	0.9194	*CREB5*	0.9188
HSPA4L	0.9198	*KRR1*	0.9193	*UBE2W*	0.9188
PIGP	0.9198	*EXOC5*	0.9193	*MPLKIP*	0.9187
ZDHHC21	0.9198	*DR1*	0.9193	*LLPH*	0.9187
ABHD2	0.9198	*SARM1*	0.9193	*PAPD5*	0.9187
GABPB2	0.9198	*ZNF740*	0.9193	*SLFN5*	0.9187
CD226	0.9197	*ANKRD11*	0.9193	*ZEB1*	0.9187
SCAI	0.9197	*PYGO1*	0.9193	*NOVA1*	0.9187
METTL8	0.9197	*CLCN5*	0.9193	*GTF2H5*	0.9187
PTPN4	0.9197	*BCL11B*	0.9193	*PDPR*	0.9186
CUX1	0.9197	*FAM26E*	0.9193	*GXYLT1*	0.9186
SHPRH	0.9197	*KIAA1456*	0.9193	*RIMS3*	0.9186
TET3	0.9197	*MAP3K13*	0.9192	*ZNF778*	0.9185
DYRK2	0.9197	*GFOD1*	0.9192	*FBXO22*	0.9185
SNX30	0.9197	*TRIM71*	0.9191	*TMEM132B*	0.9185
ATP5S	0.9196	*KCMF1*	0.9191	*PTBP2*	0.9185
C4orf32	0.9196	*FZD4*	0.9191	*PAG1*	0.9185
FMN1	0.9196	*LRPAP1*	0.919	*QKI*	0.9185
BMPR1A	0.9196	*MAS1*	0.9189	*SESTD1*	0.9185

CELF2	0.9185	*AGPAT6*	0.918	*SAMD8*	0.9174
POTEC	0.9185	*DBNL*	0.918	*PLEKHA8*	0.9174
AMER2	0.9185	*XPO4*	0.9179	*ZADH2*	0.9174
RIMKLA	0.9185	*FRRS1L*	0.9179	*MTMR9*	0.9173
NUDCD3	0.9184	*KIF6*	0.9179	*DRAXIN*	0.9173
IKZF2	0.9184	*PRKCA*	0.9179	*DGKI*	0.9172
SLC9A7	0.9184	*NTRK2*	0.9179	*CHFR*	0.9172
RAB22A	0.9184	*WHSC1*	0.9179	*SEC22C*	0.9172
ATXN1	0.9184	*TRIM66*	0.9179	*MECP2*	0.9172
BTBD7	0.9184	*ANKRD52*	0.9179	*MIER3*	0.9172
MR1	0.9184	*ZNF148*	0.9179	*ATF7*	0.9172
RPL37	0.9184	*GMPS*	0.9178	*LRIG2*	0.9171
FAM63B	0.9184	*CLVS2*	0.9178	*LNPEP*	0.9171
CENPP	0.9184	*SCN8A*	0.9178	*ADCY1*	0.9171
MBD5	0.9184	*NAA30*	0.9178	*PTPRT*	0.9171
FEM1A	0.9183	*FER*	0.9178	*SH3BP2*	0.9171
CREB3L2	0.9183	*AKAP6*	0.9177	*CTNNA3*	0.917
KIAA1715	0.9183	*NCKAP1L*	0.9177	*ARPIN*	0.917
DISC1	0.9182	*NTPCR*	0.9177	*AR*	0.917
ZNF37A	0.9182	*SLC24A4*	0.9177	*ATXN7L3B*	0.917
FAT3	0.9182	*HSBP1*	0.9177	*NKTR*	0.917
CAPRIN2	0.9182	*RNF115*	0.9177	*KLF7*	0.917
GATAD2B	0.9182	*CBL*	0.9177	*SBNO1*	0.917
CMC2	0.9181	*HRK*	0.9177	*IBA57*	0.917
PRLR	0.9181	*RNF152*	0.9177	*GPR180*	0.917
PAPPA	0.9181	*TBL1XR1*	0.9177	*AGFG1*	0.9169
TRDMT1	0.9181	*DNASE1*	0.9176	*HS6ST3*	0.9169
BMPR2	0.9181	*RBMS2*	0.9176	*RFX7*	0.9169
VGLL3	0.9181	*FKTN*	0.9176	*SOCS7*	0.9169
PSD4	0.918	*KPNA4*	0.9176	*UGGT1*	0.9169
N4BP2L2	0.918	*POLE*	0.9176	*SIK2*	0.9169
WDR3	0.918	*FOXP2*	0.9176	*ADGB*	0.9169
TMEM200C	0.918	*QSOX1*	0.9176	*BTBD9*	0.9169
RORB	0.918	*TRIM67*	0.9176	*ZKSCAN1*	0.9169
ZMAT3	0.918	*UNC13C*	0.9175	*SMC1A*	0.9169
RBM33	0.918	*PEX5L*	0.9175	*GMFB*	0.9169
NUDCD2	0.918	*NFIB*	0.9175	*DIEXF*	0.9169

MAPK13	0.9168	*KDM3A*	0.9163	*XPR1*	0.9159
PDE11A	0.9168	*SOX6*	0.9163	*OTUD3*	0.9159
FAM168A	0.9168	*ICOSLG*	0.9163	*SF3B3*	0.9158
RAB11FIP1	0.9168	*FREM2*	0.9162	*SMIM14*	0.9158
PPM1A	0.9168	*NR2C2*	0.9162	*TSPAN3*	0.9158
WNK3	0.9168	*ERCC6*	0.9162	*EPM2AIP1*	0.9158
ASAP1	0.9168	*PGM2L1*	0.9162	*APC*	0.9158
RALGPS2	0.9167	*PRKCB*	0.9162	*SMAD4*	0.9158
EXT1	0.9167	*MAN1A2*	0.9162	*OTUD7B*	0.9157
PTPRB	0.9166	*KIAA1644*	0.9162	*ZKSCAN8*	0.9157
MKL2	0.9166	*CFL2*	0.9161	*MLYCD*	0.9156
ZNF142	0.9166	*ZNF774*	0.9161	*KIF1B*	0.9156
AP5M1	0.9165	*ATRX*	0.9161	*ACOT11*	0.9156
RNF150	0.9165	*MAPK1IP1L*	0.9161	*PSD3*	0.9155
DCX	0.9165	*RNF130*	0.9161	*AMER1*	0.9155
LARGE	0.9165	*ARHGAP32*	0.9161	*CYB5R4*	0.9154
APOL6	0.9165	*UBXN7*	0.9161	*RASSF5*	0.9154
LIMD1	0.9165	*THRB*	0.9161	*A1CF*	0.9153
KCNMA1	0.9165	*IGF2BP1*	0.9161	*YLPM1*	0.9153
DNAL1	0.9164	*NA*	0.9161	*NOX5*	0.9153
BRAF	0.9164	*CREB1*	0.9161	*VAPB*	0.9153
SOX11	0.9164	*KATNAL1*	0.9161	*UHMK1*	0.9153
DAGLA	0.9164	*TEAD1*	0.916	*C1orf95*	0.9153
RNF165	0.9164	*NAV1*	0.916	*OGFRL1*	0.9153
DCUN1D1	0.9164	*IL6ST*	0.916	*KIAA0930*	0.9152
TTBK2	0.9164	*PAX5*	0.916	*DNMT3A*	0.9152
NOS1	0.9164	*HDAC2*	0.916	*DCLK1*	0.9152
PRKAA2	0.9164	*SOGA3 KIAA0408*	0.916	*PAK3*	0.9152
CADM2	0.9164	*SYNJ2BP*	0.9159	*PIK3C3*	0.9152
NWD1	0.9164	*CA5A*	0.9159	*KIAA1549L*	0.9152
INPP4A	0.9163	*CACNA1C*	0.9159	*SHE*	0.9152
ADD2	0.9163	*NT5DC1*	0.9159	*EGFR*	0.9152
FAM126B	0.9163	*ADRBK2*	0.9159	*PPP1R9A*	0.9152
TRPM3	0.9163	*CDH8*	0.9159	*RRP15*	0.9152
ADAMTS4	0.9163	*RSF1*	0.9159	*WWC2*	0.9152
USP31	0.9163	*ZBTB34*	0.9159	*MGA*	0.9151
ZHX3	0.9163	*ZC3H6*	0.9159	*PROX1*	0.9151

C14orf166	0.9151	*GSTO2*	0.9144	*SPRY3*	0.9139
ZNF641	0.9151	*TTF2*	0.9144	*C15orf59*	0.9139
MLXIP	0.9151	*ACP6*	0.9144	*KCNQ3*	0.9139
RBFOX2	0.9151	*CYP46A1*	0.9144	*ST8SIA1*	0.9139
ADAMTS5	0.9151	*PDXK*	0.9144	*CACNG8*	0.9139
MGAT5	0.9151	*WHSC1L1*	0.9144	*DICER1*	0.9139
BRWD1	0.9151	*ADAMTS6*	0.9144	*SCUBE1*	0.9138
GRIK3	0.9151	*UNC5C*	0.9144	*GEN1*	0.9138
CRB1	0.9151	*CNOT6L*	0.9144	*RAB11FIP4*	0.9138
EMP1	0.915	*MYO18A*	0.9144	*PRDM6*	0.9138
SFXN2	0.915	*SNX29*	0.9144	*PVRL1*	0.9138
TRABD2B	0.915	*NGRN*	0.9143	*FNDC3B*	0.9138
KCND3	0.915	*PAIP2B*	0.9143	*PLCXD3*	0.9138
VKORC1L1	0.915	*CPEB4*	0.9143	*ATG9A*	0.9138
TAF8	0.915	*LRRK2*	0.9143	*XPNPEP3*	0.9138
SEC31B	0.915	*TNKS*	0.9143	*MTR*	0.9138
RPS6KA3	0.915	*GTF3C4*	0.9143	*KIF26B*	0.9138
ATXN1L	0.9149	*DCN*	0.9142	*DGKE*	0.9137
RNMT	0.9149	*TRPS1*	0.9142	*RYBP*	0.9137
TTLL7	0.9149	*C18orf32*	0.9142	*KIAA2022*	0.9137
ST3GAL1	0.9149	*DCAF7*	0.9142	*KCNK10*	0.9136
CLIC5	0.9148	*SHROOM4*	0.9142	*ZNF678*	0.9136
MBP	0.9147	*IDS*	0.9142	*BCAS4*	0.9136
HSD17B2	0.9146	*LRCH3*	0.9142	*MTAP*	0.9136
NKD1	0.9146	*SOX5*	0.9141	*CLN8*	0.9136
MYEF2	0.9146	*BCAP29*	0.9141	*RBFOX2*	0.9136
LSAMP	0.9146	*GRSF1*	0.9141	*ASB1*	0.9136
GFRA1	0.9146	*PLAG1*	0.9141	*TMEM33*	0.9136
PLXNA4	0.9146	*ARHGAP19*	0.9141	*TMTC1*	0.9136
C12orf49	0.9145	*BRCA1*	0.9141	*TTC14*	0.9136
DYNLL2	0.9145	*ST6GAL2*	0.9141	*LRRC58*	0.9136
TMED8	0.9145	*TNR*	0.914	*AK4*	0.9136
STOX2	0.9145	*ROCK1*	0.914	*CLN8*	0.9136
PDE12	0.9145	*DNAJC5*	0.914	*EME2*	0.9136
TMEM237	0.9145	*GNL1*	0.914	*KCNC1*	0.9135
PTEN	0.9145	*GNL1*	0.914	*STRN*	0.9135
LPGAT1	0.9145	*GNL1*	0.914	*LUZP1*	0.9135

SSFA2	0.9135	*ATP8A2*	0.9126	*FUNDC2*	0.9123
PLXNA2	0.9135	*ITGA1*	0.9126	*FPGT-TNNI3K*	0.9123
SLC35B4	0.9135	*GPR26*	0.9126	*GFOD2*	0.9122
PI15	0.9134	*ZNF555*	0.9126	*ZFP14*	0.9122
MTMR3	0.9134	*NFASC*	0.9126	*ZNF623*	0.9122
ATL3	0.9134	*MIPOL1*	0.9126	*CCNT2*	0.9122
MTMR10	0.9134	*GULP1*	0.9126	*TTPAL*	0.9122
NEURL1B	0.9133	*SLCO5A1*	0.9125	*TSC1*	0.9122
NA	0.9133	*ZNF121*	0.9125	*GABRA2*	0.9122
SORT1	0.9132	*ELP2*	0.9125	*SLC36A1*	0.9121
ZNF831	0.9132	*NA*	0.9125	*RALGPS1*	0.9121
AP5S1	0.9131	*GOSR1*	0.9125	*RASGRF2*	0.9121
FNTA	0.9131	*KIDINS220*	0.9125	*ZBTB41*	0.9121
DCP1A	0.913	*NOVA2*	0.9125	*ZNF766*	0.912
LPCAT1	0.913	*BICD1*	0.9125	*UBXN2A*	0.912
LRP10	0.913	*RNF213*	0.9125	*INSR*	0.912
FTO	0.9129	*CNTNAP2*	0.9125	*PSMG4*	0.9119
PPM1F	0.9128	*TROVE2*	0.9125	*UBE2K*	0.9118
TNPO1	0.9128	*FSD1L*	0.9125	*NIN*	0.9117
NF1	0.9128	*MTF1*	0.9125	*NUDT4*	0.9117
TMEM170B	0.9128	*NUDT16*	0.9125	*TSHZ2*	0.9117
GDF11	0.9127	*FBXL4*	0.9124	*SLC7A14*	0.9116
DDR2	0.9127	*ATP8A1*	0.9124	*MED13L*	0.9116
CAMK1D	0.9127	*NR6A1*	0.9124	*ZNF326*	0.9116
OPA3	0.9127	*SPN*	0.9124	*KCNH5*	0.9116
IGF1R	0.9127	*SFMBT2*	0.9124	*FBXW2*	0.9116
CEP192	0.9127	*PEG10*	0.9124	*RAB3IP*	0.9116
DOK6	0.9127	*TBCK*	0.9123	*SLC7A11*	0.9116
ZNF445	0.9127	*DCUN1D3*	0.9123	*CEP78*	0.9116
CAND1	0.9127	*FRS2*	0.9123	*PGAP1*	0.9116
PIAS1	0.9126	*USP6*	0.9123	*GNE*	0.9116
ACVR1C	0.9126	*GSK3B*	0.9123	*SLC30A8*	0.9116
DAPK2	0.9126	*AFF4*	0.9123	*PLEKHG3*	0.9115
ITGA11	0.9126	*LMBRD2*	0.9123	*RPP14*	0.9115
NFIC	0.9126	*ELOVL6*	0.9123	*ARPC5*	0.9115
XKR7	0.9126	*ANTXR2*	0.9123	*BACE2*	0.9115
FOXN3	0.9126	*C22orf29*	0.9123	*FARP1*	0.9115

GDAP2	0.9115
FAM155A	0.9115
ZNF451	0.9115
MPP6	0.9114
ICE2	0.9114
THSD4	0.9114
IRAK3	0.9114
XIAP	0.9114
KIAA1549	0.9114
CINP	0.9113
HECW2	0.9113
VSTM4	0.9113
ASPH	0.9113
ZNF736	0.9113
MOB3B	0.9113
SMURF2	0.9113
SLC26A2	0.9113
RFX3	0.9113
RLIM	0.9113
KDM3B	0.9113
TCF4	0.9113

Supplementary 2: List of genes containing T-UCR in their exonic regions according to the study of Bejerano et al.

UCR number	**Length (bp)**	**Gene ID**
uc.13	237	*EIF2C1*
uc.28	355	*SFRS11*
uc.33	312	*PTBP2*
uc.45	203	*HNRPU*
uc.46	217	*HNRPU*
uc.48	298	*PUM2*
uc.49	207	*BC060860*
uc.50	222	*SFRS7*
uc.61	326	*BCL11A*
uc.77	296	*ZFHX1B*
uc.97	442	*HAT1*
uc.102	338	*PTD004*
uc.129	212	*MBNL1*
uc.135	201	*AK096400*
uc.138	419	*SFRS10*
uc.143	218	*AB014560*
uc.144	205	*HNRPDL*
uc.151	214	*ZFR*
uc.174	260	*MATR3*
uc.183	236	*FBXW1B*
uc.184	230	*CPEB4*
uc.185	411	*CLK4*
uc.186	305	*HNRPH1*
uc.189	573	*SFRS3*
uc.193	319	*SYNCRIP*
uc.194	201	*EPHA7*
uc.203	203	*AB067798*
uc.208	218	*TRA2A*
uc.233	266	*CENTG3*
uc.263	207	*HNRPK*
uc.280	220	*PBX3*
uc.282	207	*GRIN1*
uc.285	232	*CARP-1*
uc.292	217	*MLR2*
uc.313	231	*TIAL1*
uc.324	225	*C11orf8*
uc.330	207	*RBM14*
uc.331	218	*DLG2*
uc.333	270	*FLJ25530*
uc.338	223	*PCBP2*
uc.339	252	*ATP5G2*
uc.356	251	*MBNL2*
uc.375	300	*MIPOL1*
uc.376	290	*PRPF39*
uc.378	251	*NRXN3*
uc.393	275	*CLK3*
uc.395	249	*RBBP6*
uc.406	211	*NFAT5*
uc.409	244	*L32833*
uc.413	272	*BC060758*
uc.414	246	*THRA*
uc.419	289	*SFRS1*
uc.436	210	*TCF4*
uc.443	239	*HNRPM*
uc.454	208	*SLC23A1*
uc.455	245	*RNPC2*
uc.456	320	*SFRS6*
uc.471	239	*DDX3X*
uc.473	222	*NLGN3*
uc.474	210	*ZNF261*
uc.475	397	*OGT*
uc.477	209	*RAB9B*
uc.478	252	*GRIA3*