The role of vault RNA-derived small RNA in gene expression regulation



Jana Terenova

Sir William Dunn School of Pathology

University of Oxford

This dissertation is submitted for the degree of Doctor of Philosophy

Wadham College

September 2022

Preface

During the first six months of my PhD studies, I was a part of a collaboration project focused on investigating the RNA content of extracellular vesicles using RNA sequencing data isolated from *Drosophila* in Dr Baena-Lopez's research team in the Sir William Dunn School of Pathology. The obtained results were summarised in the dissertation chapter of Dr Claire Hill. In addition, for another six months, I worked on analysing three different RNA sequencing datasets generated by Ruth F. Ketley, in the laboratory of Dr Gullerova, for her PhD project. These data included investigation of short-hairpin RNA, small interfering RNA Knock Downs of TIRR and analysis of TIRR bound RNA. The results of these analyses are part of her dissertation and publication in preparation. In my PhD dissertation, I worked on the vault RNA-derived small RNA biogenesis and functional role (2 years and 5 months).

Impacts of COVID-19 pandemic

-During the national lockdown, I worked from home. Although I was able to perform computational analyses, my planned validation experiments were hindered.

-The department of Sir William Dunn School of Pathology re-opened on 8th June 2020 with strict measures, allowing a maximum of 2 members from each research group to be present in the lab at one time, later increasing to 50% capacity.

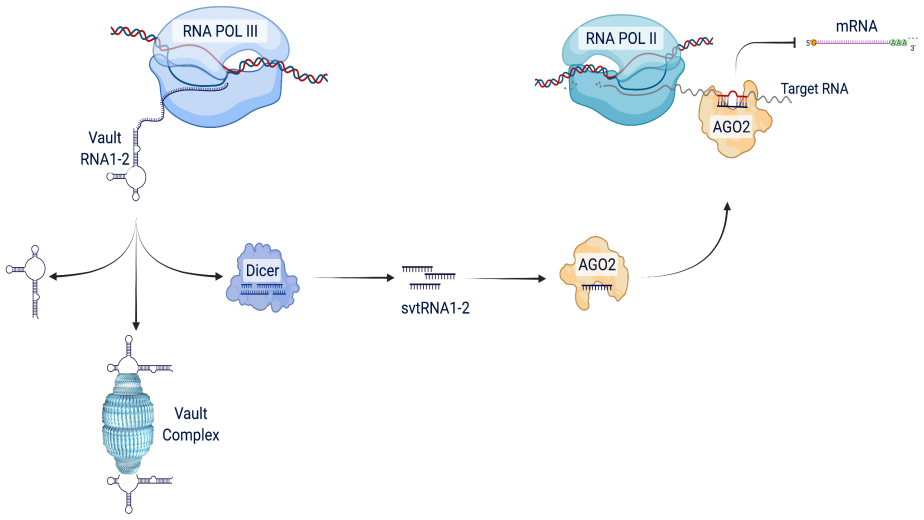
-The arrival of new students in early October 2020 forced the entire lab to work in shifts. This put me under a lot of stress, as I had only limited time to perform the key experiments and return to the computational work.

This was an unprecedented situation, which influenced my wet-lab experiments. I hope that the assessors will consider it while examining my dissertation.

Summary

Vault RNA are small non-coding polymerase III transcribed RNA, which are a crucial component of a large ribonucleoprotein complex called Vault. Despite only 5% of vault RNA being associated with the Vault complex, little is known about their Vault-independent cellular roles. Here I show that vtRNA are processed by Dicer into functional small vault RNA (svtRNA) fragments that associate with Argonaute 2. Nascent RNA sequencing revealed the highest number of predicted svtRNA1-2 target genes among up-regulated genes, in AGO2 and Dicer Knock Down cells, compared to other svtRNA. svtRNA1-2 target sites were detected predominantly in the intronic regions. Genomic deletion of vtRNA1-2 results in the severe inhibition of proliferation and up-regulation of genes significantly associated with cell membrane physiology, including ERBB4 and cell adhesion. Finally, I show that svtRNA1-2 hybridises to target genes in a dual-region complementarity and loop protrusion is required for intron targeting. These results show a novel small vault RNA1-2 with a role in gene expression regulation, possessing unique molecular features for targeting genes in introns leading to nascent RNA silencing.

Graphical Abstract



Acknowledgements

First and foremost, I would like to thank my supervisor, Monika Gullerova, for having given me the opportunity to carry out my PhD thesis work in her group and for feedback throughout this writing process. I am also grateful to Irina Chelysheva, who joined the supervision halfway through my project, for the support and advice she has provided and her feedback throughout this writing process. I would also like to acknowledge the committee assessors, Nicholas Proudfoot, Andre Furger and Daniel O'Connor for their helpful feedback during my DPhil Status transfer and confirmation. I would like to thank all the members of the Gullerova research group-both current and former for the helpful comments they have provided throughout my work. In particular, Adele Alagia and Ruth F. Ketley who contributed to the work reported in this thesis.

In addition, I would like to express my gratitude to the Wadham College students, associated researchers and staff. Being part of this welcoming community has been an honour. Also, thank you to the Dunn School community, which during the worst times (Covid-19 pandemic) showed tremendous support for their students and employees.

Importantly, I would like to show my appreciation to all my friends in Slovakia and Oxford who supported me during my PhD. I would like to thank my friends from Slovakia, Maria Mazurova and Dominik Hadzega, for all the last-minute hourly calls we shared about RNA and life. Special thank you to Patricia Zizkovicova, who was always there for me to listen and discuss everything from science to life and shared with me the bad moments as well as the happy ones; thank you very much for your support and friendship. Adriana Nemcekova, thank you for being an amazing friend and my ultimate supporter. My Oxford friend Alistair Boyd thank you for being my best college buddy. I am incredibly grateful to the Oxford University Fencing Club (OUFC) members for all the Oxford memories and experiences and for introducing me to an amazing sport that helped to maintain my sanity during my PhD. And the biggest thank you to Florent Yvon for being by my side during these years and supporting me especially during the challenging times.

Finally, I would like to express my huge gratitude to my family, who has supported me through my studies, countless internships and encouraged me to always want more. For all these, I am very grateful to them.

Table of Contents

1 Introduction **1**

1.1 Vault Complex . . . 1

1.1.1 Characterisation and Identification . . . 1

1.1.2 Structural characterisation . . . 2

1.1.3 Functional relevance . . . 4

1.2 Vault RNA . . . 7

1.2.1 Functional relevance . . . 10

1.2.2 Intact vault RNA . . . 11

1.2.3 Small vault RNA fragments . . . 13

1.2.4 The role of vtRNA modifications . . . 14

1.3 Transfer RNA . . . 16

1.3.1 Characterisation and Identification . . . 16

1.3.2 tRNA modifications . . . 18

1.3.3 tRNA-derived fragments . . . 20

1.3.4 tRNA-derived fragments with functional relevance . . . 23

1.4 Post-Transcriptional Gene Silencing, Transcriptional Gene Silencing and Nascent RNA Silencing . . . 25

1.4.1 Post-Transcriptional Gene Silencing Characterisation and Identification . . . 25

1.4.2 Canonical miRNA biogenesis pathway . . . 25

1.4.3 Regulation of Dicer processing . . . 27

1.4.4 RISC complex and miRNA loading . . . 28

1.4.5 RISC complex and miRNA interactions . . . 29

1.4.6 Template target area . . . 30

1.4.7 Non-canonical Biogenesis Pathways . . . 31

1.4.8 Transcriptional Gene Silencing Characterisation and Identification . . . 32

1.4.9 Nascent RNA Silencing Characterisation and Identification . . . 33

2 The biogenesis of vault RNA-derived small RNA fragments **34**

2.1 Introduction . . . 34
2.2 Results . . . 35
2.2.1 Drosha-independent small vault RNA biogenesis . . . 35
2.2.2 Dicer-dependent small vault RNA biogenesis . . . 39
2.2.3 svtRNA1-2 associates with Argonaute proteins . . . 47
2.2.4 svtRNA1-2 targets genes on nascent RNA level . . . 49

3 The functional role of small vault RNA1-2 53
3.1 Introduction . . . 53
3.2 Results . . . 54
3.2.1 vtRNA1-2 Knock Out leads to defects in cellular proliferation . . . 54
3.2.2 Deletion of vtRNA1-2 alters the expression of genes associated with cellular membrane function . . . 57
3.2.3 svtRNA1-2 target genes are associated with cellular membrane function . . . 61
3.2.4 svtRNA1-2 targets introns through dual-region complementarity and loop protrusion . . . 64
3.2.5 Synthetic svtRNA1-2 downregulates ERBB4 expression in breast cancer cells . . . 67
3.2.6 Steady-state transcriptome profiling of vtRNA1-2 KO cells . . . 69

4 Discussion 73
4.1 The biogenesis of vault RNA-derived small RNA fragments . . . 73
4.2 The functional role of small vault RNA1-2 fragment . . . 77

5 Conclusion 81

6 Methods 83
6.1 The biogenesis of vault RNA-derived small RNA fragments . . . 83
6.1.1 Creating small non-coding RNA annotation file . . . 83
6.1.2 Human and Mouse reference genome . . . 83

6.1.3 Identification of Drosha involvement in vault RNA processing . . . 83
6.1.4 Drosha Formaldehyde Crosslinking, Immunoprecipitation, and sequencing (fCLIP-seq) . . . 85
6.1.5 Identification of Dicer involvement in vault RNA processing . . . 86
6.1.6 Dicer PAR-CLIP Sequencing data analysis . . . 87
6.1.7 vtRNA sequence alignment . . . 87
6.1.8 PhastCons Conservation Analysis . . . 87
6.1.9 vtRNA structures prediction and visualisation . . . 88
6.1.10 Identification of small vault RNA in mouse tissue samples . . . 90
6.1.11 AGO1-4 PAR-CLIP Sequencing data analysis . . . 91
6.1.12 AGO2/3 RIP-Seq data analysis . . . 92
6.1.13 Sequencing data analysis of RNA bound to AGO2 in cytoplasm and nucleus . . . 92
6.1.14 Confocal and bright light microscopy . . . 93
6.1.15 Data analysis of chromatin-associated RNA sequencing in WT, Dicer and AGO2 KD conditions . . . 93
6.1.16 chromatin-associated RNA sequencing data pre-processing . . . 94
6.1.17 GO . . . 95
6.1.18 Canonical transcripts annotation file . . . 95
6.1.19 Gene feature annotation file . . . 96
6.1.20 Reference sequences for svtRNA target prediction . . . 96
6.1.21 Small vault RNA target prediction . . . 97
6.1.22 miRanda output processing . . . 97
6.2 The functional role of small vault RNA1-2 . . . 98
6.2.1 Generating vtRNA1-2 KO cell line . . . 98
6.2.2 Cell culture conditions . . . 99
6.2.3 Proliferation and Colony formation Assays . . . 99
6.2.4 qRT-PCR validation of svtRNA1-2 predicted target - ERBB4 . . . 99
6.2.5 vtRNA1-2 KO and WT chromatin-associated RNA samples preparation . . . 100

6.2.6 Fractions validation using Western blot 101
6.2.7 ChrRNA Sequencing data analysis of WT and vtRNA1-2 KO samples 101
6.2.8 GO and GSEA 102
6.2.9 STRING analysis 103
6.2.10 Ingenuity Pathway Analysis (IPA) (Trial version) 103
6.2.11 svtRNA-target binding pattern investigation 104
6.2.12 Target validation of ERBB4 in MCF7 cells 104
6.2.13 Total RNA sequencing sample preparation 105
6.2.14 Total RNA sequencing data analysis 105

References **109**

7 Appendix **147**

7.1 Supplementary Figures 147
7.2 Supplementary Tables 150

List of Figures

1 Cryo-EM density for the high-resolution vault reconstruction 3
2 Genome browser view of vault RNA genes 7
3 Overview of the evolutionary conserved vtRNA genes in Mammalia 8
4 Vault RNA sequences alignment 9
5 Expression profiles of vtRNA in different cell lines 10
6 The states of vault RNA 11
7 Secondary structures of vtRNA1-1 and vtRNA1-2 with highlighted chemotherapeutic compounds binding 13
8 Schematic representation of vtRNA1-1 secondary structure and identified small vault RNA1-1 fragments 15
9 Schematic representation of differences and similarities among tRNA, miRNA and vtRNA 16
10 Illustration of pre-tRNA description and processing 17
11 Illustration of a mature tRNA and tRNA derived fragments 21
12 The Microprocessor complex and Dicer structures 26
13 Illustration of the human Argonaute 2 structure 29
14 Illustration of a guide RNA description 30
15 Drosha-independent vtRNA processing using small RNA-seq data 36
16 Drosha-independent vtRNA processing using Drosha fCLIP-seq data . . . 38
17 Small RNA-seq from HEK293T cells in WT and Dicer KD conditions . . . 39
18 Differential expression analysis of vtRNA in Dicer KD vs WT conditions . 41
19 vault RNA derived fragments are produced by Dicer 42
20 vtRNA sequences alignment and conservation analysis 44
21 Identification of murine svtRNA 45
22 svtRNA1-2 is associated with Argonate proteins in cytoplasm and nucleus 48
23 chrRNA-seq WT, Dicer and AGO2 differential expression analysis 50
24 Identification and annotation of predicted binding sites 51
25 Producing vtRNA1-2 KO using CRISPR-Cas12a strategy 55

26 chrRNA-seq in WT and vtRNA1-2 KO conditions data pre-processing statistics . . . 58
27 Differential expression analysis of chrRNA-seq in WT and vtRNA1-2 KO conditions . . . 60
28 Identification and annotation of svtRNA predicted targets from the chrRNA-seq vtRNA1-2 KO up-regulated gene . . . 61
29 Annotation of svtRNA1-2 predicted targets . . . 63
30 Analysis of nascent RNA target-svtRNA binding complementarity . . . 65
31 Validation of svtRNA1-2 regulation of ERBB4 in MCF7 cells . . . 68
32 Total RNA-seq data pre-processing . . . 70
33 Total RNA-seq differential expression analysis . . . 72
34 Schematic representation of chrRNA-sequencing data analysis pipeline . . . 108
35 GC content and length distribution of human mature miRNA and svtRNA 147
36 vault RNA tertiary structures . . . 147
37 Dicer-vtRNA interaction sites . . . 148
38 *In silico* workflow overview of svtRNA1-2 target preditions from nascent RNA transcripts . . . 148
39 *In silico* workflow overview of svtRNA1-2 nascent RNA target preditions . 149

List of Tables

1 Analysed datasets 107
2 Top 20 Gene Ontology terms for 227 predicted svtRNA1-2 target genes . . 151
3 227 predicted svtRNA1-2 target genes 153
4 Top 10 Gene Ontology terms for Up-regulated genes (vtRNA1-2 KO) . . . 160
5 Top 10 Gene Ontology terms for Down-regulated genes (vtRNA1-2 KO) . . 162
6 29 predicted svtRNA1-2 gene targets 164

List of Boxes

1 svtRNA1-2 WT miRanda predicted bindings 165

List of Acronyms

4-SU	4-thiouridine
7SL	Signal recognition particle RNA
aaRS	aminoacyl-tRNA Synthetases
ADAR	Adenosine Deaminase RNA Specific
ADAT3	Adenosine Deaminase TRNA Specific 3
AGO	Argonaute
ALYREF	Aly/REF Export Factor
ANG	Angiogenin
ASL	Anticodon Stem-Loop
ATP	Adenosine TriPhoshate
CACNG7/8	Calcium Voltage-Gated Channel Auxiliary Subunit Gamma 7/8
CBL-3	Cord Blood Lymphocyte-derived ncRNA
CCL21	C-C Motif Chemokine Ligand 21
CDC6	Cell Division Cycle 6
CDS	Coding DNA Sequence
chrRNA-seq	chromatin-associated RNA sequencing
CPM	Counts Per Million
cryo-EM	cryogenic Electron Microscopy
CYP3A4	Cytochrome P450 3A4
DGCR8	DiGeorge syndrome Critical Region gene 8
DIS3L	DIS3 Like Exosome 3'-5' Exoribonuclease
DIS3L2	DIS3 Like 3'-5' Exoribonuclease 2
DMEM	Dulbecco's Modified Eagle Medium
dsRBD	double-stranded RNA-Binding Domain
DU145	Prostate cancer
DUF283	Domain of Unknown Function
EBV	Epstein-Barr Virus
EGF	Epidermal Growth Factor

List continued from previous page

EGFR	Epidermal Growth Factor Receptor
ERBB4	Receptor protein-tyrosine kinase ErbB-4
ERK1/2	Extracellular signal-Regulated Kinase
ES	Enrichment Score
EV	Extracellular Vesicle
FC	Fold Change
fCLIP-seq	formaldehyde CrossLinking ImmunoPrecipitation and sequencing
FDR	False Discovery Rate
GAPDH	Glyceraldehyde-3-Phosphate Dehydrogenase
GEO	Gene Expression Omnibus
GFF	General Feature Format
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
GTF	General Transcription Factor
GTP	Guanosine-5'-TriphosPhate
H3	Histone 3
H3K27	tri-methylation of lysine 27 on histone H3
H3K9	Histone H3 lysine 9
HCT116	human colon cancer
HEK	Human Embryonic Kidney
HEK293	Human Embryonic Kidney 293
HeLa	Henrietta Lacks (Cervical cancer cell line)
IGV	Integrated Genomics Viewer
INT	INTeraction domain
IPA	Ingenuity Pathway Analysis
JAG2	Jagged Canonical Notch Ligand 2
KD	Knock Down
KEGG	Kyoto Encyclopedia of Genes and Genomes

List continued from previous page

KO	Knock Out
La	Lupus autoantigen protein
LIN28	Lin-28 Homolog
lncRNA	long non-coding RNA
LRP	Lung Resistance Protein
m5C	5-methylcytosine
m6A	6-methylAdenosine
m7C	7-methylcytosine
MALAT1	Metastasis-Associated Lung Adenocarcinoma Transcript 1
MAPK	Mitogen-Activated Protein Kinase
MCF-7	Michigan Cancer Foundation-7 (Breast cancer cell line)
MDR	Multidrug Resistance
METLL16	METhylTransferase-Like 16
MFE	Minimum Free Energy
miRISC	miRNA-Induced Silencing Complex
miRNA	microRNA
miscRNA	miscellaneous RNA
MRE	miRNA Response Elements
mRNA	messenger RNA
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
MVP	Major Vault Protein
nc886-3p	non-coding 886-3p
ncRNA	non-coding RNA
NES	Normalized Enrichment Score
NFX1	Nuclear Transcription Factor, X-Box Binding 1
NPC	Nuclear Pore Complex
NRS	Nascent RNA Silencing
NSUN2	NOP2/Sun RNA Methyltransferase 2

List continued from previous page

PAR-CLIP	PhotoActivatable-Ribonucleoside Crosslinking and ImmunoPrecipitation
PARP12	Poly(ADP-Ribose) Polymerase Family Member 12
PAZ	Piwi/Argonaute/Zwille domain
PB1 domain	Phox and Bem1 domain
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
piRNA	piwi-interacting RNA
PKR	Protein Kinase Receptor
Pol III	RNA Polymerase III
pri-RNA	primary-miRNA
PTGS	Post-Transcriptional Gene Silencing
PUS7	Pseudouridine Synthase 7
qRT-PCR	quantitative Real-Time Polymerase Chain Reaction
RanGTP	Ras-related nuclear Guanosine-5'-TriphosPhate
RBP	RNA Binding Protein
RIIID	RNase III domain
RIP-Seq	RNA-ImmunoPrecipitation Sequencing
RISC	RNA-Induced Silencing Complex
RNA-seq	RNA sequencing
RNAi	RNA interference
RNP	RiboNucleoprotein Particles
RPM	Reads Per Million
rRNA	ribosomal RNA
RT	Reverse Transcription
scRNA	small conditional RNA
scaRNA	small cajal body-specific RNA
sgRNA	single guide RNA
SH-SY5Y	SK-N-SH neuroblastoma cell line

List continued from previous page

SHAPE	Selective 2'-Hydroxyl Acylation analyzed by Primer Extension
shRNA	short hairpin RNA
SINE	Short Interspersed Nuclear Element
siRNA	small interfering RNA
sncRNA	small non-coding RNA
snoRNA	small nucleolar RNA
SPINT1	Serine Peptidase Inhibitor, Kunitz Type 1
sRNA-seq	small RNA sequencing
SRSF2	serine/arginine SR-rich Splicing Factor 2
svtRNA	small vaultRNA
TALEN	Transcription Activator-Like Effector Nucleases
TCGA	The Cancer Genome Atlas
TEP1	TElomerase-associated Protein 1
TGIRT-seq	Thermostable group II intron reverse transcriptase-sequencing
TGS	Transcriptional Gene Silencing
tiRNA	stress-induced tRNA fragment
TOG	Terminal Oligo Guanine
tRF	tRNA-derived fragment
tRH	tRNA Halves
TRMT1	TRNA Methyltransferase 1
tRNA	transfer RNA
TUT4/7	Terminal Uridylyl Transferase 4/7
UCSC	University of California Santa Cruz
UTR	UnTranslated Region
V-loop	Variable loop
Vaultrc5	Vault RNA component 5
vPARP	vault Poly Adenosine diphosphate-RibosePolymerase
vtRNA	vault RNA

List continued from previous page

WCL	Whole Cell Lysate
WT	Wild Type
XPO5	Exportin 5
XPOT	Exportin For TRNA
YBX1	Y-Box Binding Protein 1
ZMAT2	Zinc Finger Matrin-Type 2

1 Introduction

In this chapter, I will provide a general introduction to vault RNA and tRNA and describe different types of gene silencing mechanisms that utilise small RNA fragments to regulate gene expression. Vault RNA is the central topic of this dissertation. The vault RNA-derived fragments are compared to the tRNA-derived fragments based on their biogenesis and functional relevance. As recently identified in the Gullerova laboratory, specific tRNA-derived fragments are generated by a non-canonical microRNA pathway to regulate the expression of the nascent transcripts. The additional background information is relevant to the studies presented in each chapter.

1.1 Vault Complex

The vault particles were first isolated from cytoplasmic extracts of rat liver cells together with coated vesicles in 1986. Their size was, however, smaller, and the structure was different to most known coated vesicles. As their morphology resembled the vaulted ceilings of cathedrals, *Kedersha* and *Rome* named them 'Vaults' (Kedersha et al. 1986). Vaults are the largest known ribonucleoprotein particles (RNPs) in eukaryotes, with a mass of 13 MDa or three times the size of a ribosome (Zon et al. 2001). In the last three decades, numerous studies have investigated their molecular composition, morphology and functional relevance. Yet, many aspects remain unknown, and some of them are still debated. In the following sections, I will describe vaults in general and then focus on their structural characterisation and functional role.

1.1.1 Characterisation and Identification

After the first identification of vaults from the subpopulation of coated vesicles, *Kedersha* and *Rome* sought to determined the absence of membrane around vaults, which is found in the coated vesicles, thus differentiating them from coated vesicles (Kedersha et al. 1986). Further work by the same group demonstrated that vaults are highly con-

served among higher eukaryotes, including mammals, amphibians (*Rana catesbeiana* and *Xenopus laevis*), avians (*Gallus Gallus*), Chondrichthyes (*Torpedo marmorata*/Electric ray) (Herrmann et al. 1996), Echinoidea (*Echinometra viridis*/sea urchin) (Hamill et al. 1997) and lower eukaryotes including Dictyostelia (*Dictyostelium discoideum*/slime mold) (Kedersha et al. 1990). Surprisingly, vaults have not been found in the commonly used molecular research model species such as *Saccharomyces cerevisiae*, *Caenorhabditis elegans* and *Drosophila melanogaster*. Nevertheless, vaults appear to be strongly conserved and broadly distributed, suggesting that their function is essential to eukaryotic cells (Kedersha et al. 1990).

The cellular localisation of the majority of vaults is cytoplasmic (Kedersha et al. 1990), but a small fraction was found localised in the nucleus (Chugani et al. 1993; Slesina et al. 2005; Slesina et al. 2006). Interestingly, in the sea urchin, vaults were dynamically localised in different cell compartments based on the developmental stage (Hamill et al. 1997), suggesting various roles in the cell. Their quantification showed amounts ranging from 10^4 to 10^5 particles per cell (Kickhoefer et al. 1998) with tissue-specific expression levels (Kedersha et al. 1990; Herrmann et al. 1996). Despite their abundance, this ribonucleoprotein complex is still largely uncharacterised from a functional point of view.

1.1.2 Structural characterisation

The structural features of the vault nanoparticle were first observed using an electron microscope, revealing an oval-shaped knot-like morphology of the vaults isolated from rat liver cells (Kedersha et al. 1986). Shortly after, Kong et al. 1999 uncovered in more details a hollow barrel-like structure with invaginated waist (78% of the total particle volume) and bulging caps at the ends (11% for each cap) using cryoelectron microscopy (cryo-EM) (**Figure 1**). The mature human vault particle has a molecular weight of around 13 MDa with overall dimensions of 72.5 nm x 41 nm $\times$ 41 nm (the values vary among species) (Mikyas et al. 2004).

Their structure is dominated by the major vault protein (MVP, 104 kDa) (**Figure 1a**) (Kedersha et al. 1990), which makes up more than 70% of the whole particle mass (Kedersha et al. 1991). Interestingly, it was shown that MVP can self-assemble to form vault-like

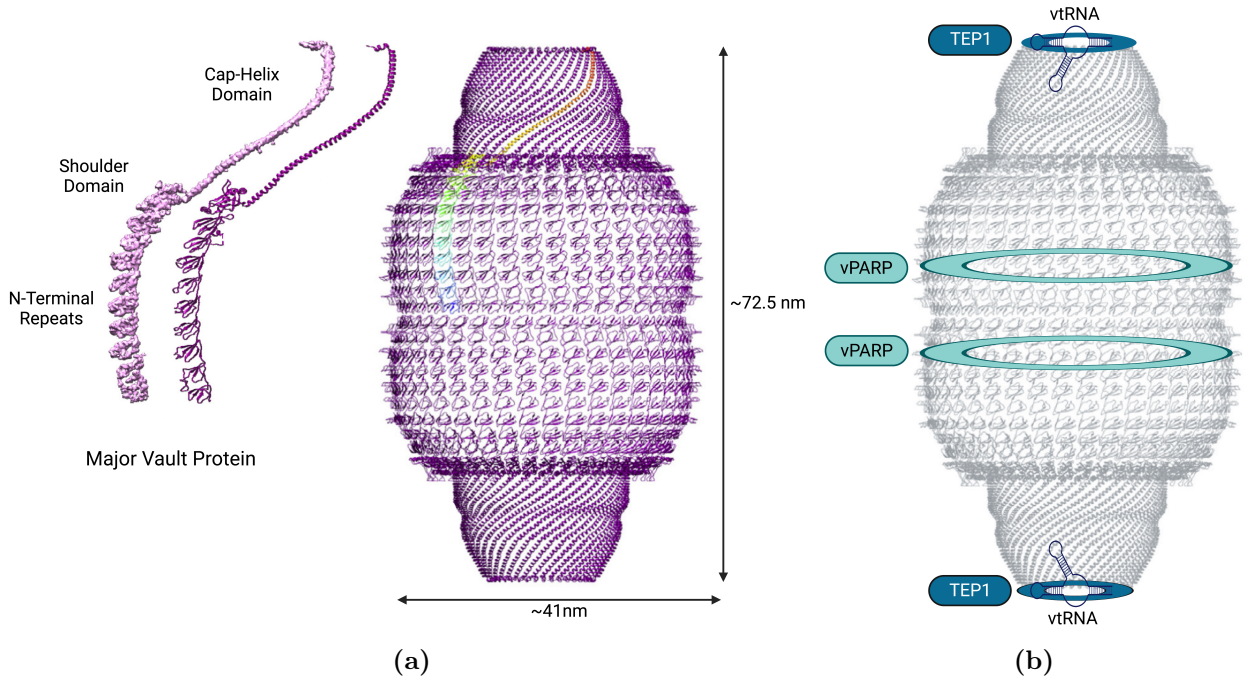


Figure 1: ***(a)*** *Side view of Vault (right), with major vault protein (MVP) molecule colored in purple (left).* ***(b)*** *Structural features of the vault nanoparticle including locations of the minor vault molecular components (Figure adapted and modified from Ding et al. 2018).*

particles (Stephen et al. 2001) consisting of two half-vaults containing 39 identical copies of MVP per half-vault. These halves can be dynamically exchanged and re-assembled into a new vault particle triggered by low pH. In this way, they can deliver the particle's content or uptake it from the location which may be relevant for their function (J. Yang et al. 2010). Vaults enclose approximately nine copies of vault poly(adenosine diphosphate-ribose)polymerase (vPARP, 139 kDa) (**Figure 1b**) (Poderycki et al. 2006), also referred to as PARP-4, which interacts with MVP via a 162-aa region at the C-terminus (amino acids 1563–1724) called interaction domain (INT) (Kickhoefer et al. 1999a). Depending on the physiological condition, vPARP can interact with MVP or be found detached in the nucleus, cytoplasm or in the proximity of the mitotic spindle during mitosis (Zon et al. 2003; Kickhoefer et al. 1999a). The vPARP protein has an enzymatically active domain which can ADP-ribosylate itself and MVP within the complex. In contrast to the regular PARP protein, vPARP's activity is not triggered by DNA damage (Kickhoefer et al. 1999a). Another described vault complex component is telomerase-associated protein 1

(TEP1, 290kDa) which associates with the vault caps (2-4 copies per vault) (Berger et al. 2009) and interacts with vault RNA (vtRNA) through the p80 homology region (Tetrahymena p80) (**Figure 1b**) (Poderycki et al. 2005), yet is not essential for the assembly of the vault multi-subunit structure (Kickhoefer et al. 1999b; Kickhoefer et al. 2001). As its name implies, telomerase-associated protein is also found as a telomerase subunit interacting with telomerase RNA, yet Y. Liu et al. 2000 showed using mTep1-deficient mice that its absence did not affect telomerase activity or telomere length, suggesting a different regulatory role. Interestingly, vtRNA and telomerase RNA are in direct competition for binding to TEP1, indicating the presence of only one RNA binding site (Poderycki et al. 2005). These vtRNA represent less than 5% of the vault mass, with approximately 16 RNA copies per vault particle (Kedersha et al. 1986) (**Figure 1b**). Equally, as TEP1 proteins, they do not influence the morphology of the ribonucleoprotein (Reviewed in Rome et al. 1991) but they are thought to have a rather functional role through RNA-Protein or RNA-RNA interactions. These results shed light on the molecular morphology and structural composition of vaults in cells.

1.1.3 Functional relevance

In the past years, the functional role of vault particles has been extensively studied, revealing a plethora of functions in normal and pathological processes, including nuclear-pore complex formation, nuclear-cytoplasmic transport, cellular signalling, innate immune response, apoptosis resistance, cell motility, DNA damage repair and development of multidrug resistance in cancer cells.

Despite the overwhelming efforts to characterise the functional relevance of the vault nanoparticles, their precise role is still unclear. Several studies investigated their role in Nuclear Pore Complex (NPC) and their potential in intracellular transport. Although the cytoplasmic localisation of vaults has been well-documented, further evidence shows their presence in the nuclear compartment in response to external stimuli, suggesting their involvement in nuclear-cytoplasmic shuttle (Slesina et al. 2005; Hamill et al. 1997). Furthermore, the physical properties of vaults, such as mass (13 MDa), shape (two halves, each with eightfold radial symmetry) and diameter (26-35 nm), were highly similar to the

central plug of the NPC and led to the hypothesis that vaults might interact with the NPC (Chugani et al. 1993). This hypothesis turned out to be incorrect. Instead, the amino acid sequence alignment of MVP and lung resistance protein (LRP) (Scheffer et al. 1995), which was found specifically over-expressed in multi-drug resistance human lung cancer (Scheper et al. 1993), revealed their same identity. This result suggested a new role for vaults in a multi-drug resistance (MDR) malignant tumours. In line with these findings, Noh et al. 2022 showed that the overexpression of MVP/LRP in glioblastoma caused a decrease in their drug sensitivity leading to less efficient target treatment, thus further supporting the role of vaults in the drug resistance.

It has been found that, MVP can also indirectly influence drug efficiency in colon cancer by transporting tumour suppressor miR-193a from tumour cells to exosomes, reducing its accumulation in cells, which was necessary for the inhibition of tumour cell growth (Teng et al. 2017). In addition, the depletion study demonstrated that MVP can inhibit ERK1/2 which promotes apoptosis in breast cancer cells (J. W. Wei et al. 2017). These data highlight diverse roles of vaults in the regulation of biological processes.

Furthermore, increased expression of MVP has been reported for several tumour tissues such as pancreas (Ramberg et al. 2021), breast (Y. S. Xiao et al. 2019), ovarian (Zhao et al. 2016) and cervical (Lloret et al. 2008) compared to the adjacent normal tissue. The elevated MVP levels can be further used as a cancer prognostic marker due to their association with poor patient prognosis. From a translational perspective, vaults present an interesting target for developing novel treatment approaches. Vault complex has gained scientific interest due to its potential in therapeutics as a drug delivery system, mainly due to its ability to encapsulate various molecules and specifically target cells. The spontaneous encapsulation of molecules by vaults is facilitated by the INT domain that was discovered on vPARP (Kickhoefer et al. 1999a). INT acts as 'zip code', which is recognised by vaults as a signal for loading and can be further used to encapsulate specific molecules of interest. Once the molecular cargo is present, vaults C-domain, located on the outer part of the particle, can be modified with an attached molecule recognised by cell-specific membrane receptors for example epidermal growth factor receptor (EGFR) (Kickhoefer et al. 2009). The cells can then uptake vaults via endocytosis (Galbiati et al. 2018) and receive

the functional molecules. Interestingly, the vaults nanoparticles are currently enrolled in pre-clinical trials in phase I for lung cancer treatment loaded with CCL21 chemokine, which acts as a chemoattractant for T lymphocytes, inhibiting cancer cell growth (Kar et al. 2011). The vault particles with customised properties and content show promising potential for manufacturing cell-specific nanocapsules for drug delivery.

1.2 Vault RNA

Vault RNA are small non-coding RNA of 80 to 150 nucleotides in length, depending on species (Stadler et al. 2009; Zon et al. 2001; Kickhoefer et al. 1993). Vault RNA are located on chromosome 5q31 within two different loci in the human genome (**Figure 2**). The VAULT-1 locus contains three genes coding for VTRNA1-1, VTRNA1-2 and VTRNA1-3 (previously known as HVG1/2/3), with sequences spacing about 7600 bp between VTRNA1-1 and VTRNA1-2 and 7200 bp between VTRNA1-2 and VTRNA1-3, which most probably arose through a tandem gene duplication (Zon et al. 2001). This locus is situated on the forward strand between the Zinc Finger Matrin-Type 2 (ZMAT2) and protocadherin-α genes responsible for synaptic adhesion molecules. The VAULT-2 locus encodes for VTRNA2-1 and is found on the reverse strand (Stadler et al. 2009) with no known sequence duplicates (Stadler et al. 2009). The RNA products of VTRNA2-1 gene transcription are still debated.

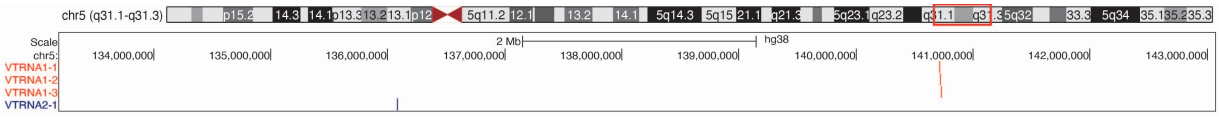


Figure 2: *Genome browser view of vault RNA genes on the chromosome 5. The VAULT-1 locus containing VTRNA1-1, VTRNA1-2 and VTRNA1-3 genes is located on plus strand (red) and VAULT-2 locus containing VTRNA2-1 gene is located on minus strand (blue).*

VTRNA2-1 was first annotated as precursor-miRNA (pre-miR-886), producing two microRNA size-like fragments, miR-886-5p and miR-886-3p (Landgraf et al. 2007). Later, it was called cord blood lymphocyte-derived ncRNA (CBL-3) because of its high expression in the CBL cell line after their infection with the Epstein-Barr virus (EBV). Further work by the same group noticed that CBL-3 share similar features with already identified vault RNA, such as expression pattern, sequence and structure, leading to the speculation that CBL-3 could be a novel fourth vault RNA (Mrazek et al. 2007). Only two years later, after naming the ncRNA CBL-3, Nandy et al. 2009 demonstrated that this non-coding RNA could associate with the vault complex and was renamed vtRNA2-1. The most recent study shows evidence of a full-length transcript called nc886 (101/102 nt in length) that can produce two short RNA nc886-3p, nc886-5p transcribed from the VTRNA2-1

gene (Y. S. Lee 2015). Nevertheless, in the Ensembl current version (107), VTRNA2-1 is listed as a gene coding for vault RNA2-1 with 128 nucleotides in length. Additionally, two vault RNA pseudogenes, VTRNA2-2P and VTRNA3-1P (previously known as HVG4) were identified on chromosomes 2 and X, respectively.

Vault RNA were found in a variety of organisms including, human, mouse, rat and bullfrog (**Figure 3**). Rats and mice have only one vtRNA gene. Vault RNA genes are transcribed by RNA Polymerase III (Pol III).

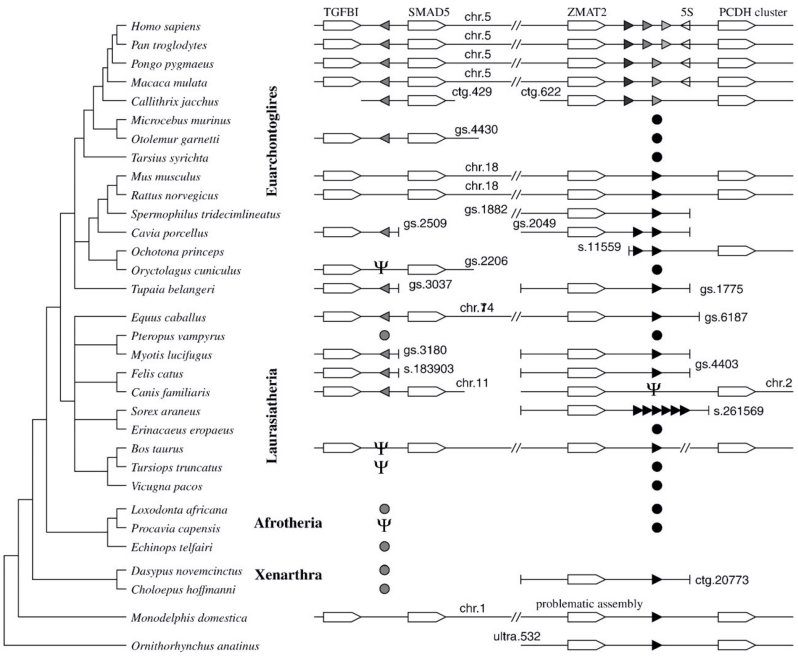


Figure 3: *Overview of the evolutionary conserved vtRNA genes in Mammalia. Black triangles indicate vtRNA and their transcription direction relative to the adjacent genes. Psi symbol represents vtRNA pseudogenes and white triangles represent 5S rRNA located adjacent to the vtRNA. Circles indicate vtRNA homologs with characteristic promoter element for one of the two loci. To construct the figure genomic data from ENSEMBL 50 and the NCBI Trace Archive (2008) were used.'Contig'(ctg.), 'Scaffold'(s.), 'GeneScaffold'(gs.), or 'UltraContig'(ultra.) (Figure adapted from Stadler et al. 2009).*

Pol III is responsible for transcribing 5S rRNA, tRNA, Y RNA and other small non-coding RNA (ncRNA). The transcription of these genes is regulated via three different promoter types. The rat vault RNA contains both internal (type 2) and external (type 3) promoter elements (single A box, two B boxes and 5'-flanking sequence), box A and B1 contributing to vtRNA expression, while the box B2 element shows expression inhibition

in the absence of the upstream sequences (Vilalta et al. 1994). The human vault RNA genes contain internal type 2 promoter motifs box A and box B (typically found within tRNA genes) (Vilalta et al. 1994).

Their sequence alignment clearly shows conserved regions at both ends, while the central region displays low homology and variable length (**Figure 4**).Moreover, all four vtRNA genes contain a functional stop sequence (TTTT) despite the fact that the VTRNA1-3 termination sequence has T-to-C mutation (TCTT) (**Figure 4**) (Zon et al. 2001).

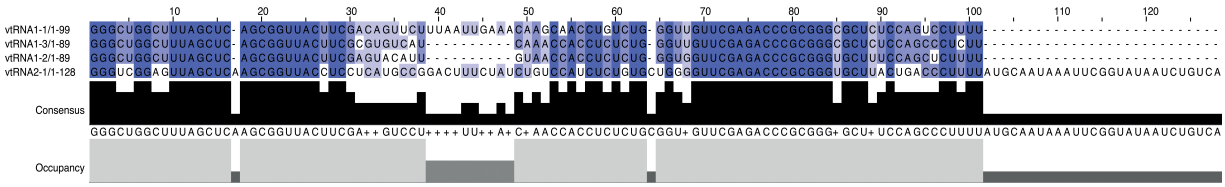
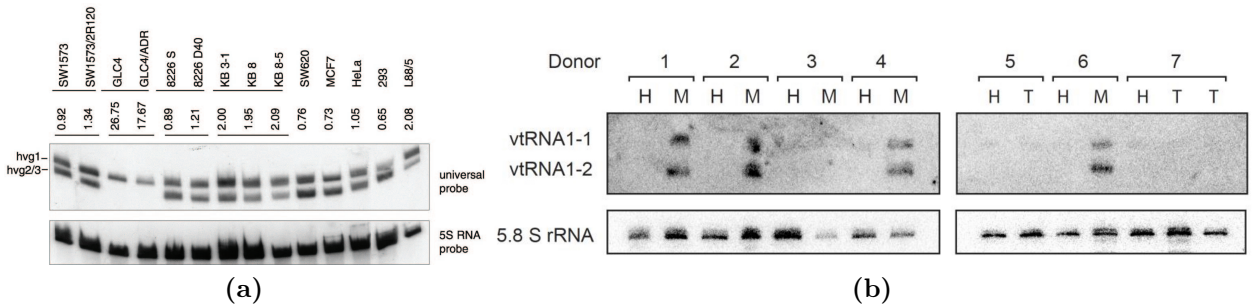


Figure 4: *Multiple sequence alignment of four human vault RNA (vtRNA1-1/2/3 and vtRNA2-1) sequences was done using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) aligner. The darker colour shows higher sequence consensus while lighter shows lower sequence consensus.*

Similarly to vaults, individual vtRNA are expressed in a tissue-specific manner. As large-scale vtRNA gene expression studies using RNA sequencing approaches were unavailable, numerous works characterised the expression of vtRNA using experimental methods from a subset of tissues and cell lines (F. Li et al. 2015; Zon et al. 2001; Stadler et al. 2009; Persson et al. 2009; Kickhoefer et al. 1998; Chen et al. 2018). For example, Zon et al. 2001 performed northern blots of expressed vault RNA in different cell lines showing that vtRNA1-1 displayed higher expression in the L88/5 (bone marrow) cell line than the other vault RNA. In contrast, the MCF7 (breast cancer) and HEK293 (human embryonic kidney) cell lines had higher levels of vtRNA1-2/3 (**Figure 5a**).

Furthermore, vtRNA1-1 and vtRNA1-2 showed high expression levels in metastatic but not in non-pathological adjacent liver patient samples, highlighting their potential roles in tumorigenesis (**Figure 5b**) (Gallo et al. 2022). A recent study by Fort et al. 2021 analysed genome-wide DNA methylation and chromatin accessibility data from The Cancer Genome Atlas (TCGA) to estimate the expression of vault RNA genes. They found that the expression of vtRNA2-1 and vtRNA1-2 was up-regulated in neoplastic



(a) **(b)**

Figure 5: ***(a)*** *Northern blots showing expression profiles of vtRNA in different cell lines. 15 µg of total RNA was blotted and then hybridized with the vtRNA-specific probes and 5S rRNA as loading control. The values above each lane indicate the hvg1:hvg2/3 ratio of universal probe. The hvg2 and hvg3 RNA migrate very close together in the gel and cannot be distinguished from each other (Figure adapted from Zon et al. 2001).* ***(b)*** *Northern blot showing expression profiles of vtRNA1-1 and vtRNA1-2 in paired patient-derived liver tissues. 5.8S rRNA as loading control. healthy (H), metastasis (M) and tumor (T) (Figure adapted from Gallo et al. 2022).*

tissues, indicating a broad oncogenic role for vtRNA1-2 and both tumour suppressor and oncogenic functions for vtRNA2-1. These studies highlight vault RNA similarities and differences, which provide necessary evidence toward a comprehensive understanding of their functional roles.

1.2.1 Functional relevance

vtRNA were identified as a part of the 13 MDa vault complex in 1986, after which they received their name (Kedersha et al. 1986). Since then, numerous independent works have focused on uncovering different biophysical characteristics, biological functions and their biogenesis. The primary states of vtRNA are unprocessed vtRNA bound to vaults, independent unprocessed vtRNA and vtRNA processed into smaller fragments (**Figure 6**).

As a part of the vault complex, vault RNA must first bind to TEP1 and then form a complex with MVP. Kickhoefer et al. 2002 found that the Lupus autoantigen (La) protein binds to vault RNA at the 3' end UUU-OH sequence in the nucleus and then to TEP1 in the cytoplasm. Interestingly, the La protein was reported as a nucleocytoplasmic shuttle (Bachmann et al. 1989; Fok et al. 2006) that could act as a vtRNA chaperone facilitating binding to TEP1. The formed complex of TEP1-vtRNA then targets the MVP

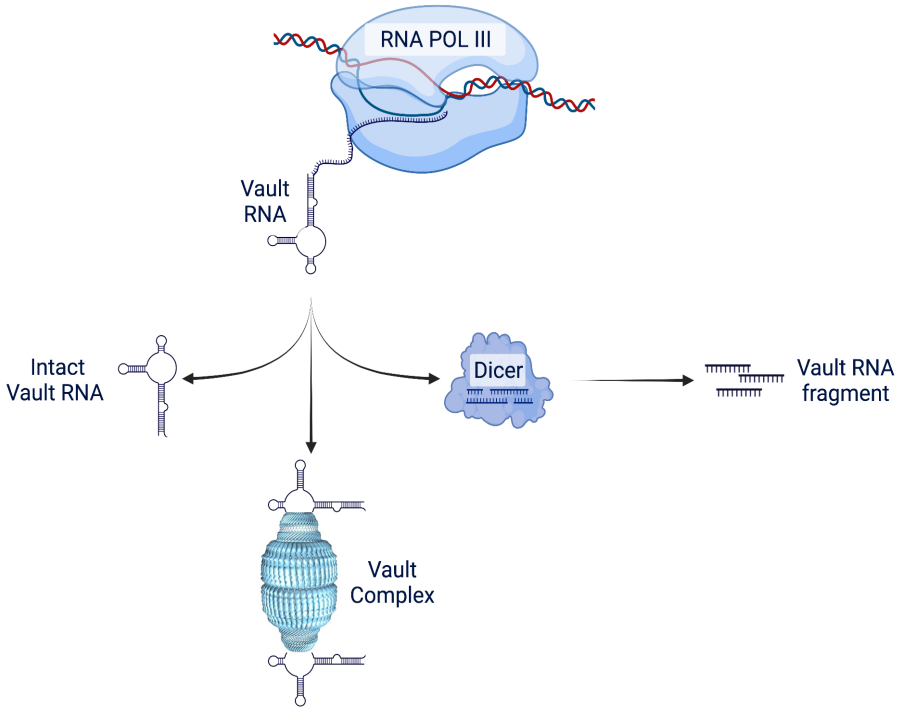


Figure 6: *The vault RNA forms.*

for the vault ribonucleoprotein assembly. It has been observed that, vtRNA1-3 isolated from immunoprecipitated vault complex showed increased levels in multidrug-resistant cell lines such as GLC4 and GLC4/ADR (adriamycin-resistant and doxorubicin-resistant human small cell lung). This effect was only observed for vtRNA1-3, suggesting specific vtRNA1-3/vaults complex involvement in drug resistance (Zon et al. 2001).

1.2.2 Intact vault RNA

Interestingly, it was found that only 5% of the vtRNA directly associates with the vault complex, whereas the remaining 95% were found distributed in the cytoplasm (Zon et al. 2001; Kickhoefer et al. 1998; Nandy et al. 2009). In recent years, several studies have investigated the independent unprocessed vtRNA (Reviewed in Hahne et al. 2021). J. Wang et al. 2020 showed that XPO5, which together with RanGTP mediates nuclear export of miRNA precursors into the cytoplasm, also facilitates vtRNA1-1 transport, yet is not the only vtRNA translocator. This has been reported as no change in the vtRNA1-1 levels nor the cytoplasmic localisation pattern was observed in the absence of

XPO5, leading to the speculation that other proteins might play a role in the cytoplasmic shuttle. In fact, the mass spectrometry data (LC-MS/MS) of proteins bound to vtRNA1-1 identified XPO1, NFX1 and Aly/REF (Sajini et al. 2019; Bracher et al. 2020), which play a role in RNA nucleoplasmic transport. Despite NXF1 and ALYREF preferentially binding to messenger RNA (mRNA), they have also been shown to bind to small non-coding RNA (Viphakone et al. 2019), further supporting alternative vtRNA cytoplasmic localisation. Furthermore, vtRNA can be exported from the cell through exosomes (Shurtleff et al. 2017), suggesting their potential role in inter-cellular communication.

A recent study by Horos et al. 2019 showed that vtRNA1-1 directly binds to the autophagy receptor sequestosome-1/p62 and blocks the oligomerisation required for the functional autophagy process (Reviewed in Buscher et al. 2020). Structural probing of vtRNA1-1 revealed precise nucleotides responsible for vtRNA1-1 binding to the PB1 domain and adjacent linker region of p62, which were located in the central region of the vtRNA1-1 sequence (Buscher et al. 2022). In contrast with these data demonstrating that the vtRNA1-1 overexpression favours autophagy-dependent apoptosis via p62, a study by Amort et al. 2015; Bracher et al. 2020; Ferro et al. 2022 provided findings of pro-survival, anti-apoptotic effects of the vault RNA expression with the regulatory region also located in the variable part of the vtRNA1-1 sequence.

Due to the fact that VTRNA2-1 was previously misannotated as pre-microRNA (pre-miR-886, miR-886-3p, miR-886-5p) and more recently called (nc886), it has been extensively studied. nc886/VTRNA2-1/pre-miR-886 was identified as a tumour suppressor mainly inferred by its expression pattern. It has been shown that CpG hyper-methylation of the VTRNA2-1 promoter caused its downregulation which was correlated with poor patient outcomes in small cell lung cancer, oesophagal squamous cell carcinoma and gastric cancer, highlighting its potential as a useful prognostic marker (Cao et al. 2013; H. S. Lee et al. 2014; K. S. Lee et al. 2014; Fort et al. 2018; K. Lee et al. 2011; M. C. Yu et al. 2020).

However, contrasting results show that nc886/VTRNA2-1/pre-miR-886 can act as an oncogene with increased expression inhibiting cellular processes in different cancers such as prostate cancer, thyroid cancer and renal cell carcinoma (Pan et al. 2022; E. K. Lee et al. 2016; Lei et al. 2017). These findings highlight that nc886/VTRNA2-1/miR-886 can

act as an oncogene and tumour suppressor gene in a tissue-specific manner.

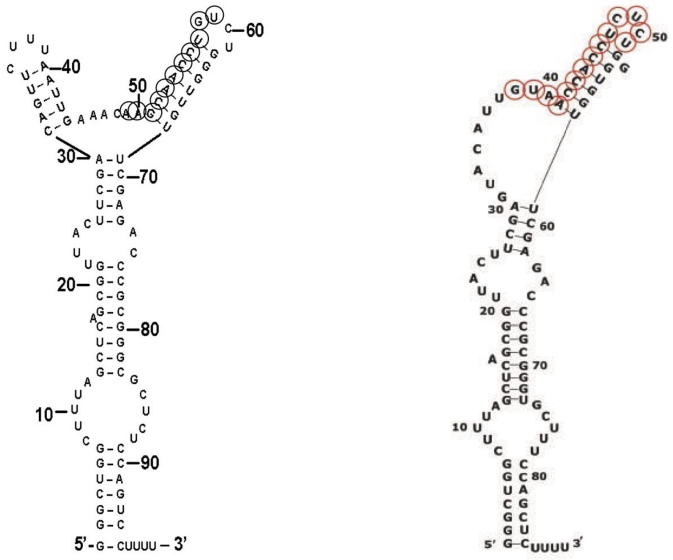


Figure 7: *Secondary structures of vtRNA1-1 and vtRNA1-2. Circled bases show binding regions of vtRNA with chemotherapeutic compounds (Figures adapted from Mashima et al. 2008; Gopinath et al. 2005).*

In addition to the functions attributed to vault RNA presented in the previous paragraphs, vtRNA1-1 (Gopinath et al. 2010) and vtRNA1-2 (Mashima et al. 2008; Gopinath et al. 2005) can also interact with chemotherapeutic compounds such as doxorubicin, mitoxantrone and etoposide to decrease their chemotherapy efficacy. Their interaction is facilitated by the central region of the vtRNA sequence and the specific chemical structure of these drugs (**Figure 7**). These observations led to interesting speculation that vault RNA might act together with vaults in the detoxification process; thus function in the multidrug resistance of cancer cells (Gopinath et al. 2005; Gopinath et al. 2010).

1.2.3 Small vault RNA fragments

Vault RNA can be processed into small-vault RNA (svtRNA) fragments to regulate gene expression by different mechanisms from those in the canonical miRNA pathway. The first identified svtRNA were derived from vtRNA1-1 (Persson et al. 2009). The expression of the highly abundant fragment of vtRNA1-1 from the 3' end is modulated by the methyl group on the cytosine 69 (Hussain et al. 2013). Its production was shown to be

dependent on Dicer but independent of the Drosha protein. Subsequently, the identified small fragments of vtRNA1-1, associated with the Argonaute (AGO) proteins, silence the expression of target genes (e.g. CYP3A4 and CACNG7/8) involved in drug metabolism and encoding voltage-gated calcium channels essential for cell differentiation via sequence complementarity (Hussain et al. 2013; Persson et al. 2009), respectively. Interestingly, Hasler et al. 2016 showed that the La protein acts as a gatekeeper, ensuring correct tRNA maturation and protection from Dicer processing to functional tRNA-derived fragments. Even though it was not investigated in their study, it is possible that vault RNA as a substrate of the La protein might undergo a similar process, leading to vtRNA processing into smaller functional fragments in the absence of the La protein.

Interestingly, vtRNA2-1 was equally found to be processed into svtRNA by Dicer independently of Drosha into two different svtRNA2-1 3' and 5' end fragments which could be loaded on AGO proteins (Minones-Moyano et al. 2013; Fort et al. 2020). These data suggest a new alternative miRNA-like pathway using svtRNA for post-transcriptional regulation of gene expression. Furthermore, fragments derived from the 3' end of vtRNA2-1 (previously known as miR-886-3p) and its precursor RNA (miR-886) were found to suppress important cellular regulators such as Protein Kinase RNA-activated (PKR) and Cell Division Cycle 6 (CDC6) in variety of cancer cell lines (K. Lee et al. 2011; Xiong et al. 2011).

1.2.4 The role of vtRNA modifications

Post-transcriptional modification of RNA at the base level can alter their metabolism. The 5-Methylcytosine (m^5C) modification was shown to drive vault RNA fragmentation into svtRNA (**Figure 8**) (Hussain et al. 2013). Sajini et al. 2019 further investigated the importance of (m^5C), discovering a novel vtRNA1-1 binding partner serine/arginine rich (SR) splicing factor 2 (SRSF2) in the absence of the m^5C modification. The binding of SRSF2 to the unmethylated vtRNA1-1 transcript prevented its processing into svtRNA, which in turn mediated the cell epidermal differentiation programme.

Furthermore, vault RNA can be modified by the TUT4/7 poly-U uridylyltransferase, by adding uridines to their 3' end, which are in turn recognised by the DIS3 Like 3'-5'

Exoribonuclease 2 (DIS3L2) to mediate their decay; thus monitoring the levels of vault RNA in the cell.

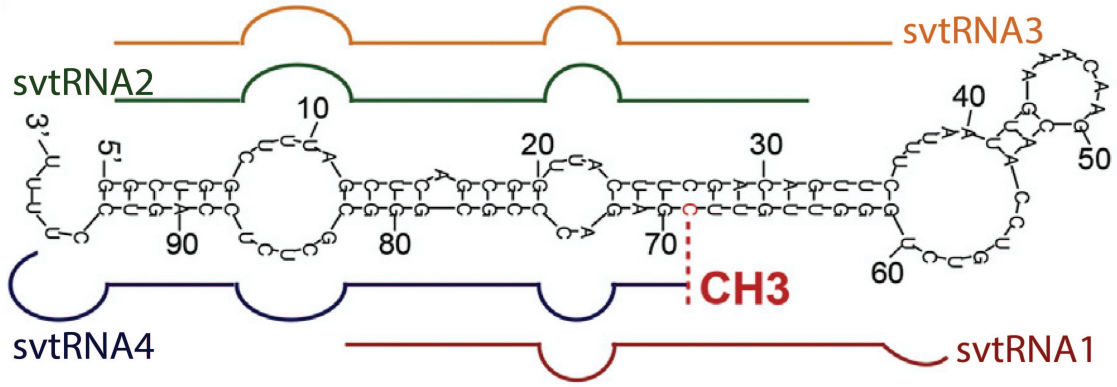


Figure 8: *Schematic representation of vtRNA1-1 secondary structure and identified small vault RNA1-1 fragments. CH3, cytosine-5 methylated site at position 69. (Figure adapted from Hussain et al. 2013).*

Additionally, several other modifications were described for vtRNA, such as 6-methyladenosine (m^6A) facilitated by methyltransferase-like 16 (METLL16) (Warda et al. 2017) and pseudouridine (ψ) introduced by PUS7 (Guzzi et al. 2018), which regulates the processing of a tRNA subset into tRNA fragments (tRF); however, their biological implications remain to be identified.

In conclusion, these findings shed light on the biochemical and structural properties of the vtRNA, revealing the importance of not only evolutionary conserved 3' and 5' ends but also to the less conserved central variable region of vault RNA. vtRNA have been attributed a variety of functions such as modulating multidrug resistance of cancer cells; protecting cells from undergoing apoptosis; regulating autophagy; modulating cancer cell growth by interaction with protein kinase R; substrate for production of small regulatory fragments, and a possible role in intercellular communication via extracellular vesicles.

However, these results do not point in a unique direction, suggesting individual roles for each vault RNA depending on tissues. It is possible to speculate that the common features of vault RNA and tRNA genes such as internal promoter elements, Pol III transcription, La protein binding, modification events, sequence length and processing into smaller fragments might help to elucidate the functional role of vault RNA (**Figure 9**). However, all the above-mentioned functions were demonstrated predominantly for vtRNA1-1 and

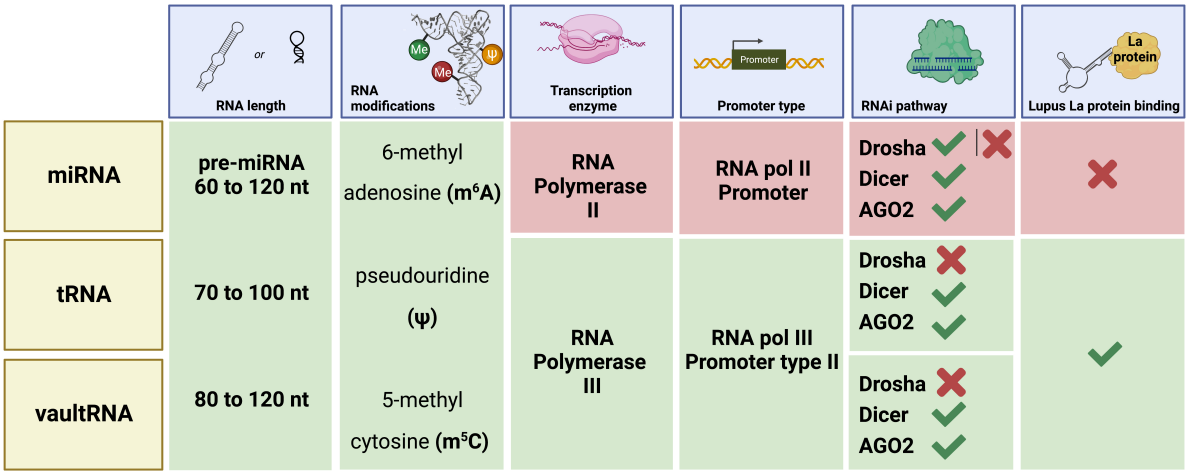


Figure 9: *Differences and similarities among tRNA, miRNA and vtRNA.*

vtRNA2-1, leaving vtRNA1-2 and vtRNA1-3 poorly characterised. Our understanding of vault RNA molecular functions remain limited, and further studies of this group of small non-coding RNA will be required to reach a comprehensive understanding.

1.3 Transfer RNA

The transfer RNA (tRNA) is a non-coding RNA which serves as a physical link between the mRNA and the protein amino acid sequence. They are complex molecules with a signature structure folding and extensive base modifications. In the following sections, I will describe their structure, function, which goes beyond translation regulation, and significant post-transcriptional changes and their influence on different aspects of tRNA.

1.3.1 Characterisation and Identification

Transfer RNA (tRNA) are small non-coding RNA of 73 to 93 nucleotides in length, responsible for translating genetic information into a protein sequence by transferring amino acids to ribosome machinery during translation. Their secondary structure folds into a characteristic, highly conserved cloverleaf shape with four stem-loops (**Figure 10**). In order from 5' to 3', the first is the amino acid acceptor stem with attached amino acid. Then the deoxyuridine stem-loop (D-arm) with signature dihydrouridine bases stabilises

the tRNA tertiary structure and interacts with aminoacyl-tRNA synthetases (aaRSs) for correct amino acid attachment. The opposite of the acceptor stem is the anticodon stem-loop (ASL) which mediates the recognition of complementary codons on mRNA during protein synthesis.

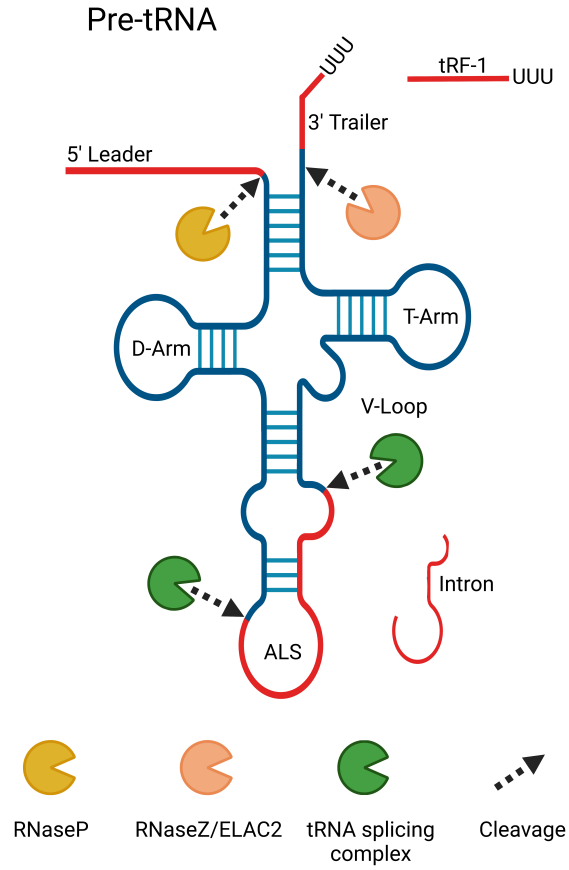


Figure 10: *Illustration of pre-tRNA. RNaseZ produces tRF-1 fragmnet from 3' trailer ending of tRNA. The tRNA intron is spliced out by tRNA splicing complex, and the ends are ligated together, creating Anticodon Stem-Loop (ALS). V-Loop represents the variable loop region of tRNA.*

Next is the variable loop (V-loop). The T(ψ)C stem-loop (T-arm) contains conserved modified bases thymidine, pseudouridine and cytidine and plays a role in tRNA binding to the ribosome during translation. The secondary cloverleaf structure folds into a tertiary L-shaped structure through the base pairing of nucleotides in the D-arm and T-arm. In comparison, mitochondrial tRNA are smaller (59-75 nt) and fold into irregular structures.

tRNA precursors are transcribed by Polymerase III, generating a variable length 3' end poly-U tail recognised by the lupus autoantigen (La), stabilising and protecting tRNA from

exonuclease digestion and misfolding (Hasler et al. 2016). On the other end, the 5' leader sequence is cleaved by RNase P, generating an acceptor stem with the signature seven base pairs (**Figure 10**). Afterwards, the 3' trailer sequence is trimmed by RNAse Z, leaving a 3'-OH group accessible for the CCA-adding enzyme (tRNA nucleotidyltransferase), which catalyses the addition of a CCA triplet. A small group of tRNA precursors contain short introns which are spliced out by the tRNA splicing complex. During the splicing, first the endonuclease releases the intron and then the ligase joins the flanking exons. After their processing in the nucleoplasm, tRNA are transported into the cytoplasm via the nuclear pore complex, which serves as the quality control to ensure that only correctly processed tRNA can pass into the cytoplasm. It is worth noting that eukaryotic tRNA can also shuttle back into the nucleus (Zasloff 1983, Reviewed in Chatterjee et al. 2018). Their transport is mediated by Los1 in yeast and Xpot in vertebrates (Arts et al. 1998; Hellmuth et al. 1998) in a Ran-GTP-dependent manner (Cook et al. 2009). In the cytoplasm, tRNA is subsequently modified with the amino acid specific aminoacyl-tRNA synthetases (aaRSs), into an amino acid-enzyme-tRNA complex. The 'charged tRNA' is released and prepared for delivery of the attached amino acids for the ribosomal incorporation into the growing polypeptide chain.

1.3.2 tRNA modifications

Post-transcriptional modifications are highly abundant on transfer RNA, some modifications are shared among tRNA, such as dihydrouridine (D-arm) or pseudouridine (T-arm), but some are specific for an individual tRNA or group of tRNA. The number of modifications on tRNA varies, but on average human tRNA contains between 11 to 13 modifications (Saikia et al. 2010). The base modification process can occur in the nucleus or cytoplasm, depending on the localisation of the modifying enzymes (Huh et al. 2003). The highest concentration of identified tRNA modifications was found in the anticodon stem-loop, notably highly altered bases at positions 32, 34, 37 and 38 with a significant impact on growth due to their role in translation. For example, inosine at position 34 (wobble position) expands decoding to A, U or C (Reviewed in Torres et al. 2014), profoundly impacting codon-anticodon recognition. Another modification at position 34 can restrict

decoding to prevent mistranslation by changing the base-pairing preference from G to A. An example of such modification is Lysidine (lysine-containing cytidine). This alteration converts the original codon from AUG to AUA and the amino acid from methionine to isoleucine (Muramatsu et al. 1988).

Other groups of modifications play a role in maintaining tRNA structures and stability. The secondary cloverleaf structure folds into a tertiary L-shaped structure through nucleotide base pairing in the D-arm and T-arm (Reviewed in Vare et al. 2017). Nobles et al. 2002 showed that modified T-arm bases at positions 54, 49 and 55 (U54 modified to ribothymidine, C49 modified to 5-methylcytidine or U55 modified to pseudouridine) increased interaction with D-arm compared to unmodified T-arm, thus promoting accurate tRNA folding into tertiary L-shaped structure. Several other modifications, such as 7-methylguanine and 5-methylcytidine and N4-acetylcytidine were demonstrated to prevent tRNA degradation through increased stability of the acceptor stem and T-arm (Whipple et al. 2011). Furthermore, pseudouridine has been shown to stabilise tRNA secondary structure through base-stacking interactions (Davis 1995; Durant et al. 1999; Yarian et al. 1999). Subsequent works further supported these observations, showing a positive correlation between the number of modifications and increased growth temperature (Watanabe et al. 1976; Agris et al. 1973; Kowalak et al. 1994; McCloskey et al. 2001). Taken together, these data show that modifications are an integral part of tRNA, contributing to dimensions of tRNA folding.

From a disease perspective, tRNA modifications are critical players in the regulation of important cellular processes. For example, modified positions 32 and 34 on tRNALeu, tRNAPhe and tRNATrp by methyltransferase FTSJ1 (Angelova et al. 2020; Feder et al. 2003) were linked to intellectual disabilities when mutated (Freude et al. 2004; Takano et al. 2008). The mutation influenced transcription, small RNA silencing and disease associations, which was further confirmed in mice and *drosophila* (Angelova et al. 2020; Jensen et al. 2019). Similarly, several other modifying enzymes, such as ADAT3 (Thomas et al. 2019) or TRMT1 (K. Zhang et al. 2020), were also found to play a role in intellectual disability. A different study demonstrated that the loss of (m^5C) enhanced the affinity between tRNA and angiogenin, leading to the tRNA processing into tRNA fragments

(tRF), which repressed protein synthesis and promoted squamous tumorigenesis (Blanco et al. 2014). These losses of modifications can be caused by modification 'eraser' enzymes such as ALKBH1 which causes a removal of N1-methyl group from adenine, whereas ALKBH1 reduction increased it. This methylation–demethylation modification cycle is particularly relevant during glucose deprivation when the expression of ALKBH1 is increased, leading to a global decrease in translation initiation (F. Liu et al. 2016).

Interestingly, the abundance of modifications can also influence their functionality. It has been shown that the absence of enzymes responsible for the nucleobases modification outside of the tRNA anticodon loop did not cause any abnormal phenotypes, likely because of the redundancy among the modifications that stabilise the tRNA structure (Reviewed in Phizicky et al. 2010). This observation was explored by Alexandrov et al. 2002, confirming that the removal of yeast TRM8 or TRM82, which methylate G46, did not affect the cell growth; however, the removal of enzymes modifying multiple bases such as TRM4, TRM10 (methylate cytosine or guanine) (Jackman et al. 2003; Motorin et al. 1999) or PUS7, DUS1, DUS2, DUS3, DUS4 (pseudouridine or dihydrouridine) (Behm-Ansmant et al. 2003; Xing et al. 2002; Xing et al. 2004) resulted in low or lethal interactions (Alexandrov et al. 2006).

1.3.3 tRNA-derived fragments

Initially, they were thought to be non-functional residues of random tRNA degradation, but advances in high-throughput sequencing allowed the identification of a novel class of small regulating RNA, tRNA-derived fragments (tRF). This field still lacks a clear consensus on the nomenclature and classification of the tRNA-derived fragments, but they are typically divided into two categories based on their length and cleavage sites.

In the first category are stress-induced tRNA fragments (tiRNA), with 31 to 40 nucleotides in length (Y. Xie et al. 2020), generated by angiogenin (ANG, a member of the ribonuclease RNase A superfamily). The processing is triggered by stress conditions such as hypoxia, UV radiation or heat shock and occurs within the anticodon loop of mature tRNA, generating tRNA halves (tRH) (Yamasaki et al. 2009; Fu et al. 2009; Elkordy et al. 2018; Z. Su et al. 2019).

The second category is made of tRNA-derived fragments cleaved by Dicer, a member of

the RNase III family (Babiarz et al. 2008; Cole et al. 2009; Yeung et al. 2009; Maute et al. 2013), at or near D or T-arms generating fragments of 14 to 30 nt in length corresponding to 3' tRF or 5' tRF as well as the internal part 2-tRF (Telonis et al. 2015). tRF-1 are small fragments derived from precursor tRNA's 3' trailer end (**Figure 11**). These tRF have poly-U ending and length ranging from 15 to 22 nt (Haussecker et al. 2010; Kumar et al. 2014). tRF-3 or tRF-5 are further classified into tRF-5a(∼15 nt), tRF-5b (∼22 nt), tRF-5c (∼31 nt), tRF-3a (∼18 nt), and tRF-3b (∼22 nt) (Kumar et al. 2015). Contrasting evidence presented by Kumar et al. 2014 and Kuscu et al. 2018 showed that Dicer was not the exclusive tRF cleaving enzyme, indicating an involvement of other RNase(s).

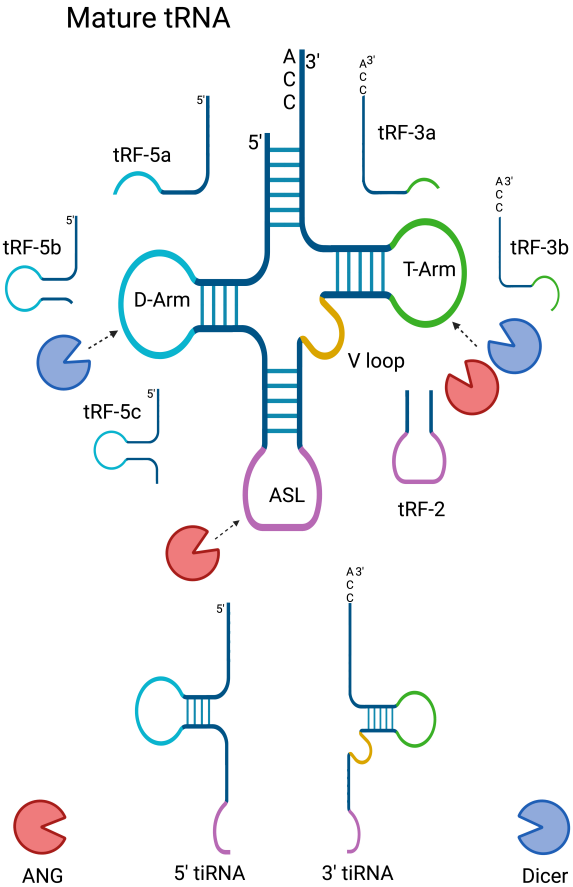


Figure 11: *Illustration of a mature tRNA. tRF-2 contains the stem sequence and the anticodon loop portion of the mature tRNA. tRF-3 is the 3' end fragment of tRNA produced by cleavage of the mature tRNA T-arm by nucleases. tRF-5 is produced by Dicer enzyme cleavage of the D-arm of mature tRNA. 5' tiRNA and 3' tiRNA are 5' and 3' fragments of mature tRNA produced by cleavage from the anticodon loop, respectively. 5' and 3' tiRNA are produced by cleavagge of the anticodon loop.*

These results suggest that tRNA can be processed into multiple functional tRNA-derived

fragments.

Although very useful, high-throughput techniques still face challenges in tRNA fragment detection, such as high sequence similarity of tRF, which require precise detection methods, ideally with base resolution. Next, tRF, particularly tRNA halves, have chemically modified termini due to RNase cleavage, incompatible with sequencing adaptor ligation. A different challenge is when tRF inherit internal modification, which strongly hinders reverse transcription (RT) during cloning, leading to inaccurate quantification of tRF. For small RNA sequencing, this effect could lead to either underestimation (RT after adaptor ligation) or overestimation (RT before adaptor ligation) of the tRF levels (Reviewed in Z. Su et al. 2020). Because of these limitations, combining different detection approaches is vital for correct tRF detection.

Cleavage of tRNA does not occur randomly; on the contrary, the expression of tRNA fragments in cells and tissues is highly specific (Kumar et al. 2014); this feature was also observed for vtRNA. For instance, specific tRNA modifications were found to regulate the biogenesis of tRNA fragments. One such post-transcriptional modification is (m^5C), located on positions 48/49 deposited by NSUN2 methyltransferase. The loss of this modification resulted in tRNA secondary structure rearrangements, allowing tRNA cleavage. In neurons, a decrease of (m^5C) 48/49 on tRNA increased 5' tRNA half levels due to the higher binding affinity of hypomodified tRNA with ANG (Blanco et al. 2014). This was an interesting observation from the perspective of vault RNA, which displayed the opposite phenotype; NSUN2 vtRNA modification increased the levels of 3' end fragments. In contrast, the loss of tRNA modifying pseudouridine synthase PUS7, resulted in the downregulation of novel PUS7-activated 18-nt long tRNA fragments from tRNAAla, tRNACys, and tRNAVal, leading to an impaired translation initiation complex (Guzzi et al. 2018).

Overall, these examples highlight the importance of tRNA modifications on tRF biogenesis. It is worth noting that the levels of tRNA-derived fragments in the cell do not manifest the amount of the corresponding tRNA (Kumar et al. 2014; Reifur et al. 2012). Several tRF functions have now been well described, and more of them are emerging as critical regulators of a broad and diverse set of biological processes (Keam et al. 2015; Y. Xie et al. 2020).

1.3.4 tRNA-derived fragments with functional relevance

It has become clear that tRF have diverse functions and mechanisms of action, not necessarily with a unifying rule. The localisation of 5' and 3' tRF was mainly in the cytoplasm (Kumar et al. 2014; Haussecker et al. 2010); however, Kumar et al. (2014) showed enrichment of 5' tRF in nucleoli of HeLa cells using deep sequencing, demonstrating a functional role in both cellular compartments. As small non-coding RNA such as microRNA are known to regulate gene expression, this led to speculation that tRNA-derived fragments might also participate in the same or similar mechanism. Several groups have independently focused on characterising tRF translation inhibition by sequence-specific mRNA silencing of translation mechanism components (Kuscu et al. 2018; Gebetsberger et al. 2012; Ivanov et al. 2011). This tRF-mediated silencing is dependent on AGO2-mediated target inhibition via RNAi. Sobala et al. 2013 showed that tRF-5 derived from tRNAGln were able to reduce translation in the absence of base-pairing only with conserved 3' 'GG' dinucleotide. Taken together these data suggest tRF-specific inhibition of global translation dependent on the presence of multiple guanines.

It has been shown that tRF are increased in cancer. The investigation of specifically elevated tRF-1-Ser-TGA (19 nt) in the colon cancer cell line (HCT116) revealed its importance for cell proliferation and cell cycle propagation (Kumar et al. 2014). Another study showed 3' tRF-Leu (17 nt) as a tumour suppressor involved in the inhibition of proliferation and colony formation in colorectal cancer cells. This was demonstrated as 3' tRF-Leu suppressed the Notch signalling pathway that is important for cancer progression, metastasis and proliferation through direct binding and inhibition of 3' UTR of Jagged Canonical Notch Ligand 2 (JAG2) (Huang et al. 2017). In contrast, H. K. Kim et al. (2017) showed that inhibition of 3' end tRF-Leu (22 nt) caused apoptosis of the tumourigenic cells, demonstrating that this tRF plays an essential role in cancer cells (HCT116 and HeLa) viability and can be used as a potential drug target. These data show completely different roles for two similar tRNA-derived fragments, highlighting the complexity of tRF-mediated processes in tumour progression.

Interestingly, a recent study showed that similarly to vault RNA, tRF can be loaded

into EV by YBX1 and provide a promising source of biomarkers and a novel mediators in cell-to-cell communication. This observation further expands the previously identified repertoire of RNA loaded into vesicles and provides another common element between tRNA and vtRNA (Jeppesen et al. 2019; Srinivasan et al. 2019; Z. Wei et al. 2017).

1.4 Post-Transcriptional Gene Silencing, Transcriptional Gene Silencing and Nascent RNA Silencing

1.4.1 Post-Transcriptional Gene Silencing Characterisation and Identification

RNA Interference (RNAi) can control the expression of specific genes on the level of mRNA transcripts through post-transcriptional gene silencing (PTGS). This process can occur in the cytoplasm or nucleus and is responsible for direct sequence-specific degradation of mRNA or blocking mRNA translation into protein. During PTGS, small non-coding RNA (sncRNA) targets the processing complex to the target site. Argonaute proteins and sncRNA, and other cofactors form the processing complex. The sncRNA can be divided into three main groups: small interfering RNA (siRNA), microRNA (miRNA), and piwi-associated RNA (piRNA) (N. C. Lau et al. 2006); however, recent studies revealed novel alternative substrates such as tRNA and vault RNA. In the following paragraphs, I will introduce the post-transcriptional gene silencing mechanism and describe important regulatory factors.

1.4.2 Canonical miRNA biogenesis pathway

miRNA are a class of non-coding RNA known to mediate post-transcriptional gene silencing. miRNA are produced through a multistep process; predominantly transcribed by Pol II as primary-miRNA (pri-miRNA), which can be several kilobases long. Their maturation starts in the nucleus with the Microprocessor complex, which contains nuclear RNAse III protein Drosha and co-factor DiGeorge syndrome critical region 8 (DGCR8) and cleaves the pri-miRNA to generate precursor-miRNA (pre-miRNA) of 60–90 nucleotides in length. These transcripts are further processed and exported into the cytoplasm by Exportin 5 (XPO5) in a RanGTP-dependent manner. However, in a recent study J. Wang et al. 2020 investigated XPO5-associated RNA species, revealing novel interacting partners such as vault RNA and pri-miRNA (mir-17 and mir-15b). Moreover, in the presence of XPO5 the processing efficiency of pri-mir-15b by DROSHA/DGCR8 was strongly enhanced, uncovering another regulatory layer in the miRNA biogenesis (**Figure 12a**).

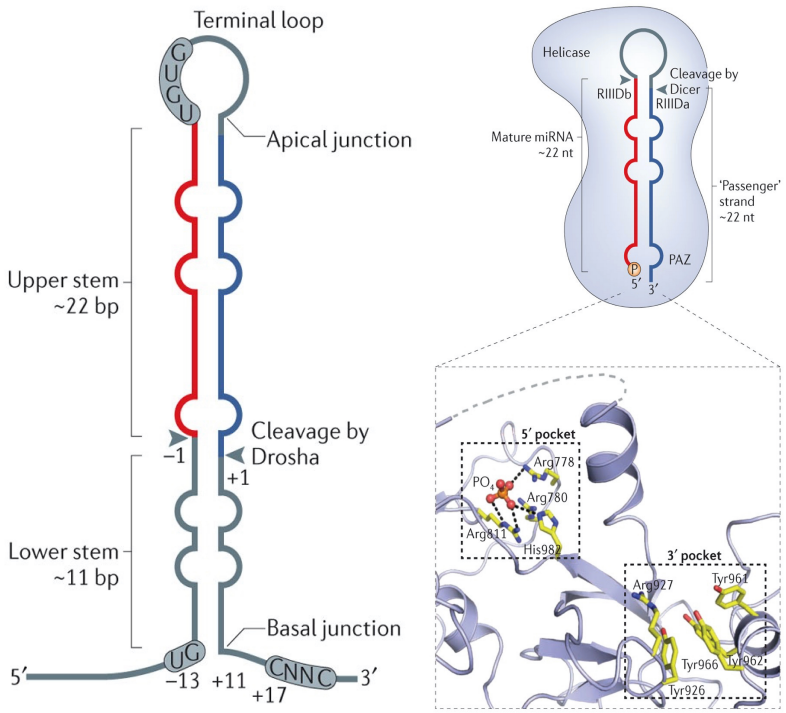


Figure 12: ***(a)*** *The Microprocessor complex recognizes the single-stranded RNA tails and then cleaves primary microRNA at 11 bp from the basal junction and 22 bp from the apical junction.* ***(b)*** *Dicer recognizes pre-miRNA by the PAZ (PIWI –Argonaute (AGO)–ZWILLE) domain which contains two basic pockets. Dicer measures a set distance from both termini (like a 'molecular ruler') and cleaves the pre-miRNA into small dsRNA 23 bp (Figure adapted from Ha et al. 2014).*

In the cytoplasm, pre-miRNA are cleaved by Dicer, a dsRNA-processing enzyme (RNase III family), into miRNA duplexes around 24 nt in length, including 2-nt 3' overhang on each end. This is the final stage of maturation for most miRNA, although Dicer-independent microRNA have been previously described (Cheloufi et al. 2010). Dicer is a 219 kDa multidomain enzyme with an L-shaped structure composed of a head, body, and base (H. W. Wang et al. 2009; P. W. Lau et al. 2012). Within the head of the molecule is the PAZ (Piwi/Argonaute/Zwille) domain which has binding pockets for the 3' end 2-nt overhang and 5' phosphate end of pre-miRNA (**Figure 12b**). The body consists of two RNase III domains (RIIIDa and RIIIDb), which form the catalytic centre for cleaving two strands of the pre-miRNA. N-terminal DExD/H-box helicase domain forms the base of Dicer and is responsible for dsRNA unwinding and translocation. This domain, combined

with the C-terminal double-stranded RNA-binding domain (dsRBD), mediates substrate recognition and binding (Reviewed in Vergani-Junior et al. 2021). Interestingly, dsRBD showed to be important for the Dicer nuclear localisation (Doyle et al. 2013). Finally, the role of DUF283 (domain of the unknown function) remains unclear; however, it has been suggested to regulate RNA-RNA annealing activity (Ciechanowska et al. 2021).

1.4.3 Regulation of Dicer processing

Emerging evidence suggests that Dicer cleavage activity depends on various factors such as RNA substrate secondary structure elements, modifications and RNA binding accessibility. A recent structural study showed that single-nucleotide bulges located at the 22nd nucleotide (22-bulges) from the 3' end are important for accurate and efficient pre-miRNA Dicer processing. Furthermore, it was also found that shorter stem lengths (23-24 bp) exhibited a higher single cleavage activity than longer stems (26-27 bp) independently of loop sizes (Nguyen et al. 2022). Another important processing modulator is RNA modifications. For instance, RNA editing caused by the double-stranded RNA-specific adenosine deaminase (ADAR) catalysed the conversion of adenosine to inosine in the stem region of pri-miR-151, which in turn decreased Drosha and Dicer processing efficiency (Kawahara et al. 2007). Similarly, methylation of primary, precursor, and mature microRNA 125b (miR-125b) by methyltransferase NSUN2 inhibited the biogenesis of miR-125b in *Drosophila*, which led to a repressed gene expression silencing by miR-125b (Yuan et al. 2014).

RNA binding proteins can also influence the maturation of pre-miRNA. In particular, it was observed that LIN28 could suppress Drosha and Dicer processing of miR-let7 through terminal loop binding, inducing oligouridylation and leading to miRNA degradation (Piskounova et al. 2011). The length of the Dicer end product depends on the molecular ruler, which corresponds to the distance between the PAZ domain and the catalytic centre (P. W. Lau et al. 2012), producing RNA duplexes of 20-25 bp from dsRNA substrates (D. Yang et al. 2002). Apart from pre-miRNA, Dicer-mediated cleavage has been demonstrated for several other small non-coding RNA such as tRNA (Di Fazio et al. 2022; Cole et al. 2009; Haussecker et al. 2010), vault RNA (Persson et al. 2009; Minones-Moyano et al. 2013),

H/ACA box (Langenberger et al. 2013), 7SL RNA (Ren et al. 2012) or short induced RNA (exogenous siRNA or endogenous siRNA) (Moore et al. 2010; Paddison et al. 2002). These data highlight the variety of Dicer cleavage substrates.

1.4.4 RISC complex and miRNA loading

After Dicer cleavage, mature miRNA are incorporated into the RNA-induced Silencing Complex (RISC) in an ATP-dependent fashion (Kawamata et al. 2009). RISC is a heterogeneous molecular complex that consists of an Argonaute protein (AGO1-4 in humans) which plays a vital role in RISC-mediated gene silencing (Muller et al. 2019). Argonaute protein comprises four domains: an N-terminal domain, which can split duplexes during the RISC Assembly (Kwak et al. 2012); the MID and PAZ domains that facilitate small RNA binding and 5' and 3' end recognition (Ma et al. 2004; Parker et al. 2005), and the PIWI domain that is capable of target cleavage (Parker et al. 2004). Although the four human AGO share 80% amino acid identity (Sasaki et al. 2003), only AGO2 (Meister et al. 2004) and AGO3 (M. S. Park et al. 2017; M. S. Park et al. 2020) are catalytically active. The target-specific regulation of AGO depends on the interaction with the miRNA duplex (**Figure 13**). This duplex undergoes unwinding, and only one strand (guide) is selected and anchored into the Argonaute protein, while the other one (passenger) is subjected to degradation (Leuschner et al. 2006).

However, how does Argonaute choose which strand of the miRNA duplex should be loaded? There are two significant drivers of strand selection: 5' end nucleotide identity and thermodynamic end stability. Early works suggested that the 5' nucleotide-binding pocket in the MID domain preferentially recognise adenine or uracil at the first nucleotide position over cytosine or guanine (Mi et al. 2008; Frank et al. 2010). This has been further underlined by the fact that uracil is the most commonly observed 5' nucleotide of miRNA guide strands (Hu et al. 2009). Another miRNA characteristic that can favour one strand over the other is thermodynamic stability. The more stable end interacts with the PAZ domain, while the less thermodynamically stable binds the MID domain (Y. Wang et al. 2008; Frank et al. 2010). In particular, it was observed that the PAZ domain actively participates in passenger strand removal by shaking the 3' end of the guide strand (J. H.

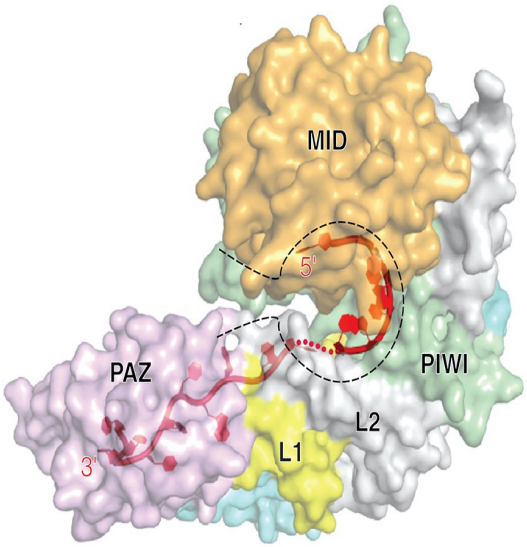


Figure 13: *Illustration of the human Argonaute 2 structure. The linkers, L1 and L2, are colored in yellow and gray, respectively. The guide RNA (red) is depicted as a ribbon model. The disordered parts of the guide are shown as dotted lines (Figure adapted from Nakanishi 2016).*

Park et al. 2015). Taken together, these results strongly suggest that the selection of the guide strand is not a random event and has sophisticated controlling mechanisms.

1.4.5 RISC complex and miRNA interactions

The interaction between the RISC complex and miRNA forms 'activated RISC' (miRISC), which can regulate mRNA translation and stability (J. Liu et al. 2004). The canonical miRNA-mRNA target sequence complementarity, which functions as target recognition, was found between the miRNA 'seed' region (2-7 nucleotides) (Bartel 2018) at the 5' end and the 3' UTR of target mRNA. The nucleotide at position 8 was further confirmed to influence the specificity of miRNA–target interactions (Brancati et al. 2018). For comparison, siRNA (21 nt) requires perfect base pairing to the target sequence. In addition to the canonical complementarity, base-pairing at the 3' end (nucleotides 13-18, supplementary region of 23 nt long miRNA) can regulate the specificity of targeting (**Figure 14**), the miRNA stability can compensate for mismatches in the seed (Broughton et al. 2016; Grimson et al. 2007; Helwak et al. 2013; Y. Xiao et al. 2020). This extended pairing is initiated after the seed pairing and AGO conformational change, exposing nucleotides in

the 3' half of the miRNA (Schirle et al. 2014). In line with these findings, mutation studies of small hairpin RNA miB revealed that miRNA-mRNA base pairing beyond the seed triggered distinctive AGO-binding patterns through more stable base pairing, increasing the probabilities of target silencing (Yan et al. 2018). The pairing downstream of the seed occurs when the PAZ domain becomes less tightly bound to guide RNA; the free 3' end facilitates the sequence complementarity process of the supplementary region, forming the duplex on both sides. The complementarity in the central region facilitates efficient cleavage (Yan et al. 2018). Moreover, the location and size of the centre are the essential factors affecting the efficiency and specificity of gene regulation mediated by miRNA (Ye et al. 2008).

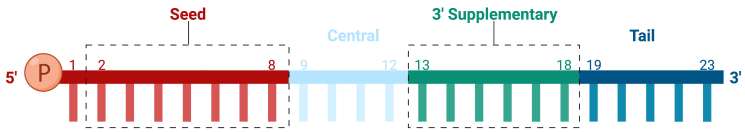


Figure 14: *Illustration of a guide RNA. A guide RNA (length 23) can be split into four regions, seed (red), central (light blue), 3 supplementary (green) and tail (dark blue). The numbers above represent the nucleotides positions (Figure adapted from Nakanishi 2022).*

These findings highlight the importance of sequences beyond the seed region in controlling the function and existence of miRNA. In conclusion, the extended seed region and multiple target binding sites can increase AGO efficacy.

1.4.6 Template target area

The target region is an equally crucial factor to consider when assessing miRNA functionality. Kertesz et al. 2007 used mutational analysis in the experimentally validated miRNA binding sites to observe target site accessibility. They demonstrated that the introduced mutational changes caused reduced target accessibility and significantly decreased miRNA-mediated translational repression (Kertesz et al. 2007). Interestingly, it was found that a single miRNA targets multiple transcripts (Lal et al. 2009; Lal et al. 2011) and can act in combination with another miRNA to regulate the target expression (Cursons et al. 2018). The majority of miRNA target sites are located in the 3' UTR, although pairing to 5' UTR, Coding DNA Sequence (CDS), introns (Meng et al. 2013) and

Intergenic regions were also detected (Plotnikova et al. 2019). Particularly, in *C. elegans*, dsRNA induced silencing of a pre-mRNA by targeting intronic regions, thus demonstrating that RNAi can also regulate nuclear transcripts (Han et al. 2009). Similarly, Hoffer et al. 2011 showed in *A. taliana* that nuclear PTGS could control pre-mRNA levels using siRNA-mediated degradation of target RNA. Furthermore, in nuclei of human cells, PTGS was demonstrated to modulate the expression of long non-coding RNA (lncRNA) metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) via miR-9 (Leucci et al. 2013). These results show that PTGS can occur in the nucleus and that mRNA features are essential for miRNA function.

miRNA show predominantly two modes of regulation of their target sequences, such as mRNA degradation or translation inhibition. The effect of both modes is decreased expression of the protein encoded by the target RNA. The target specificity of miRISC is determined by interaction with the target mRNA sequence, called miRNA response element (MRE). The degree of this interaction defines whether there is AGO2-dependent slicing of target mRNA or miRISC-mediated translational inhibition and target mRNA decay (Jo et al. 2015). A perfect base-pairing miRNA-MRE induces AGO2 endonuclease activity and initiates mRNA cleavage (Jo et al. 2015). Recent studies revealed other regulatory mechanisms, such as mRNA deadenylation, decapping, and exonucleolytic digestion of target mRNA (Behm-Ansmant et al. 2006; Giraldez et al. 2006; Wu et al. 2006).

1.4.7 Non-canonical Biogenesis Pathways

These pathways exclude individual proteins involved in the canonical pathway, mainly Drosha, Dicer, Exportin 5, and AGO2. The non-canonical miRNA biogenesis can be divided into two groups Drosha/DGCR8-independent and Dicer-independent pathways. The canonical Microprocessor maturation of pri-miRNA described above has alternative biogenesis mechanisms independent of Drosha. For instance, Mirtrons are a type of microRNA produced during mRNA splicing from introns of other genes; they can fold into pre-miRNA, omitting Drosha cleavage (Ruby et al. 2007; Babiarz et al. 2008). Similarly, a group of pre-miRNA with the 7-methylguanosine (m^7C)-cap can bypass the Drosha cleavage and is directly exported to the cytoplasm (M. Xie et al. 2013). Additionally, other

sncRNA such as tRNA, snoRNA (J. S. Yang et al. 2011), vault RNA (Persson et al. 2009) or short interspersed nuclear elements (SINEs) (Babiarz et al. 2008) can be processed into smaller fragments independently of Drosha by Dicer cleavage. Dicer-independent miRNA are, on the other hand, processed by Drosha into pre-miRNA from endogenous shRNA, although they still require AGO2-dependent slicing to complete their maturation within the cytoplasm (J. S. Yang et al. 2010). These findings indicate alternative ways of generating small RNA fragments from non-canonical RNA templates.

1.4.8 Transcriptional Gene Silencing Characterisation and Identification

Transcriptional gene silencing (TGS) is a process of mediated gene expression through promoter epigenetic modifications regulated by binding of small non-coding RNA. In the following paragraphs, I will describe the mechanisms of this regulation and, with examples, highlight its importance.

Transcriptional gene silencing was, for the first time, observed in transformed tobacco plants by Matzke et al. 1989 as a process which suppressed phenotype of a transgene. Careful analysis revealed suppression of the targeted gene by DNA methylation. Similarly to PTGS, TGS also requires guide RNA and members of the Argonaute family, thus leading to the question of how different is TGS from PTGS? The main differences between the two gene silencing processes are in their mechanism and regulatory regions.

The mechanism of TGS in human cells is directed by siRNA, which is translocated into the nucleus via AGO1/2 (D. H. Kim et al. 2006; Janowski et al. 2006). The AGO-bound siRNA is guided to the promoter region of the target gene, while in PTGS, the siRNA targets gene features such as intron (in the nucleus), exon or UTRs. Hawkins et al. 2009 performed an experiment with two different siRNA; one was designed to target the promoter, and the other targeted exon 1 of the coding transcript. The results showed that the target genes were functionally repressed in a TGS and PTGS manner, respectively, to the used siRNA. In the nucleus, siRNA interacts with target transcript via RNA-RNA hybridisation, which triggers recruitment of DNA methyltransferase to histone 3 lysine's 9 and 27 (H3K9 and H3K27, respectively) at the promoter sites of target genes leading to the transcription shut down (Janowski et al. 2005; Castanotto et al. 2005). Furthermore,

the gene expression changes caused by epigenetic modifications have long-term effect in comparison with PTGS.

Apart from miRNA and siRNA, later studies identified a novel regulator of TGS; long non-coding RNA (Reviewed in Morris 2009). In particular, antisense lncRNA binds to the promoter sequence and stimulates DNA methylation, thus causing the repression of transcription. For example, lncRNA overexpression was shown to downregulate the expression of tumour suppressor genes such as p15 (W. Yu et al. 2008) and p21. Taken together, these data display different mechanisms from PTGS with novel regulatory molecules.

1.4.9 Nascent RNA Silencing Characterisation and Identification

As introduced earlier, tRF can influence gene levels either via mRNA degradation or translation repression in the cytoplasm. A recent work by Di Fazio et al. 2022 investigated the involvement of tRNA-derived fragments in nuclear gene expression regulation. The findings presented in this study introduced a novel mechanism of tRF-mediated nascent RNA silencing (NRS) of target genes in nuclei. The described mechanism required Dicer-dependent processing of tRNA into smaller fragments and catalytically active AGO2 for successful target cleavage. The predicted tRF binding sites on nascent RNA were located in the intronic regions, which was further validated by CRISPR-Cas9 Knock Out of the genomic segment with predicted binding locations in SPINT1, one of the tRF predicted targets. Thus causing an increase in SPINT1 expression at both nascent and steady-state RNA levels as well as protein levels. The NRS mechanism was further validated using synthetic single-stranded tRF targeting EGFR (tsEGFR) in BT549 breast cancer cells, resulting in the repression of EGFR at steady-state, nascent and also protein levels. Moreover, another identified tRF against BCL2 showed the exact effect in MCF7 breast cancer cells. In summary, these data demonstrated novel group of tRNA-derived fragments, which can regulate gene silencing by the novel NRS mechanism differently from already described PTGS and TGS mechanisms.

2 The biogenesis of vault RNA-derived small RNA fragments

The work reported in this dissertation is summarised into Terenova et al. and deposited on bioRxiv[1]*. All the computational analyses have been done in the laboratory of Dr Gullerova by myself. The experimental sample preparation for small RNA-seq in Dicer Knock Down (KD) and Wild Type (WT) conditions was carried by Dr Davis and chrRNA-seq samples in Dicer and AGO2 KD and WT conditions were prepared by Dr Pong. Dr Adele Alagia performed the confocal microscopy using fluorescently labelled svtRNA1-2. I have developed the project idea of vault RNA, while analysing different RNA sequencing datasets, I noticed their significantly changed expression. The published literature showed their involvement in gene expression regulation, therefore I decided to investigate it. I have performed all the computational data analysis and interpretation.*

2.1 Introduction

Advances in high-throughput sequencing have revealed that cleavage of sncRNA can generate smaller fragments, which are products of not specific RNA degradation. In the past years, several studies showed that other small non-coding RNA (<300 nt) apart form microRNA, such as tRNA (Di Fazio et al. 2022; Kuscu et al. 2018), scaRNA (McLaurin et al. 2020) or snoRNA (Reviewed in Wajahat et al. 2021) can also be processed into smaller functional fragments via alternative RNAi pathways. These fragments were then found to be involved in regulating gene expression through various silencing mechanisms, including PTGS, TGS and NRS or contributing to the regulation of alternative splicing, depending on their cellular localisation. Despite the recent efforts in their identification, we still lack systematic analysis methods for their recognition. Vault RNA are a type of sncRNA which was also demonstrated to be processed into smaller fragments independently of Drosha but dependent on Dicer, namely vtRNA1-1 (Persson et al. 2009) and vtRNA2-1

[1]https://doi.org/10.1101/2022.08.12.503724

(Minones-Moyano et al. 2013). For instance, the identified 5' and 3' end fragments of vtRNA1-1 were shown to regulate the expression of CYP3A4 (Persson et al. 2009) and CACNG7/8 mRNA (Hussain et al. 2013) respectively, through sequence complementarity in the cytoplasm. However, vtRNA1-2 and vtRNA1-3 have not been well explored, neither as intact RNA nor their fragments. In this study, I aimed to investigate the biogenesis of small vault RNA derived fragments including all four vtRNA using computational and experimental approaches. Due to vtRNA's high sequence and structure similarity, I hypothesised that the small RNA biogenesis might be shared among all the vtRNA, including vtRNA1-2 and vtRNA1-3. I found novel small RNA fragments derived from vtRNA1-2 that are generated in Drosha-independent but Dicer-dependent manner. Interestingly, I observed that svtRNA1-2 is associated with Argonaute proteins in the cytoplasm and the nucleus. Knock Down experiments, and computational predictions revealed a group of genes in which nascent transcripts can be targeted by svtRNA1-2 in the nucleus causing alterations in their expression. These results provide evidence for svtRNA1-2 biogenesis and introduce the idea of its role in nascent RNA silencing.

2.2 Results

2.2.1 Drosha-independent small vault RNA biogenesis

The majority of vtRNA are not associated with vault complex; instead, they were demonstrated to participate in different biological processes, bound by different RNA Binding Proteins (RBPs). vtRNA1-1 and vtRNA2-1 are the most studied vault RNA in which fragments were identified and described. In this work, I considered the remaining vtRNA1-2 and vtRNA1-3, which might also produce regulatory fragments. In order to investigate vtRNA processing independently of Drosha, I utilised two small RNA sequencing (sRNA-seq) datasets in WT and Drosha KD and KO conditions from the publicly available data repository Gene Expression Omnibus (GEO)(GSE77989, GSE55333, **Table 1**). The obtained raw read files from Drosha KO and WT conditions were processed and aligned to the human reference genome (hg38).

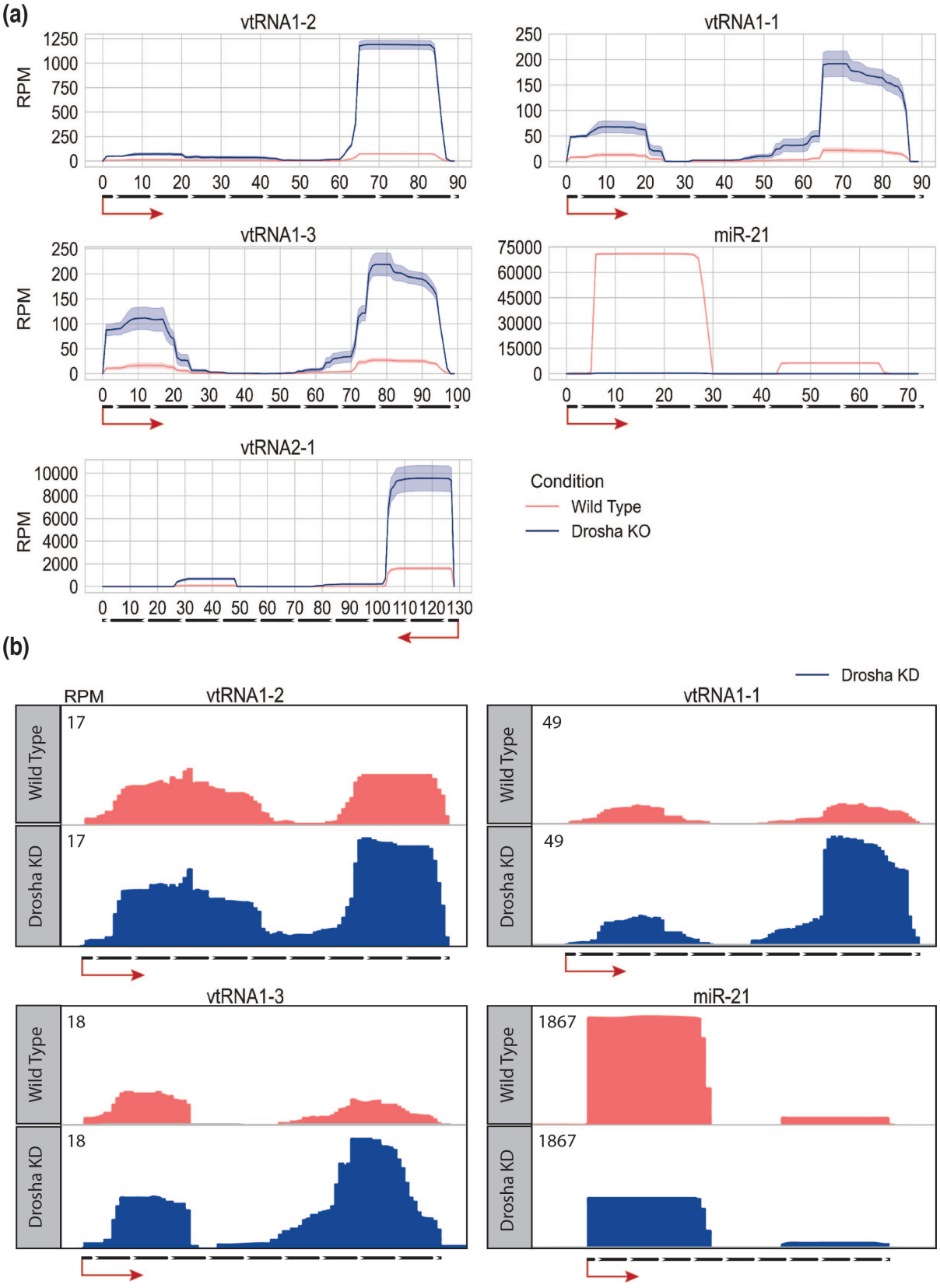


Figure 15: Caption on next page

Figure 15: *(previous page)* ***(a)*** *Per base read coverage of vtRNA. The x-axis shows the length of transcripts, and the y-axis shows the normalised number of mapped reads in RPM (Reads Per Million). miR-21 acts as a positive control. The blue line shows results in Drosha KO [n=2] and the red line in WT [n=2] conditions using the HCT116 cell line. The shaded area around the line represents the standard deviation of the mean. The red arrow shows the direction of transcription. The results were visualised using python3.* ***(b)*** *Per base read coverage of vtRNA. The x-axis shows the length of transcripts, and the y-axis shows the normalised number of mapped reads in RPM. miR-21 acts as a positive control. Blue line shows results in Drosha KD [n=1] and red line in WT [n=1] conditions using the HEK293 cell line. The results were visualised using IGV.*

The alignment output was then used to calculate normalised read count values for all four vtRNA based on their genomic coordinates (see Methods, section 6.1.3). I found an accumulation of reads at the 3' and 5' end of vtRNA upon Drosha KO (**Figure 15a**), indicating that vtRNA are processed independently of Drosha. I reasoned that Dicer has less pre-microRNA substrate to process due to Drosha's absence, leading to a higher amount of Dicer processed vtRNA fragments. For comparison, the expression level of microRNA-21 (miR-21), dependent on Drosha processing from pri-miRNA to pre-miRNA, had vastly reduced number of reads mapped to the 3' end. The second processed dataset contained small RNA isolated from HEK293T cells in siRNA-induced Drosha KD and WT conditions. I applied a similar downstream analysis using the provided raw reads, but due to the lack of biological replicates, I used the Integrated Genomics Viewer (IGV) for qualitative data analysis. Similarly to the previous results, vtRNA1-1/2/3 showed up-regulated and miR-21 down-regulated expression of 3' but also 5' end fragments, which were not detected in the previous dataset upon Drosha KD condition (**Figure 15b**). vtRNA2-1 is not displayed as it did not have sufficient read coverage. Additionally, I also analysed formaldehyde crosslinking, immunoprecipitation, and sequencing (fCLIP-seq)(GSE93653, **Table 1**) of RNA bound by Drosha in HEK293 cells to investigate the presence of Drosha cleavage sites on vtRNA. The Drosha RNA cleavage of pri-microRNA generates 5' and 3' flanking fragments and pre-microRNA. These fragments have a longer length therefore a paired-end library of 75 bp was used to sequence the isolated RNA. After the data processing, I visualised the mapped reads in IGV together with sRNA-seq of RNA isolated from HEK293 WT cells (mentioned in the analysis above).

I observed the cleavage products for miR-21 as well as for another Drosha-dependent

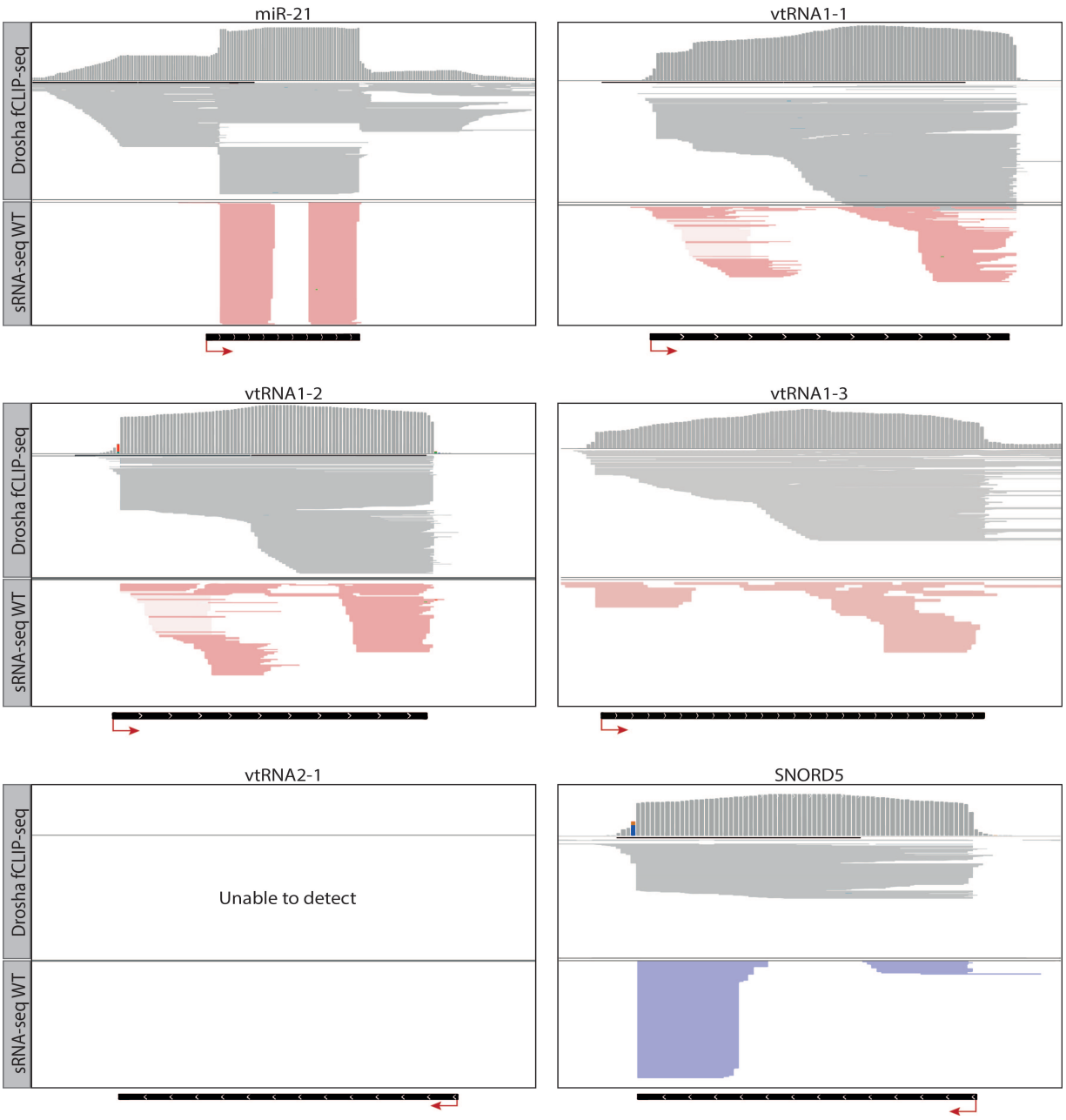


Figure 16: *Mapped reads from Drosha fCLIP-seq (gray) and reads from WT HEK293T small RNA-seq (red/blue) to the pri-miR-21 (used as positive control) vtRNA1-1, vtRNA1-2, vtRNA1-3 and snoRNA SNORD5 (used as negative control).vtRNA2-1 was unable to detect. The x-axis shows the length of transcripts, and the y-axis shows mapped reads. The results were visualised using IGV.*

miRNA (**Figure 16**). However, I did not observe any evidence of such fragments for vtRNA. Moreover, Drosha-independent snoRNA showed a comparable lack of any cleavage fragments, confirming the vtRNA Drosha-independent processing. These results support

the previously published results of Drosha absence in vtRNA1-1 and vtRNA2-1 fragmentation and extend the list of non-cleavable Drosha substrates for vtRNA1-2 and vtRNA1-3.

2.2.2 Dicer-dependent small vault RNA biogenesis

In the results of the previous analyses, I noticed the presence of small RNA fragments at the extremities of vtRNA. To investigate Dicer's involvement in vtRNA cleavage, I analysed sRNA-seqs isolated from HEK293T cells in WT and short hairpin RNA-induced Dicer Knock Down conditions (**Figure 17a**) (GSE126751, **Table 1**).

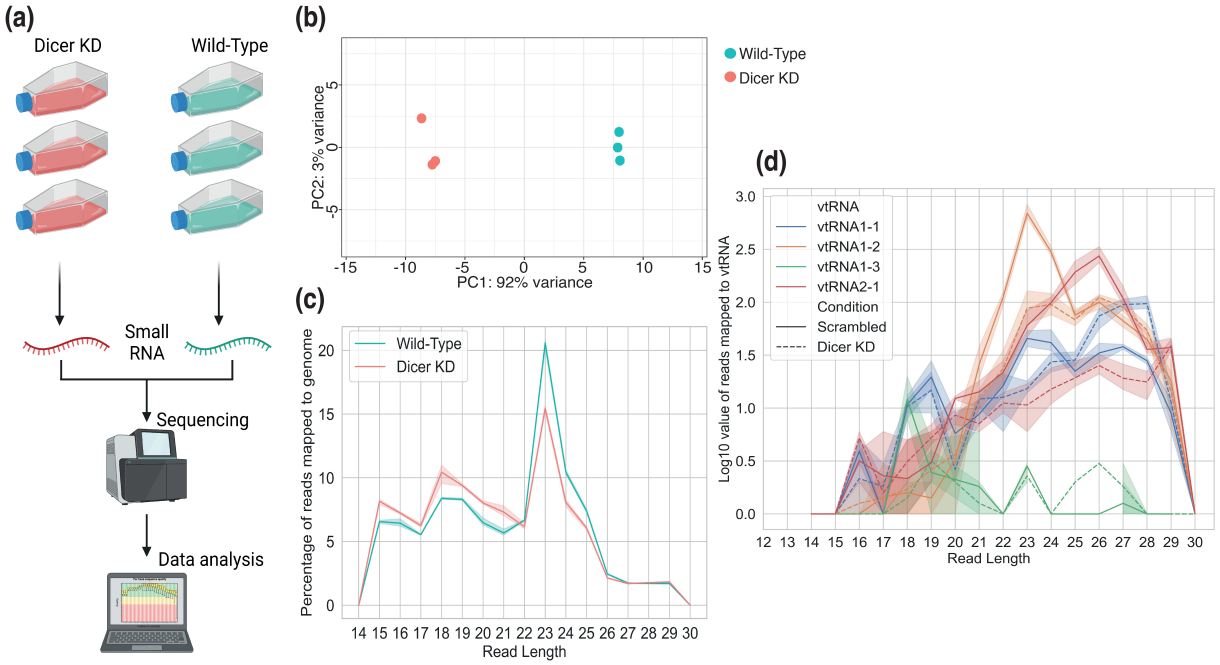


Figure 17: ***(a)*** *Schematic representation of small RNA-seq workflow from HEK293T in Dicer KD (red) and WT (green) conditions (see Methods, section 6.1.5 for more details). Samples prepared by Dr Davis.* ***(b)*** *Principal component analysis plot shows three biological replicates corresponding to the samples from HEK293T cells in WT and Dicer KD conditions. The RNA composition of cells in the WT differed significantly from those in the Dicer KD.* ***(c)*** *Size distribution of mapped reads to the human genome in Dicer KD and WT.* ***(d)*** *Size distribution of mapped reads to the vault RNA in Dicer KD (dashed line) and WT (full line). The shaded area around the line represents the standard deviation of the mean.*

The sequencing was performed using single-end libraries generating 50 bp-sized reads. These reads were trimmed and size selected (15-29), as these lengths fall into the average mature miRNA sizes (**Supplementary Figure 35b**), and mapped to the human reference

genome (hg38) using the STAR aligner (Dobin et al. 2013) with recommended settings for mapping of small RNA (see Methods, section 6.1.5).

A principal component analysis (PCA) outlined the apparent differences between the two conditions and the tight clustering of the replicates (**Figure 17b**). The size distribution of the mapped reads (**Figure 17c**) showed a distinct Dicer-dependent peak between 22 to 25 nt, including reads corresponding to mature miRNA (21-23 nt). The small difference in this peak between WT and Dicer KD is most likely affected by the presence of induced shRNA in the Dicer KD sample (used to generate Dicer KD), which might mask the absence of a proportion of mature miRNA. Interestingly, the fragments derived from vtRNA1-2 were mostly 22 to 24 nt long, suggesting that they might be Dicer products (**Figure 17d**).

I built a custom sncRNA annotation file including miRNA, miscRNA, rRNA, scRNA, snoRNA, snRNA, tRNA and vtRNA, which allowed me to investigate the global profile of small non-coding RNA expression changes upon Dicer KD. Surprisingly, these results revealed a mild effect on the downregulation of miRNA (**Figure 18a**). The large proportion of small non-coding RNA consists of microRNA and tRNA; as expected, these two groups were also found prevalent in the analysed dataset (810 and 488 genes, respectively) (**Figure 18a**). Among the identified sncRNA, I also identified the four species of vault RNA. In order to determine whether vtRNA expression becomes affected in response to Dicer KD, I performed a differential expression analysis, revealing 205 significantly up-regulated sncRNA, including the higher proportion of miRNA (45.4%) followed by snoRNA (22.4%). The 304 significantly down-regulated sncRNA in majority consisted of miRNA (77.6%) (adj. p-value < 0.05 and Fold Change (FC) $> 1.5 \mid < -1.5$) (**Figure 18b**). Interestingly, vault RNA were identified in both conditions. Specifically, vtRNA1-1 was significantly up-regulated, and vtRNA1-2 and vtRNA2-1 were significantly down-regulated upon Dicer KD, suggesting Dicer is important for their biogenesis (**Figure 18c**). Moreover, the expression of vtRNA1-1 significantly increased upon Dicer KD compared to WT, whilst the expression of vtRNA1-3 was not significantly altered (**Figure 18c**), indicating that in HEK293T cells, Dicer is only involved in the processing of vtRNA1-2 and vtRNA2-1.

To verify the source location of the fragments, I calculated the per-base read coverage

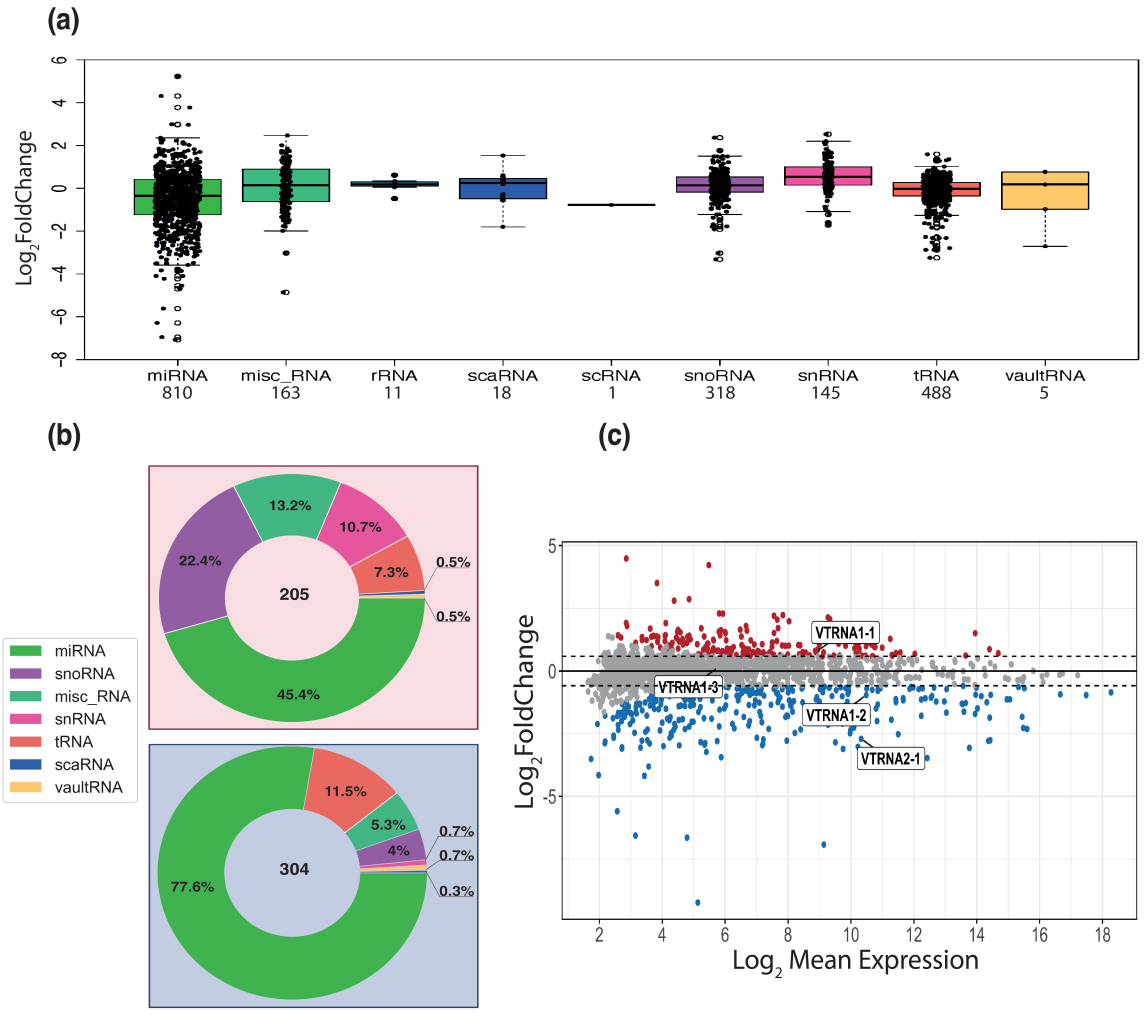


Figure 18: ***(a)*** *Box plots showing the Log_2Fold Change (y-axis) of various small non-coding RNA types (x-axis) upon Dicer Knock Down. The number of genes represented for each sncRNA type is shown below the name.* ***(b)*** *Doughnut chart of percentage of sncRNA biotypes. The top one (highlighted in red) shows biotypes of significantly up-regulated sncRNA (205). The bottom doughnut chart (highlighted in blue) shows biotypes of significantly down-regulated sncRNA (304).* ***(c)*** *MA plot showing differentially expressed genes between Dicer KD [n=3] and WT [n=3]. Statistically significant changes are coloured in red (205) and blue (304) (increased and reduced levels, adj. p-value < 0.05, $Log_2FC > 0.5 \mid < -0.5$, respectively). Points in grey do not have significant values. The labels are provided for vtRNA of interest.*

of the differentially expressed vtRNA in both WT and Dicer KD conditions. The obtained results showed regions along the vtRNA gene sequence with high read coverage located at the 3' end of vtRNA1-2 and the 5' end of vtRNA2-1, which is decreased in Dicer KD, indicating that vtRNA1-2 and vtRNA2-1 are processed into small-vault RNA fragments in a Dicer-dependent manner (**Figure 19a**).

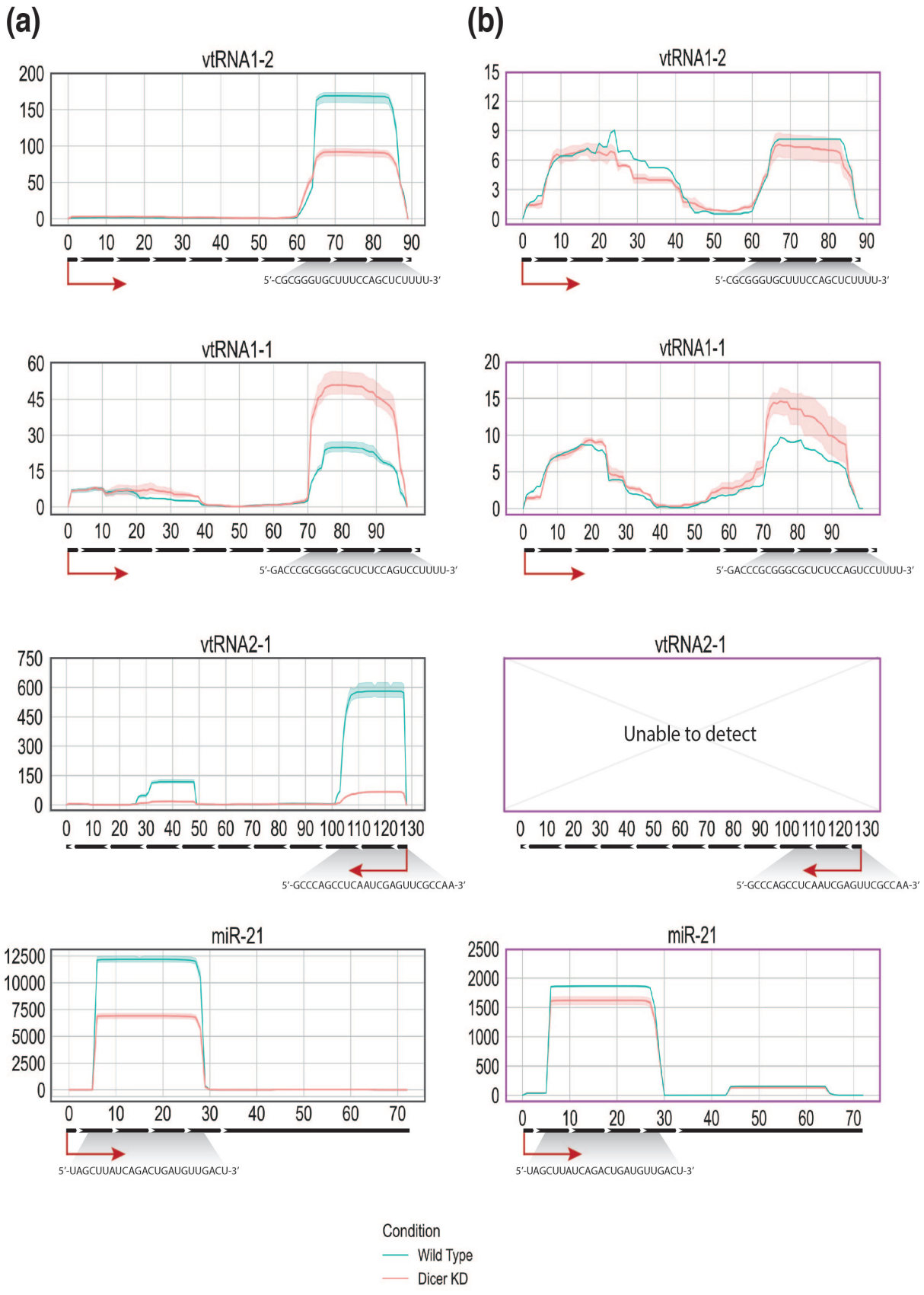


Figure 19: Caption on next page

Figure 19: (previous page)***(a)*** *Per-base read coverage plots showing mapped reads to vtRNA1-1/2 ,vtRNA2-1 and miR-21 transcripts, obtained from small RNA transcriptomes of shRNA induced Dicer KD (red) [n=3] and WT (green) [n=3] using the HEK293 cell line. The y-axis shows the expression as RPM and the x-axis shows length of transcript. The black rectangular at the bottom depicts the gene length, and the white and red arrows represent the strand identity. The zoomed sequence shows the identified fragments.* ***(b)*** *same as* ***(a)*** *but data generated from siRNA induced Dicer KD (red) [n=2] and WT (green) [n=1] (HEK293 cell line) small RNA sequencing. Note: no read coverage detected for vtRNA2-1.*

miR-21 acts as a positive control for Dicer KD and shows a decrease in the 3' end mature miRNA upon Dicer KD. The determined fragment located in the 5' region of vtRNA2-1 has been previously identified (Minones-Moyano et al. 2013; Fort et al. 2020). However, the small RNA derived from the 3' end of vtRNA1-2 is a newly found vtRNA fragment. These results were strengthened by further analysis of another small RNA sequencing dataset from the HEK293T cell line in siRNA-induced Dicer KD and WT conditions (**Figure 19b**) (GSE126751, **Table 1**) and Dicer PAR-CLIP dataset from the HEK293 cells, which identified distinct Dicer binding sites along the vault RNA with a strong preference for vtRNA1-2 (**Supplementary Figure 37**) (GSE55333, **Table 1**).

The identified fragments were mostly found at the extremities of vtRNA. This finding prompted me to investigate conserved features which might play a role in the vtRNA processing and function. The alignment of all four vtRNA sequences detected higher similarity at the 3' and 5' end regions but a lower base consensus in the central domain (between nucleotides 30-60) (**Figure 20a**), which could be relevant for particular binding partners or pathological contexts. Interestingly, those 5' and 3' end regions of vtRNA are highly conserved across vertebrates, whilst the middle part is not (**Figure 20b**), in accordance with the sequence alignment of all four vtRNA. In addition, I used the RNAfold software (Lorenz et al. 2011) with SHAPE data folding constraints from (Luo et al. 2021) to predict and display secondary structures which were then used for prediction of their tertiary structures (**Supplementary Figure 36**) to examine their structural similarities (**Figure 20c**). I observed that the conserved regions fold into stem structures, whereas the less conserved central domain folds into more open-loop structures. Interestingly, in line with my observations, recent work showed that the more accessible unpaired nucleotides

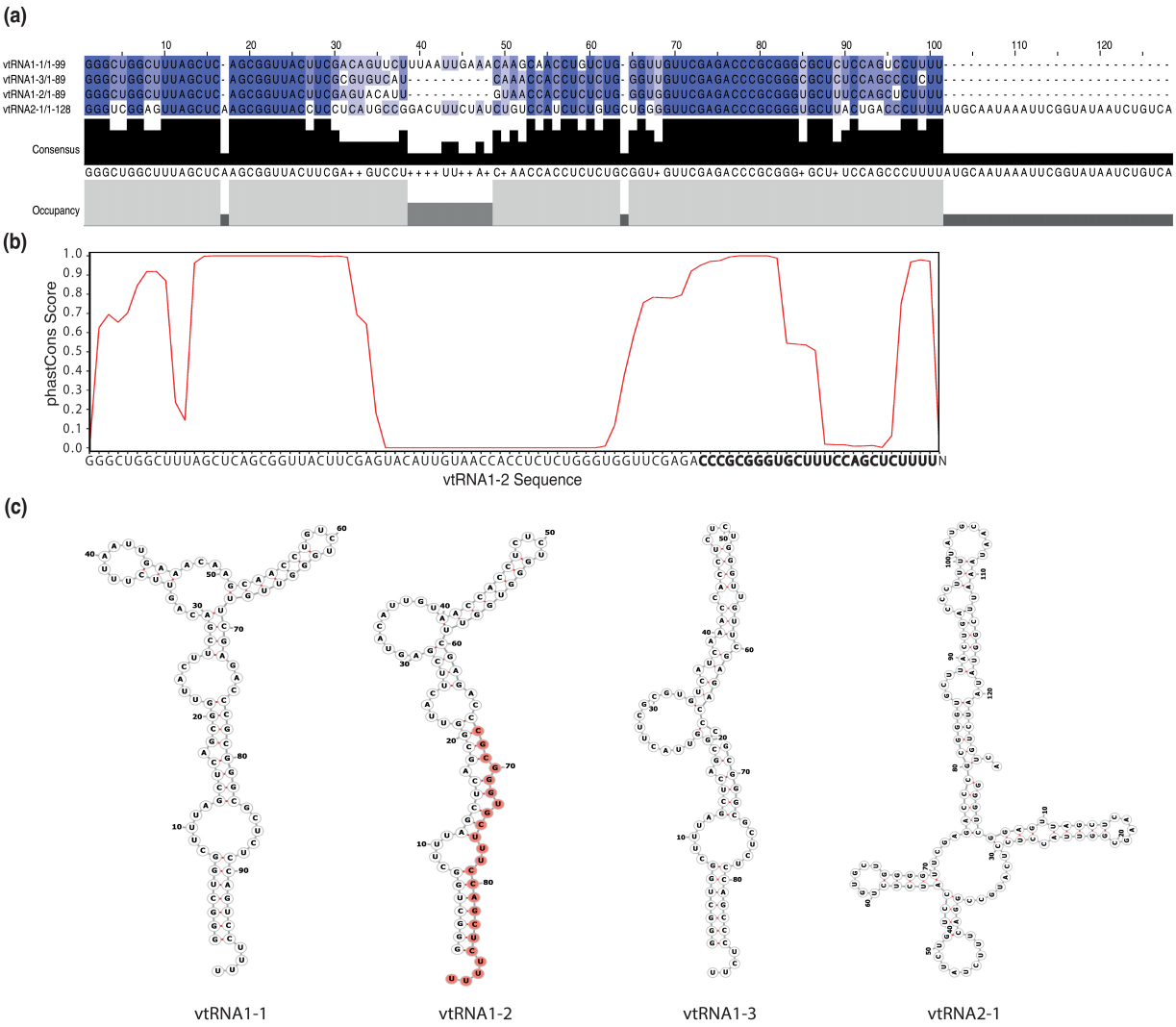


Figure 20: ***(a)*** *Multiple sequence alignment of vault RNA. The residues are colored according to percentage identity, from dark blue: 100% identity to white: less than 50% identity.* ***(b)*** *Conservation analysis of vtRNA1-2. The x-axis shows the sequence of vtRNA1-2, and the y-axis shows the conservation scores calculated from the phastCons multiple alignments of 100 vertebrate species. The 3' end sequence in bold shows the identified svtRNA1-2.* ***(c)*** *Predicted secondary structure of vtRNA based on minimum free energy (MFE). The highlighted region in vtRNA1-2 shows the newly identified 3' end fragment.*

in the central domain of vtRNA1-1 played a crucial role in p62 binding (Horos et al. 2019).

In mice, vtRNA with unknown function is encoded by only one gene, VAULTRC 5, which is longer (143 nt) than any of the human vtRNA. I thought that the processing of human vtRNA into svtRNA might not be a unique feature but rather a general rule for vtRNA. To verify this hypothesis, I analysed a collection of small RNA-seq data from mice tissues (spleen, liver, lungs, muscle, brain, pancreas) (GSE119661, **Table 1**).

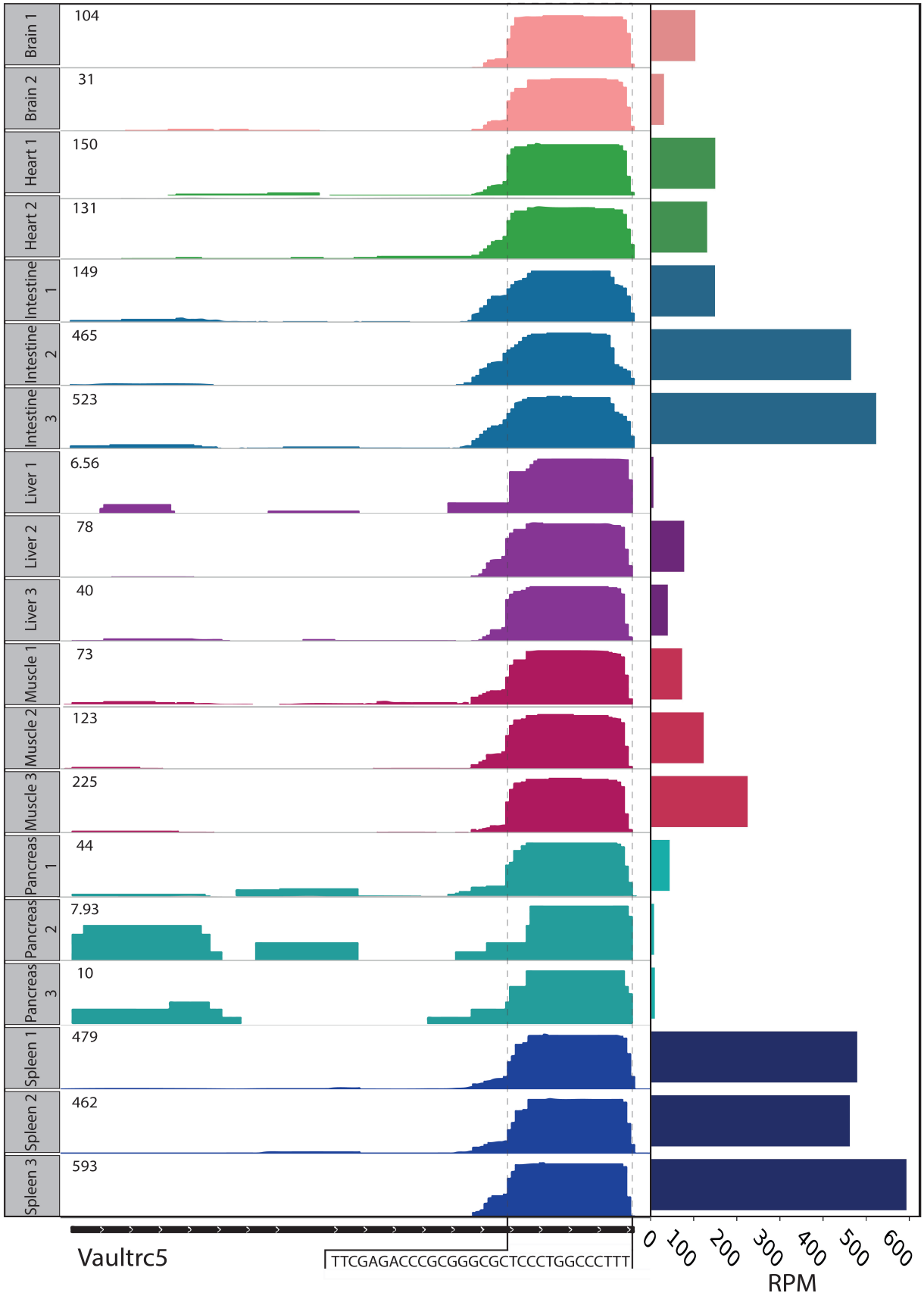


Figure 21: Caption on next page

Figure 21: (previous page) *The IGV tracks showing coverage plots for mouse vault RNA gene (Vaultrc5) from different tissues (Brain, Heart, Intestine, Liver, Muscle, Pancreas, Spleen) in duplicates/triplicates. The tissues are colour-coded and the expression is shown as RPM in the top left corner. The black rectangular at the bottom displays the length of the gene, and the white arrows represent the strand identity. The framed sequence shows the identified 3' end fragment of Vaultrc5.*

Interestingly, I found reads localised to a specific region at the 3' end on the parent vtRNA template, revealing novel small vtRNA fragments (**Figure 21**). Furthermore, I noticed the different expression patterns throughout the analysed tissues, which further supports the tissue-specific expression of vtRNA. These results show the presence of vtRNA-derived fragments in mice tissues, supporting the evidence that svtRNA are not degradation products but instead specific, biologically-generated species. Furthermore, the observed different expression levels of svtRNA in different tissue are in agreement with the previously shown vault RNA tissue-specific expression.

In summary, these results revealed novel Dicer-dependently processed vtRNA1-2 and validated the cleavage of vtRNA2-1 into smaller fragments. At the same time, vtRNA1-1 was found up-regulated upon Dicer KD, thus displaying different expression patterns in contrast with previously published results. Furthermore, vtRNA stem structures are highly conserved, while the more open accessible structures show lower conservation, suggesting individual functional relevance.

2.2.3 svtRNA1-2 associates with Argonaute proteins

The research group of Dr Gullerova previously identified a group of tRNA-derived fragments that can mediate nascent RNA silencing (Di Fazio et al. 2022). As vtRNA share many similarities with tRNA, I wished to investigate the potential role of svtRNA1-2 in gene silencing. A fundamental feature of tRNA-derived fragments in NRS is its association with Argonaute proteins. Therefore, to assess if the svtRNA fragments meet this criterion, I first sought to analyse Argonaute bound targets via photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) (GSE21918, **Table 1**). This method uses incorporation of photoreactive ribonucleoside analogue 4-thiouridine (4-SU) into nascent transcripts and the subsequent UV-crosslinking process to induce T-to-C transitions. The processed reads were mapped to the human reference genome (hg38) and clustered into groups enriched for the T-to-C mutations using Paralyzer (Corcoran et al. 2011). When I displayed the resulting alignment file in IGV, I found svtRNA1-1 and svtRNA1-2 fragments bound across all four AGO proteins. In addition, to support these results, I analysed AGO2/3 RNA-immunoprecipitation sequencing (RIP-Seq) (GSE55333, **Table 1** (see Methods, section 6.1.12), which also detected svtRNA1-1/2 but not svtRNA2-1 bound by the Argonaute family (**Figure 22a**). The observation of svtRNA2-1 absence in Argonaute proteins is in line with previously published results which demonstrated that svtRNA2-1 from the 5' end of vtRNA2-1 did not preferentially associate with the Argonaute proteins (DU145 cell line) (Fort et al. 2020). miR-21 was used as a positive control.

It has been shown that AGO2 can carry out its function in both the nucleus and the cytoplasm (Gong et al. 2021; Sarshad et al. 2018; Kalantari et al. 2016b). Furthermore, tRF-mediated NRS functions in the nucleus (Di Fazio et al. 2022). Therefore, I further explored the possibility that svtRNA1-2 might be present in a nuclear functional complex by analysing RNA bound to cytoplasmic and nuclear AGO2 in MCF-7 cells (GSE66665, **Table 1**). Surprisingly, the svtRNA1-2 fragment was enriched in both nuclear and cytoplasmic AGO2 fractions (**Figure 22b**).

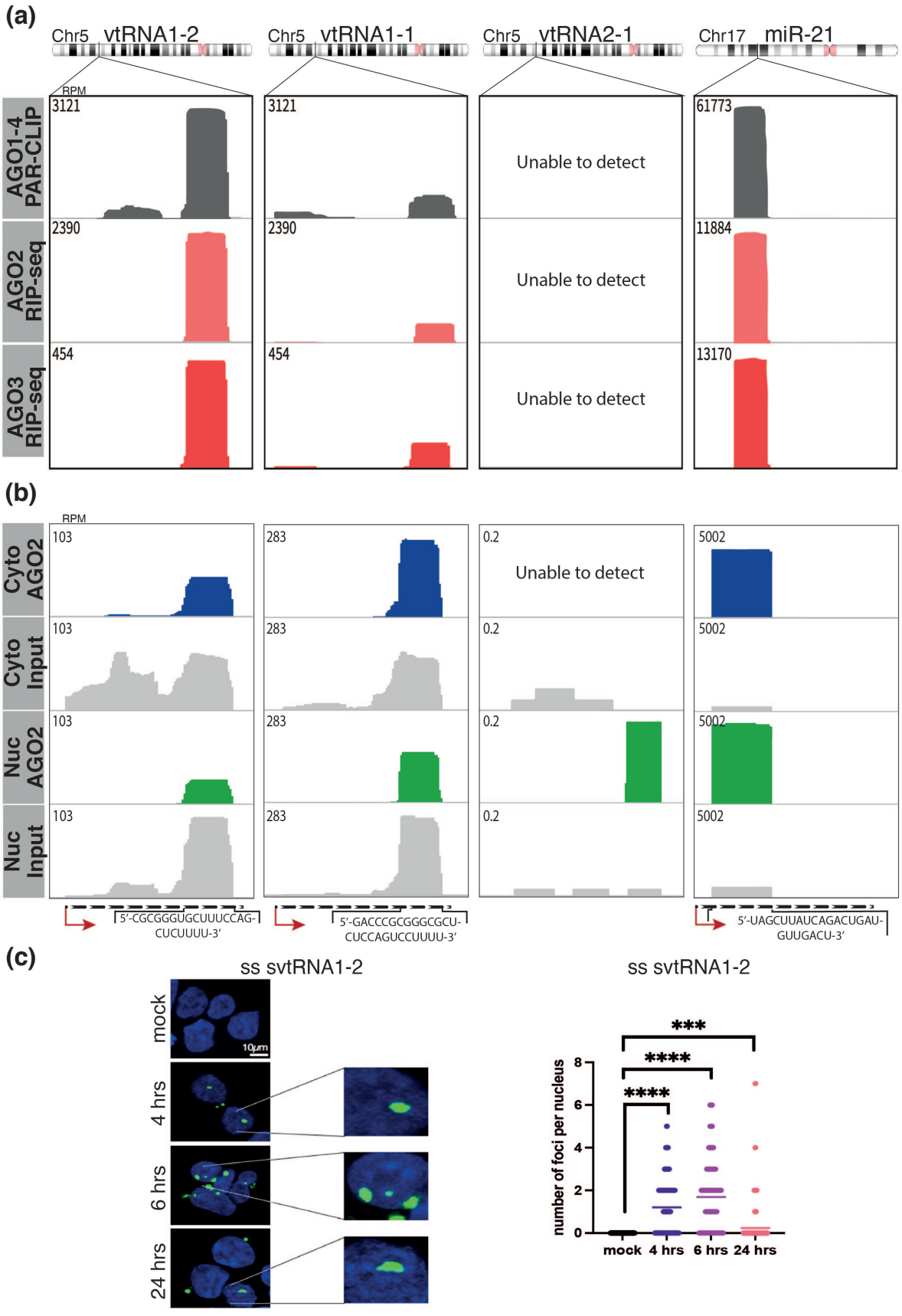


Figure 22: Caption on next page

Figure 22: (previous page) ***(a)*** *The IGV tracks depicting normalised AGO1-4 PAR-CLIP (HEK293 cell line) and AGO2/3 RIP (HEK293 cell line) sequencing small RNA reads aligned to vtRNA1-1/2, vtRNA2-1 and miR-21. Chromosome ideograms are used to depict locations of selected vtRNA and microRNA genes. The y-axis shows the top range of normalised read count as RPM. Note: no reads were detected for vtRNA2-1.* ***(b)*** *IGV tracks showing nuclear and cytoplasmic AGO2 RIP-seq (MCF-7 cell line) reads mapped to vtRNA and miR-21. Description same as* ***(a)****. The bottom black line indicates gene annotation and the arrows indicating the 5' to 3' orientation of the encoded RNA. The identified small RNA fragment parts are expanded below representing the interacting sequences.* ***(c)*** *Confocal microscopy images show nuclear localisation of single-stranded (ss) and double-stranded (green) at different time points. DAPI was used to stain the nuclei of the cells. Images were quantified using Cell profiler software (Stirling et al. 2021). Significance was calculated using t-test (****p-value < 0.0001; ***p-value < 0.001) by Dr Alagia.*

To experimentally validate these observations, HEK293T cells were transfected with fluorescently labelled single and double-stranded svtRNA1-2 and its cellular localisation was followed at 4, 6, and 24 hours post-transfection. I observed primarily nuclear localisation of single-stranded (**Figure 22c**) and double-stranded (**Supplementary Figure 38**) svtRNA1-2 at all time points tested. Taken together, these data show that svtRNA1-2 fragments bind to AGO2 in both the nucleus and cytoplasm, suggesting that they might play a role in the regulation of nuclear processes.

2.2.4 svtRNA1-2 targets genes on nascent RNA level

The Dicer dependency and association of svtRNA1-2 with AGO2, along with its nuclear localisation, prompted the idea to investigate whether svtRNA1-2 and not the rest of the identified svtRNA could regulate the gene expression on a nascent RNA level, similar to tRF. In order to explore this hypothesis, I first needed to characterise the nascent transcriptome of cells in WT, Dicer and AGO2 KD conditions (GSE126751, **Table 1**). I used the chromatin-associated RNA sequencing (chrRNA-seq) datasets, which provide information about nascent transcripts attached to chromatin by RNA polymerase II and other RNA bound to chromatin, such as lncRNA (**Figure 23a**).

The reads quantification was performed on the gene level and not exon level to capture reads mapped to the intronic regions (see Methods, section 6.1.16). The differential expression analysis revealed that AGO2 and Dicer KD had a broad effect on the transcriptome,

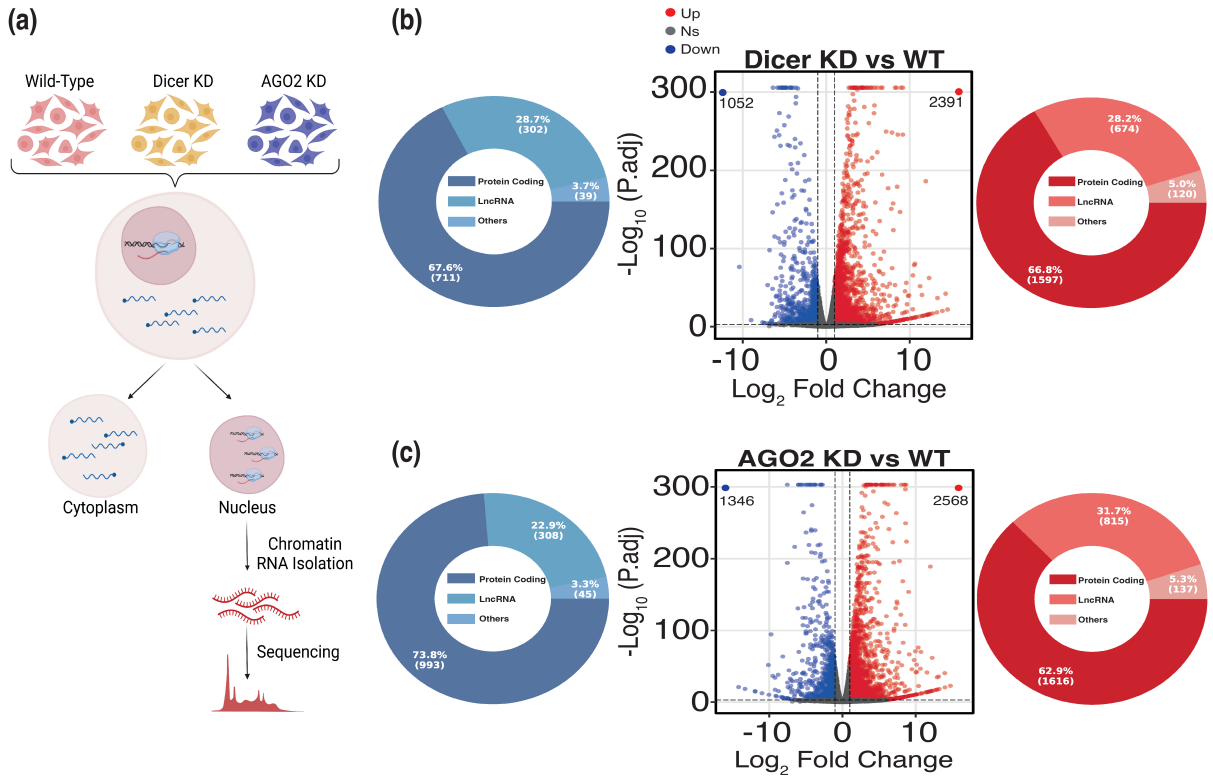


Figure 23: ***(a)*** *Schematic representation of chrRNA-seq method using HEK293 cell line. Samples prepared by Dr Pong.* ***(b)*** *The volcano plot showing the differential expression of genes in Dicer KD vs WT plotted against the -Log$_{10}$ of the adj. p-values. The red points show up-regulated genes and blue points the down-regulated genes with adj. p-value < 0.001 and Log$_2$FC $> 1 \mid < -1$, respectively. Grey points do not have significant values (Ns). Doughnut charts show different gene types of up (red) and down (blue) regulated genes.* ***(c)*** *As in* ***(b)*** *using AGO2 KD samples.*

with 1052 genes significantly down-regulated and 2391 genes significantly up-regulated in Dicer KD (**Figure 23b**). Similarly, the Knock Down of AGO2 led to significant down-regulation of 1346 and upregulation of 2568 genes (**Figure 23c**) (adj. p-value < 0.001 and Log$_2$FC $> 1 \mid < -1$). After the gene annotation, protein-coding genes were the most prevalent gene type among the differentially expressed genes.

I implemented a custom computational analysis pipeline to identify the role of svtRNA1-2 in gene silencing (**Supplementary Figure 39**). From the significantly differentially expressed genes, I considered only the up-regulated protein-coding genes from the intersection of both AGO2 and Dicer KD (1272 genes), as I showed that svtRNA1-2 is Dicer-dependent and binds to AGO2 (**Figure 24a**). Then, I employed well-known and

widely used miRanda software (Enright et al. 2003) (Deng et al. 2022; X. Y. Su et al. 2020; Dey et al. 2022) for *in silico* prediction of svtRNA1-2 gene targets through sequence complementarity, similar to miRNA binding sites, with restricted seed length and binding energy (see Methods, section 6.1.21).

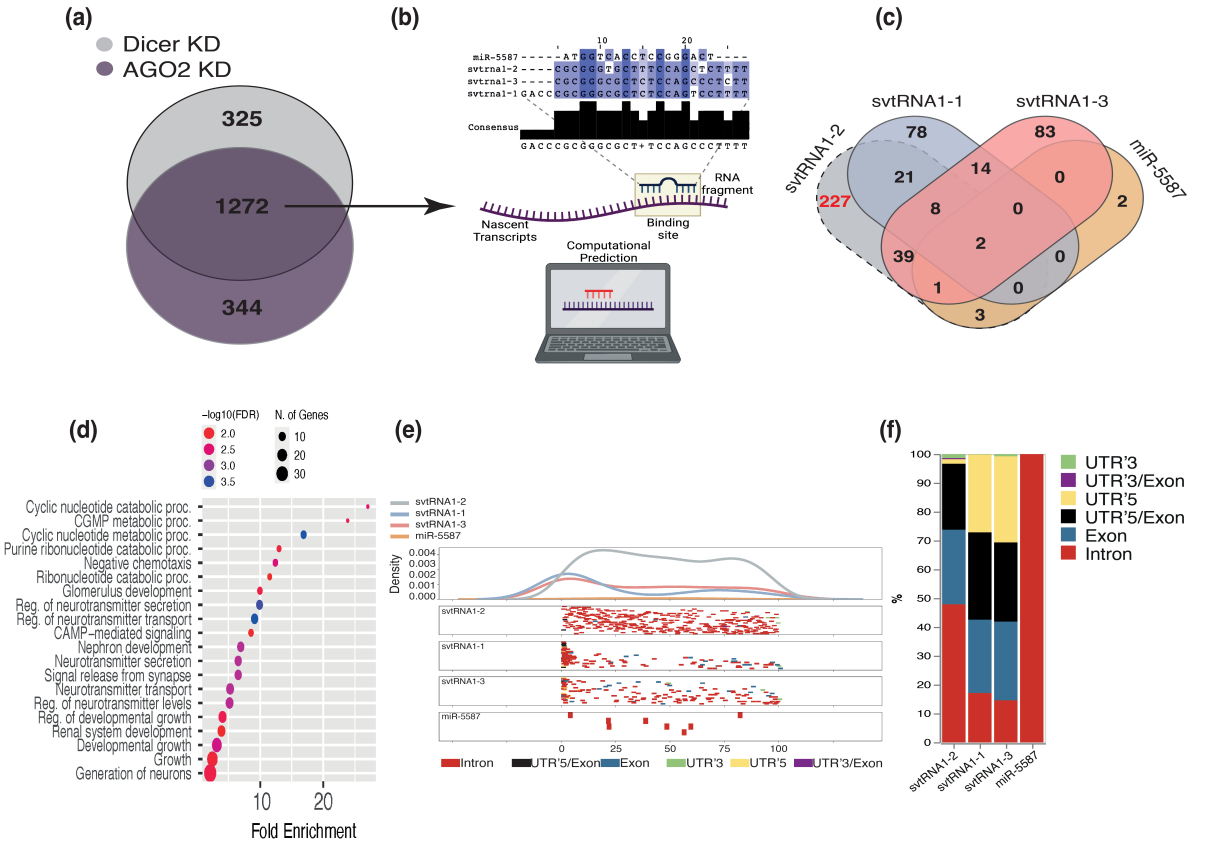


Figure 24: *(a) Venn diagram showing the overlap of up-regulated protein-coding genes between Dicer and AGO2 KD conditions.* ***(b)*** *Diagram representing the computational method for svtRNA binding targets prediction.* ***(c)*** *Venn diagram showing the number of predicted target genes for svtRNA1-1/2/3 and miR-5587. svtRNA1-1/3 and miR-5587 were used as controls as their sequences have the highest similarity with svtRNA1-2.* ***(d)*** *The scatter plot show GO enriched terms for 227 genes that were predicted as unique targets for svtRNA1-2 in both AGO2 and Dicer KDs.* ***(e)*** *Distribution of predicted svtRNA binding sites mapped along normalised target transcripts length and their corresponding gene features.* ***(f)*** *Stacked bar graphs showing the percentage of predicted binding sites for the indicated small RNA fragments mapped to different categories of genomic features. UTR'3/Exon and UTR'5/Exon represent an overlap of the predicted mapped sites on both features (see Methods, section 6.1.22).*

I used 3' end sequence fragments of vtRNA1-1 and vtRNA1-3 as controls because their sequences differ only by a few bases from svtRNA1-2 (**Figure 24b**). miR-5587 was used

as a negative control, as it has a similar sequence to svtRNA1-2 and can also control the role of miRNA in NRS.

Interestingly, I found that svtRNA1-2 had the highest number of predicted unique gene targets (227 genes) when compared to svtRNA1-1 (78 genes) or svtRNA1-3 (83 genes) (**Figure 24c**). Gene ontology analysis of the predicted svtRNA1-2 target genes revealed enrichment for pathways involved in cell growth, development and motility (**Figure 24d**) (**Supplementary Table 7.2**). Among the top 10 predicted target genes with the highest Log_2Fold Change was transmembrane receptor tyrosine kinase ERBB4 (**Supplementary Table 3**).

ERBB4 plays an essential role as cell surface receptor for neuregulins and EGF family members. It can regulate various processes such as gene transcription, cell proliferation, differentiation, migration and apoptosis. To investigate the svtRNA predicted binding regions in more detail, I developed a custom transcript annotation file which contained strand-dependent coordinates of selected transcripts and their features. This allowed me to explore the distribution of the binding sites along the transcript body as well as to identify the precise locations in different features such as intron, exon and UTRs. Interestingly, the predicted bindings of svtRNA1-2 displayed a distinct distribution pattern along the transcript body. In contrast, the other svtRNA fragments were mapped mainly to the beginning of the transcripts (**Figure 24e**). This could potentially suggest different roles for svtRNA1-2 and 1-1. Furthermore, I observed that svtRNA1-2 mapped to intronic regions with the highest occurrence suggesting a potential role of svtRNA1-2 in NRS (**Figure 24f**). In summary, I identified genes that are up-regulated upon Dicer and AGO2 KD on a nascent level and are *in silico* predicted to be targeted by svtRNA1-2, primarily within introns.

3 The functional role of small vault RNA1-2

The experimental work has been performed as follows: vtRNA1-2 CRISPR-Cas12a KO, Colony formation assay, MTS assay and vtRNA1-2 KO cell adhesion microscopy were performed by Dr Adele Alagia. The vtRNA1-2 KO chrRNA-seq sample preparation and MCF7 Colony formation assay were done by Ruth F. Ketley. The western blot chrRNA-seq fraction validation, qRT-PCR validations, vtRNA1-2 KO colony screening and counting of colonies using ImageJ were performed by myself. The total RNA-seq sample preparation was carried out partly by Dr Adele Alagia and myself. All the bioinformatics analyses have been done by myself apart from Figure 29h which was done by Dr Irina Chelysheva. I have driven the analyses of all the sequencing datasets.

3.1 Introduction

Small non-coding RNA, as well as their derived fragments, can function as regulatory molecules in different cellular processes. For instance, the canonical role of tRNA is the translation of messenger RNA into proteins; however, tRNA-derived fragments have other functions. Similarly to microRNA, tRF have also been found associated with Argonaute proteins and repressing the expression of genes via complementary binding to the target mRNA in the cytoplasm (e.g. tRF-3027b regulation of RPA1 expression (Maute et al. 2013)). Furthermore, Di Fazio et al. 2022 showed a subset of nuclear AGO2 interacting tRF, leading to decreased expression of their target gene (e.g. SPINT1) on a nascent transcript level. This transcript was silenced by the sequence complementarity to the target intronic region, demonstrating a novel functional mechanism. Out of the four vault RNA, vtRNA1-1 is the most researched one. Similarly to tRNA fragments, vtRNA1-1 fragments were shown to regulate the expression of genes involved in drug metabolism (Persson et al. 2009) and voltage-gated calcium channels (Hussain et al. 2013) via PTGS in the cytoplasm; however, no further investigations were performed to determine svtRNA1-1 localisation in different cellular compartments. Moreover, several research groups studied the intact

vtRNA1-1, demonstrating its functional involvement in cellular processes (e.g. apoptosis (Ferro et al. 2022), autophagy (Horos et al. 2019) and cell epidermal differentiation (Sajini et al. 2019). These results highlight the importance of vault RNA in regulating essential cellular processes, yet half of the vault RNA group (vtRNA1-2 and vtRNA1-3) remains uncharacterized. In the previous chapter, I showed that out of the four human vtRNA only vault RNA1-2 showed Dicer-dependent processing and subsequent AGO2 fragment loading in HEK293 cells. Moreover, the fragment derived from vtRNA1-2 is a newly identified small vault RNA with an unknown function. In this chapter, I aim to investigate the functional role of small vault RNA1-2 fragments using experimental approaches and computational analyses of datasets generated from vtRNA1-2 KO cells. First, the HEK293T-based vtRNA1-2 KO cell line was generated using the CRISPR-Cas12a system. Then the computational analysis of the nascent RNA transcriptome of vtRNA1-2 KO cells confirmed a subset of predicted svtRNA1-2 targets with downregulated ERBB4 expression, one of the previously identified targets, which was subsequently experimentally validated. Similarly to tRF nuclear nascent RNA silencing, svtRNA1-2 equally regulated the expression levels of target gene (ERBB4) on nascent and steady-state levels. Interestingly, pairing between svtRNA1-2 and target nascent RNA displayed critical complementarity requirements for both seed and 3' supplementary svtRNA1-2 regions with the predicted targets. Moreover, the Knock Out of vtRNA1-2 caused a decrease in cell proliferation, further indicating the essential role of vtRNA1-2 in the cell. These results provide important functional insights into the role of vtRNA1-2 and svtRNA1-2 in the cell and the basis for further research.

3.2 Results

3.2.1 vtRNA1-2 Knock Out leads to defects in cellular proliferation

In the previous chapter, I analysed chromatin-bound RNA sequencing datasets generated by Knocking Down important RNAi factors to study the svtRNA potential targets expression in the absence of essential regulatory molecules. To further investigate the svtRNA1-2 molecular function, RNA-guided CRISPR-Cas12a system targeting vtRNA1-2 with a single guide RNA (sgRNA) in HEK293T cells was employed in order to remove

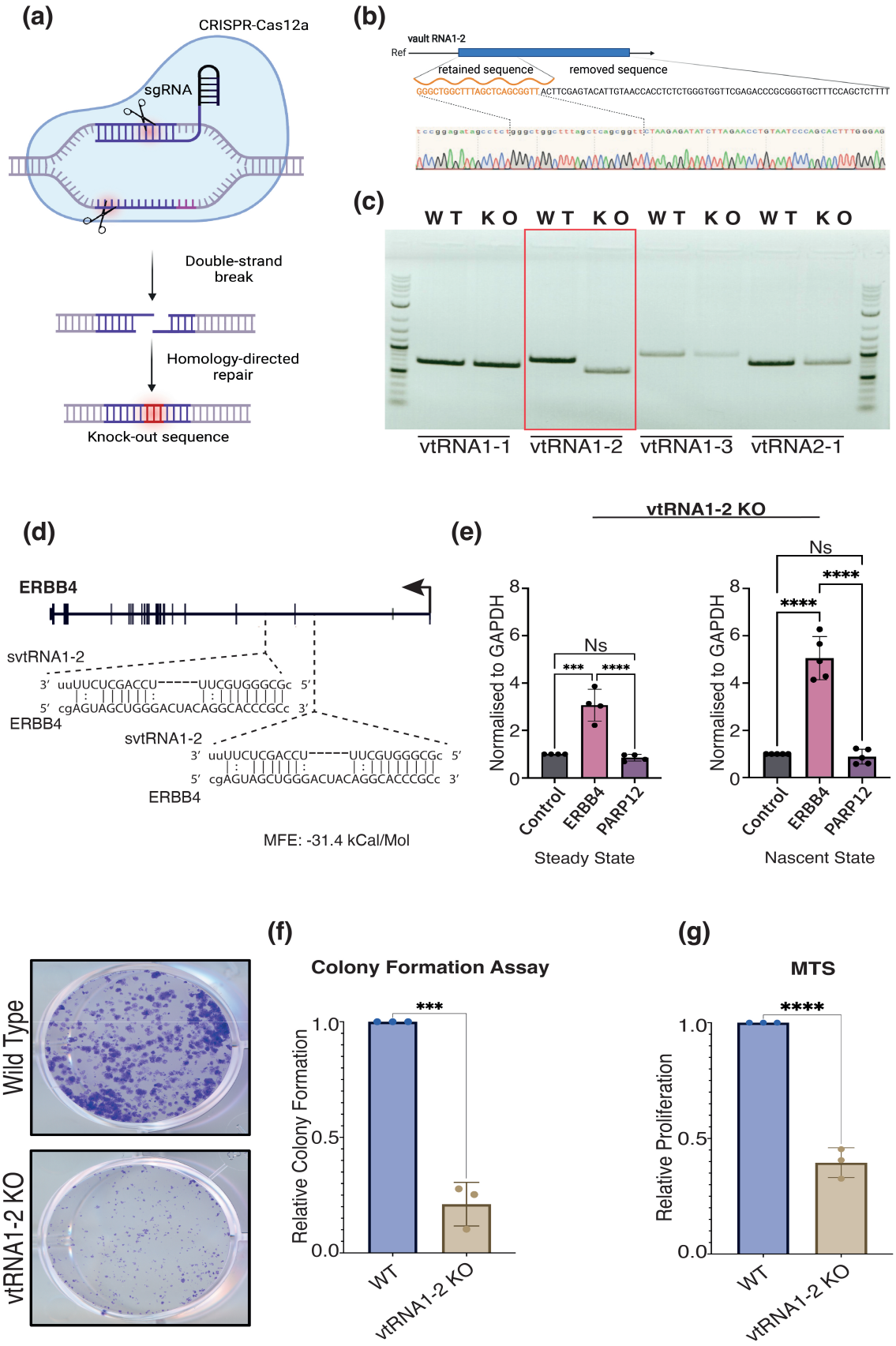


Figure 25: Caption on next page

Figure 25: (previous page) ***(a)*** *Schematic representation of CRISPR-Cas12a vtRNA1-2 KO strategy.* ***(b)*** *Sanger sequencing results confirm vtRNA1-2 Knock Out in HEK293 cells. The retained sequence is highlighted in orange; the removed part is without the highlight.* ***(c)*** *PCR analysis from purified genomic DNA from WT and vtRNA1-2 KO showing each vault RNA. The red rectangular frames the results of vtRNA1-2 confirming the Knock Out.* ***(d)*** *Representation of svtRNA1-2 WT predicted binding to ERBB4 at two intronic locations (miRanda output). The Minimal Free Energy of the predicted binding was calculated to be -31.4 kCal/Mol for each binding.* ***(e)*** *qRT-PCR quantification of ERBB4 as a predicted target of svtRNA1-2 and PARP12 as a predicted target of svtRNA1-1 (act as control) in svtRNA1-2 KO cells normalised to GAPDH. The data represent the mean fold changes from three independent experiments (****p-value < 0.0001; ***p-value < 0.001; n.s not significant).* ***(f)*** *Clonogenic assay of vtRNA1-2 KO and WT cells. The cells were stained and counted after 10 days of growing (***p-value < 0.001).* ***(g)*** *MTS assay showing proliferation of WT and vtRNA1-2 KO cells after 7 days (****p-value < 0.0001). Clonogenic and MTS assay performed by Dr Alagia.*

its genomic sequence (**Figure 25a**). Individual vtRNA1-2 Knock Out (KO) clones, which contained a partial deletion of vtRNA1-2, including the region containing the 3' end fragment, were successfully generated and confirmed by Sanger sequencing (**Figure 25b**) and PCR (KO results in a shorter genomic fragment). As vault RNA loci show relatively high sequence similarity, it was important to test whether the other vtRNA loci remained intact. Indeed, vtRNA1-1, 1-3 and 2-1 genomic loci were not affected by vtRNA1-2 deletion (**Figure 25c**).

Once the HEK293T-based cell line with the deleted svtRNA1-2 region was generated, ERBB4 as one of the top 10 most significantly differentially expressed genes in Dicer and AGO2 KDs vs WT with two putative binding sites in intronic regions was chosen to investigate the interactions with svtRNA1-2 (**Figure 25d**). I performed qRT-PCR to validate the effect of the vtRNA1-2 KO on the selected predicted targets (**Figure 25e**). I measured the expression levels of ERBB4, a unique target for svtRNA1-2 and PARP12, a unique target for svtRNA1-1, was used as a negative control.

I found that the ERBB4 expression was significantly increased at both the nascent and steady-state RNA levels in vtRNA1-2 KO cells, whilst the expression of PARP12 remained unaffected, confirming that vtRNA1-2 or svtRNA1-2 modulate ERBB4 expression levels (**Figure 25e**). Interestingly, I observed a severe growth defect in the vtRNA1-2 KO cells compared to WT cells, as shown by the clonogenic survival assay and a colourimetric

MTS proliferation assay (**Figures 25f and g**). The decreased cell viability and reduced proliferation of vtRNA1-2 KO cells suggest that vtRNA1-2 plays an important role in normal cell function.

3.2.2 Deletion of vtRNA1-2 alters the expression of genes associated with cellular membrane function

The validated target of svtRNA1-2 (ERBB4) displayed expression changes at the nascent and steady-state levels upon vtRNA1-2 KO. In the effort to fully explore the role of svtRNA1-2 in nascent RNA gene silencing, the nascent RNA was isolated from HEK293T WT and vtRNA1-2 KO cells in triplicates and chromatin-associated RNA sequencing was performed with paired-end 150bp libraries (**Figure 26a**) (GSE198298, **Table 1**). The purity of the fractions was confirmed by western blotting using Histone 3 (H3) as a chromatin marker and Tubulin as a cytoplasmic marker (**Figure 26b**). On average, the sequenced samples contained 103 million reads (**Figure 26c**). I used the STAR aligner (Dobin et al. 2013) to map them to the human reference genome (hg38). The majority of the reads mapped uniquely to the genome, with only a few reads that mapped to too many loci or did not map at all (**Figure 26d**). A principal component analysis outlined clear differences between the two conditions and tight clustering of the replicates (**Figure 26e**). The mapped reads were quantified using featureCounts (Liao et al. 2014), and the differential expression analysis was performed using the DESeq2 (Love et al. 2014) R package (see Methods, section 6.2.7). I found 547 significantly down-regulated and 944 significantly up-regulated (adj. p-value < 0.001 and $Log_2FC > 1 \mid < -1$), mostly protein-coding genes in vtRNA1-2 KO cells (**Figures 27a and c**). Closer inspection showed that ERBB4 was among the up-regulated genes in vtRNA1-2 KO (adj. p-value=$1.2x10^{-25}$ and Log_2FC 1.4), providing further validation for svtRNA1-2 mediated downregulation of ERBB4 (**Figure 27b**).

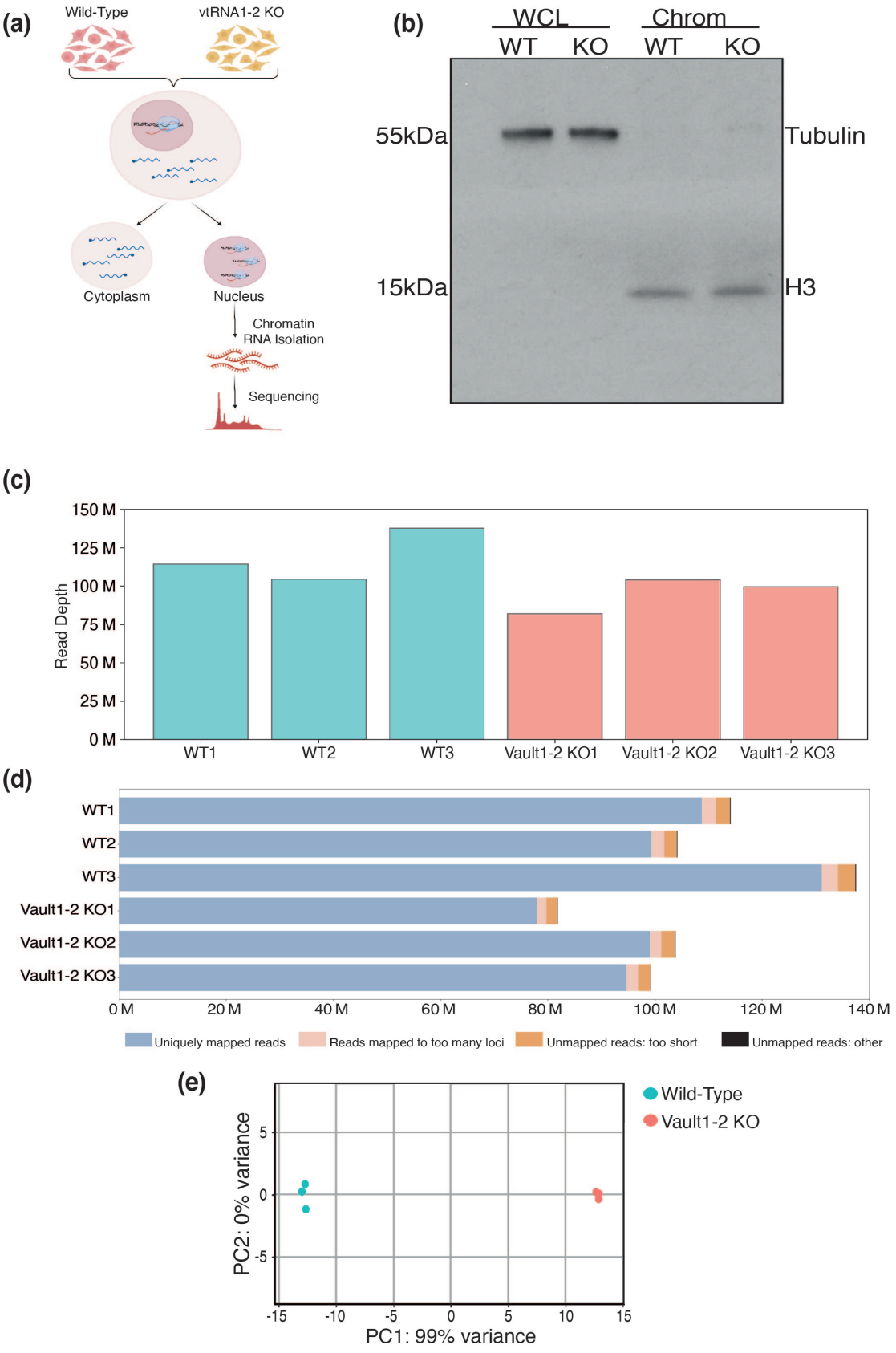


Figure 26: Caption on next page

Figure 26: (previous page) ***(a)*** *Schematic representation of chrRNA-seq workflow in HEK293 cell line. Sequencing samples prepared by Ruth F. Ketley.* ***(b)*** *Western blot confirming the chromatin and whole cell lysate (WCL) fractions components isolated for chrRNA sequencing.* ***(c)*** *Bar chart showing the number of reads sequenced for each sample.* ***(d)*** *A Total number of reads mapped to the human genome (hg38) per sample using STAR aligner (Dobin et al. 2013).* ***(e)*** *Principal component analysis plot shows three biological replicates corresponding to the samples from HEK293T cells in WT and vtRNA1-2 KO conditions. The RNA composition of cells in the WT condition differed significantly from those in the vtRNA1-2 KO condition.*

To better understand the functional relevance of the significantly differentially expressed genes, I performed a Gene Ontology (GO) enrichment analysis (**Figure 27d**), which revealed a strong association of up-regulated protein-coding genes with biological processes such as *Regulation of membrane transport and potential* (GO:0043269 and GO:0042391, False Discovery Rate (FDR) = $3.9\text{x}10^{-12}$ and FDR= 6.7x^{-14}, respectively). Surprisingly, out of the top 10 enriched biological processes, 9 were found to influence either transport processes or membrane potential (**Supplementary Table 4**), while down-regulated genes were not enriched in these categories (**Supplementary Table 5**). To identify novel paths and gain a global view of the transcriptome profile regulation connected to the vtRNA1-2 KO, I performed Gene Set Enrichment Analysis. Interestingly, the deletion of vtRNA1-2 showed significant enrichment of the Early and Late estrogen response gene set (NES=2, p.val < 0.001 and NES=1.65, p.val < 0.001), most likely due to deregulated ERBB4 (**Figure 27e**). Consistent with this interpretation, the KEGG pathway gene set analysis confirmed an altered *ERBB4 signalling pathway* (NES=1.77, p.val < 0.001) (**Figure 27f**). These data propose that vtRNA1-2 is, directly and indirectly, involved in regulating membrane proteins through svtRNA1-2 silencing abilities.

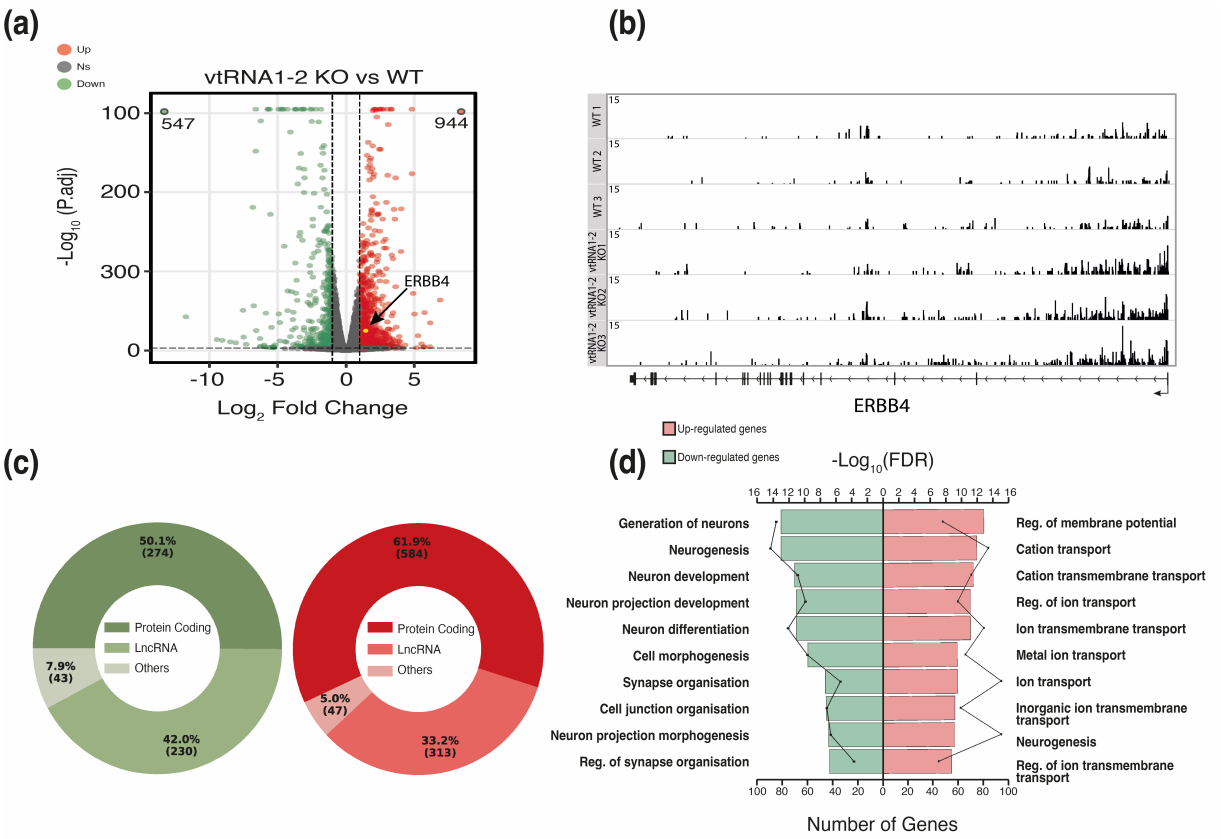


Figure 27: *(a) Volcano plot showing differentially expressed genes between vtRNA1-2 KO [n=3] and WT [n=3]. Analysis revealed 944 up-regulated and 547 down-regulated genes in vtRNA1-2 KO condition compared to WT. Significantly up are indicated in red and green points indicate down-regulated genes (adj. p-value < 0.001 and Log_2FC > 1/ < -1). Grey points do not have significant values. The yellow point shows the location of ERBB4.* ***(b)*** *Genome browser view of chromatin-associated RNA sequencing signals across the ERBB4 locus for HEK293T WT and vtRNA1-2 KO cells.* ***(c)*** *Doughnut charts show different gene types of up (red) and down- (green) regulated genes.* ***(d)*** *GO enrichment analysis of significantly differentially expressed genes in vtRNA1-2 KO vs Wild Type. The green bars represent the fold enrichment of each term for down-regulated genes and the red bars for the up-regulated genes (top x-axis). The black dotted line represents the number of genes annotated in each GO term (bottom x-axis).* ***(e)*** *Gene Set Enrichment plots show the upregulation of the early and late estrogen response pathways (NES=2, p-value < 0.001 and NES=1.45, p-value < 0.001, respectively).* ***(f)*** *GSEA using KEGG pathways shows upregulation of ERBB signalling pathway upon vtRNA1-2 KO (NES=1.77, p-value < 0.001).*

3.2.3 svtRNA1-2 target genes are associated with cellular membrane function

To further investigate svtRNA1-2 targets, I used the miRanda prediction tool to search for target binding sites with svtRNA sequences along significantly up-regulated protein-coding genes in the vtRNA1-2 KO condition. Interestingly, I observed that the svtRNA1-2 fragment had the highest number of uniquely predicted gene targets, a similar prediction pattern to the previous target search (**Figure 28a**). Close inspection of miRanda prediction revealed 139 target genes in total and 205 mapping sites for svtRNA1-2 (**Figures 28b and c**), indicating that transcripts might contain single or multiple predicted binding sites. Next, I explored the locations of the binding sites using metagene alignment, which showed an even distribution of svtRNA1-2 target sites along the gene length, whilst svtRNA1-1 and 1-3 target sites were accumulated at the beginning of the genes, similar

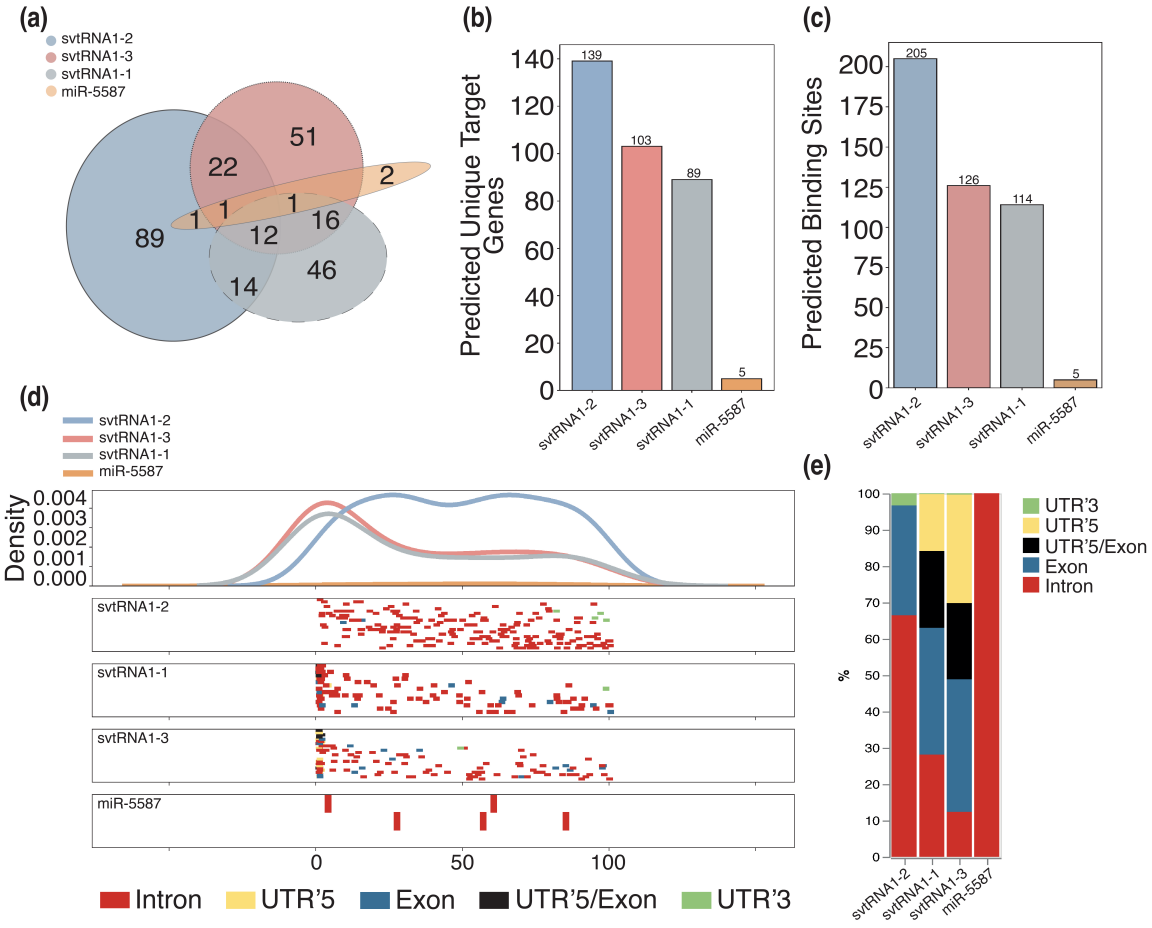


Figure 28: Capture on next page

Figure 28: (previous page)***(a)*** *Qualitative overlap of target genes predicted for each svtRNA and miR-5587.****(b)*** *Bar chart showing predicted unique target genes (left) and binding sites (right) for each svtRNA and miR-5587.****(c)*** *Distribution of predicted svtRNA binding sites mapped along normalised target transcripts length and their corresponding gene features.****(d)*** *Stacked bar graphs showing the percentage of predicted binding sites for the indicated small RNA fragments mapped to different categories of genomic features. UTR'5/Exon represents an overlap of the predicted mapped sites on both features (see Methods, section 6.1.22).*

to what I observed in **Figure 22e** (**Figure 28d**). Normalised values of target gene features confirmed that svtRNA1-2 predominantly mapped to the intronic regions. These results are consistent with the previous predictions (**Figure 28e**).

After the target prediction, I finally had all the sets of possible svtRNA1-2 targets. In order to narrow down the specific targets and exclude the false positives, I overlapped the predicted svtRNA1-2 unique targets from AGO2 and Dicer KDs (n=227) with the significantly up-regulated genes in vtRNA1-2 KO (n=584) and with predicted unique svtRNA1-2 targets from vtRNA1-2 KO up-regulated genes (n=89). I found 29 common genes which resulted from the overlap of all three sets of genes, including ERBB4 (**Figure 29a**) (**Supplementary Table 6**). The comprehensive annotation of the proteins encoded by these genes revealed that they are mostly localised in the cellular membrane (**Figure 29b**). To further resolve the functional trends of the 29 overlapped genes, I used STRING (Szklarczyk et al. 2019) to obtain an integrated pathway analysis. I identified 14 out of 29 targets to be involved in signalling and 17 out of 29 targets to be glycoproteins (**Figure 29c**). Glycoproteins are membrane proteins that play a key role in cell migration, proliferation and adhesion (Gu et al. 2012).

Three GO terms were obtained for 29 common targets *Negative regulation of homotypic cell-cell adhesion* (GO:0034111, FDR=$4.3\text{x}10^{-02}$), *Negative regulation of platelet aggregation* (GO:0090331, FDR=$4.3\text{x}10^{-02}$) and *Negative chemotaxis* (GO:0050919, FDR=$3.0\text{x}10^{-02}$). As shown in (**Figure 29d**), GO:0034111 is related to the negative regulation of biological processes such as cell-cell adhesion (GO:0022408) and cell adhesion (GO:0007162). An equal amount of WT and vtRNA1-2 KO cells was seeded to validate the result and monitor their adhesion abilities.

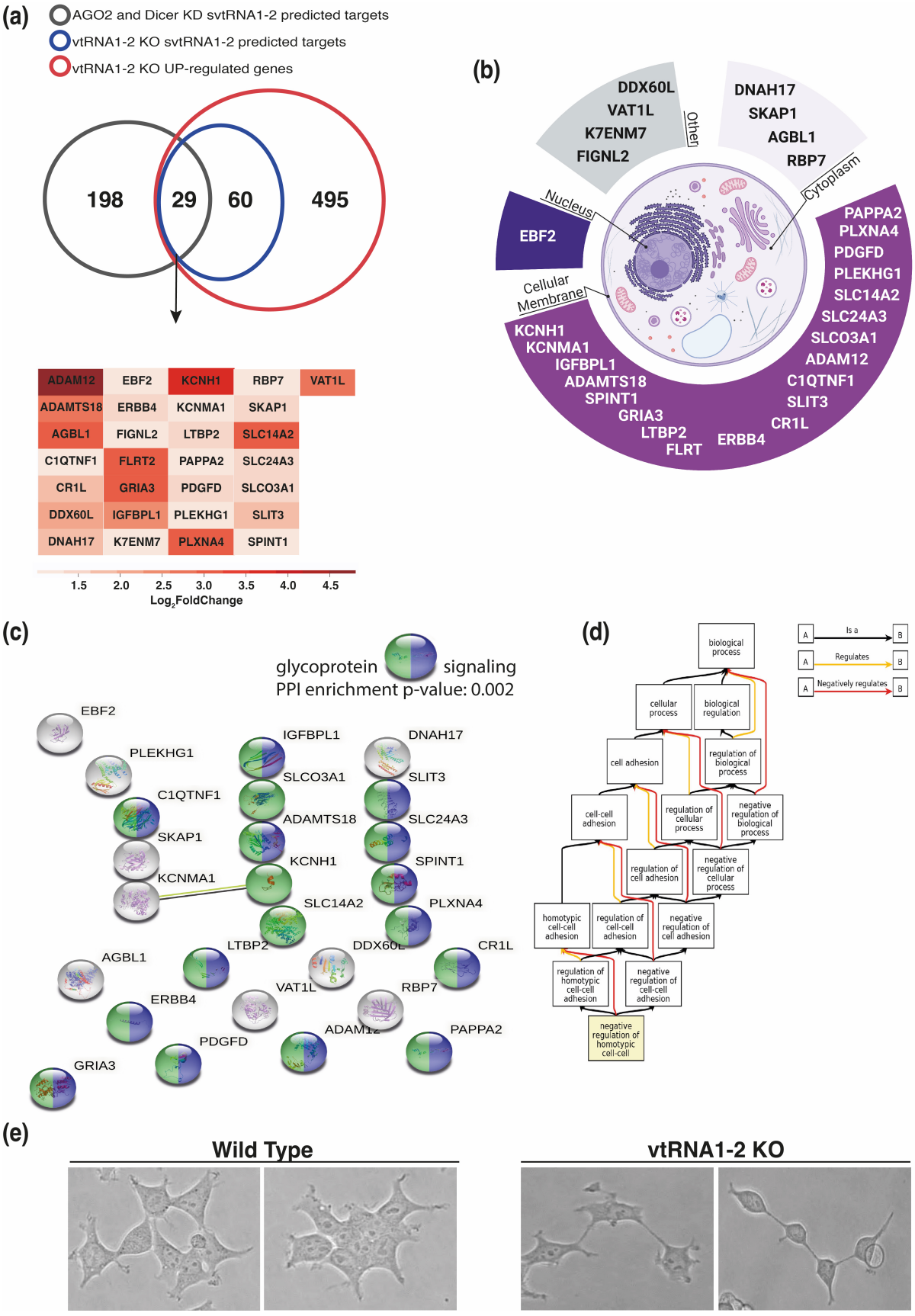


Figure 29: Caption on next page

Figure 29: *(previous page)* ***(a)****(top) Venn diagram showing the number of protein-coding genes which are identified as an overlap of up-regulated protein-coding genes from vtRNA1-2 KO (compared to WT) (red), svtRNA1-2 predicted targets in vtRNA1-2 KO (blue), and svtRNA1-2 predicted targets from up-regulated genes in AGO2 and Dicer KD conditions (grey). The table shows the names of 29 svtRNA1-2 targeted genes and their expression level (Log_2FC) in the vtRNA1-2 KO condition.* ***(b)*** *Fragmented doughnut chart representing cellular localisation of the 29 target genes.* ***(c)*** *STRING analysis showing the interactions among 29 svtRNA1-2 targets. Proteins involved in signalling are marked in blue, and glycoproteins are marked in green.* ***(d)*** *Ancestor chart view of the GO term enriched for the 29 targets showing the association of svtRNA1-2 target genes with negative regulation of cell adhesion.* ***(e)*** *Representative examples of bright field microscopy showing differences in the cell aggregation phenotype in WT and vtRNA1-2 KO cells. Images obtained by Dr Adele Alagia.*

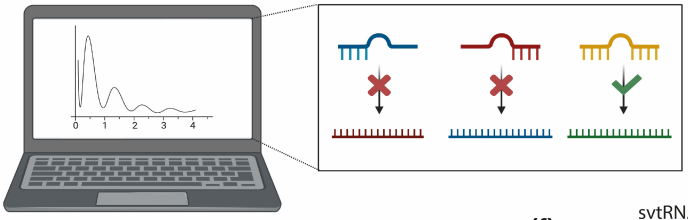
Interestingly, bright field microscopy revealed that vtRNA1-2 KO cells with reduced cell-to-cell contact exhibit elongated morphology, most likely as a potential consequence of negatively regulated cell-to-cell adhesion (**Figure 29e**). Overall, these data demonstrate that svtRNA1-2 is involved in the regulation of nascent RNA of protein-coding genes that are associated with cell membrane physiology.

3.2.4 svtRNA1-2 targets introns through dual-region complementarity and loop protrusion

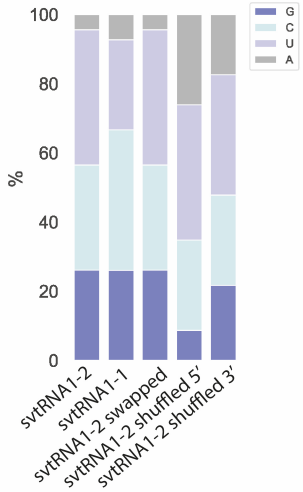
Closer inspection of miRanda predicted svtRNA1-2 targets revealed that, unlike miRNA, svtRNA1-2 maps to its targets through the 5' seed and a 3' supplementary regions (**Box 1**). This observation led to the hypothesis that complementary base pairing between svtRNA1-2 and target binding sites might be essential for the svtRNA binding and nascent target silencing. In order to verify the validity of this hypothesis, *In silico*-designed svtRNA1-2 sequences with 5' seed region shuffled (shuffled 5')(red), 3' region shuffled (shuffled 3')(red), svtRNA1-2 5' seed (highlighted in green) and 3' region swapped (svtRNA1-2 swapped) (**Figure 30a**) and svtRNA1-1 acted as negative controls, were used to scan the significantly up-regulated genes from vtRNA1-2 KO vs WT (chrRNA-seq) (see Methods, section 6.2.11). These sequences were constructed to preserve the base composition similar to the standard svtRNA1-2 to prevent a biasing artefact (**Figure 30b**). After the miRanda target scanning, I retrieved the predicted gene targets and compared them to each other (**Figure 30c**).

(a)

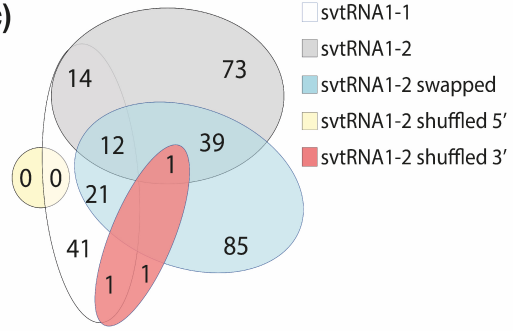
Small vault RNA	Sequence
svtRNA1-1 wild-type	3′ –UUUUCCUGACCUCUCGCGGGCGCCCAG-5′
svtRNA1-2 wild-type	3′ –UUUUCUCGACCUUUCGUGGGCGC- 5′
svtRNA1-2 swapped	3′ –UUUUCGUGGGCGUUCUCGACCUC-5′
svtRNA1-2 shuffled 5′	3′ –UUUUCUCGACCUCAAAGUUUAAC-5′
svtRNA1-2 huffled 3′	3′ –UUUUCCAUCAAAUUCGUGGGCGC-5′



(b)



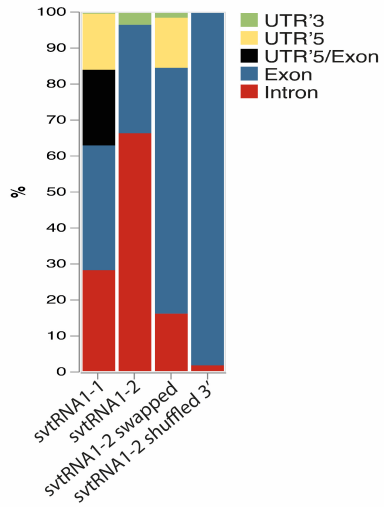
(c)



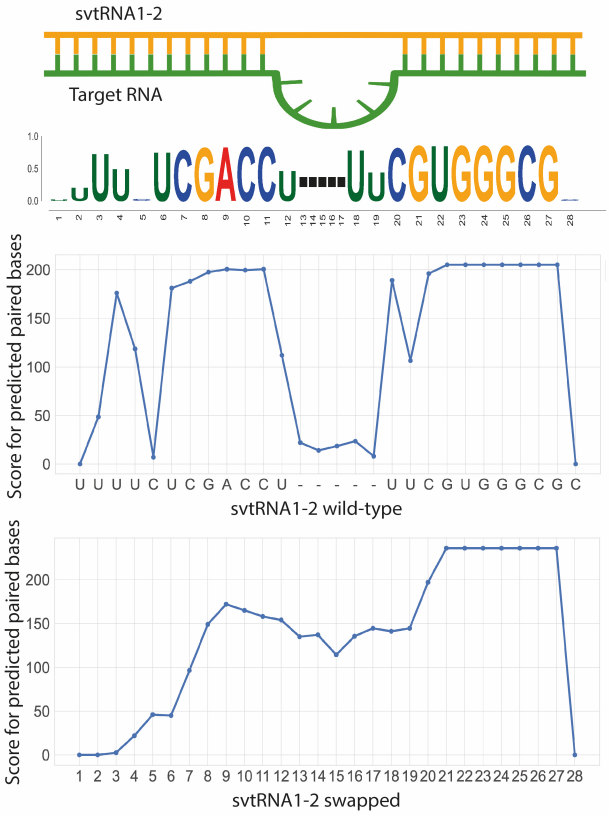
(d)

Small vault RNA	Number of Targets	Predicted Binding Sites
svtRNA1-1 wild-type	89	114
svtRNA1-2 wild-type	139	205
svtRNA1-2 swapped	158	236
svtRNA1-2 shuffled 5′	0	0
svtRNA1-2 huffled 3′	3	3

(e)



(f)



(g)

	svtRNA1-2 wild-type		svtRNA1-2 swapped	
	Length in nt (with gaps)	Number of Targets	Length in nt (with gaps)	Number of Targets
Loop	23	10	23	**98**
	24	5	24	49
	25	5	25	29
	26	2	26	18
	27	4	27	34
	28	**179**	28	8
Total		205		236

Figure 30: Caption on next page

Figure 30: *(previous page)* ***(a)*** *Sequence representation of svtRNA1-1, svtRNA1-2, svtRNA1-2 swapped (swapped regions are highlighted in green and orange, indicated with arrows), svtRNA1-2 shuffled 5' end, and svtRNA1-2 shuffled 3' end (shuffled bases have red colour). svtRNA1-1 acts as a control as its sequence has high sequence similarity to svtRNA1-2.* ***(b)*** *The bar plot represents the base composition of each sequence displayed in percentage.* ***(c)*** *Schematic representation of svtRNA target prediction search, focusing on the guide RNA seed and supplementary region complementarity.* ***(d)*** *Venn diagram showing the overlap of the target genes predicted for each sequence.* ***(e)*** *The table reporting the number of predicted targets and the number of predicted binding sites per sequence.* ***(f)*** *Stacked bar plot showing svtRNA predicted binding sites mapped to different gene features with normalised counts to the length of a feature, displayed in overall percentage.* ***(g)*** *(top) Sequence logo of the consensus svtRNA1-2 binding motif located on the target sequences.(middle) Line plot representing scores for predicted paired bases of svtRNA1-2 to their target sites. The x-axis shows the sequence of svtRNA1-2 WT, and the y-axis the number of predicted bindings.(bottom) The second line plot shows predicted paring for swapped svtRNA1-2 to their target sites. The x-axis shows the length of svtRNA1-2 swapped, and the y-axis the number of predicted bindings (bottom).* ***(h)*** *Table showing correlation analysis of the loop/no loop binding and the number of predicted target sites. 28 nt sequence complementarity of svtRNA1-2 WT contains a 5 nt long gap across the alignment with the highest number of predicted targets (highlighted in red), 23 nt sequence complementarity of svtRNA1-2 swapped represents alignment with no loop and lower number of predicted targets (highlighted in red).*

The mapping of the 5' shuffled svtRNA1-2 sequence resulted in complete loss of targets compared to the svtRNA1-2 WT sequence, whilst 3' shuffled svtRNA1-2 mapped to only two targets. The swapped svtRNA1-2 mapped to 158 targets, but these were largely different from svtRNA1-2 WT. The observed partial overlap of targets between WT and swapped svtRNA1-2 was caused by swapped svtRNA1-2 mapping to svtRNA1-2 WT targets, but at different binding sites (**Figures 30d and e**).

Interestingly, swapped svtRNA1-2 was mapping mostly to the exons, unlike svtRNA1-2 WT, which was mapping predominantly to introns (**Figure 30f**). Furthermore, the predicted svtRNA1-2 and swapped svtRNA1-2 binding diagrams from miRanda output were processed into scoring matrices that assigned values based on the complementarity (Watson–Crick–Franklin base pairing = 1, Wobble base pairing = 0.5 and no pairing = 0). The obtained findings revealed a specific target binding-motif of svtRNA1-2 WT (**Figure 30g**), which contained high frequency of GC bases at the 3' end and U-rich 5' end. Moreover, a gap between the 5' and 3' regions was observed, which created a loop structure in the target nascent RNA sequence. This loop structure was lost in targets

predicted for the swapped svtRNA1-2 (**Figure 30g**).

Finally, to test whether the loop structure associates with a certain number of targets for svtRNA1-2 WT and swapped, the length of the gaps was compared among all targets of svtRNA1-2 WT and the swapped svtRNA1-2. The original length of svtRNA1-2 without gaps in the alignment was 23 nt. In the svtRNA1-2 WT predicted bindings, the size of the gaps between seed and supplementary regions ranged from 1 up to 5 nt, establishing loop structures with variable lengths in target-svtRNA1-2 base-pairing. Importantly, most frequently observed was the longest gap (5 nt) in 179 targets out of 205, which resulted in 28 nt-long target-svtRNA1-2 binding pattern. On the contrary, the highly prevalent predicted target-swapped svtRNA1-2 binding patterns showed none or 1 nt long gap in 147 out of 236. (**Figure 30h**). Altogether these analyses suggest that svtRNA1-2 uniquely maps to intronic regions of its targets through double-region complementarity, creating a 5 nt-long loop structure in the target's nascent RNA.

3.2.5 Synthetic svtRNA1-2 downregulates ERBB4 expression in breast cancer cells

ERBB4 (HER4) activation stimulates several intracellular signalling pathways, such as RAS/MEK/ERK or PI3K/AKT, leading to increased cell proliferation and migration. Increased expression or activation of ERBB proteins is frequently observed in various cancers (Arteaga et al. 2014). Downregulation of ERBB4 has been shown to inhibit the proliferation in invasive breast cancer cells MCF7 (Okazaki et al. 2016). To further assess the regulatory effect of svtRNA1-2 on ERBB4, human breast cancer MCF7 cells were transfected with synthetic single-stranded WT svtRNA1-2, 5' shuffled svtRNA1-2 and 3' shuffled svtRNA1-2. After the transfection, a qRT-PCR analysis was performed to measure the transcription of ERBB4. The obtained values showed significantly decreased nascent RNA levels of ERBB4 in the presence of WT, but not 5' or 3' shuffled svtRNA1-2 at 24 and 48 hours post-transfection (**Figure 31a**). This result indicates that svtRNA1-2 could be employed to silence ERBB4 in MCF7 cells. In addition to these results, a colony formation assay was applied to determine whether the induced downregulation of ERBB4 by synthetic WT svtRNA1-2 might affect the proliferation of MCF7 cells.

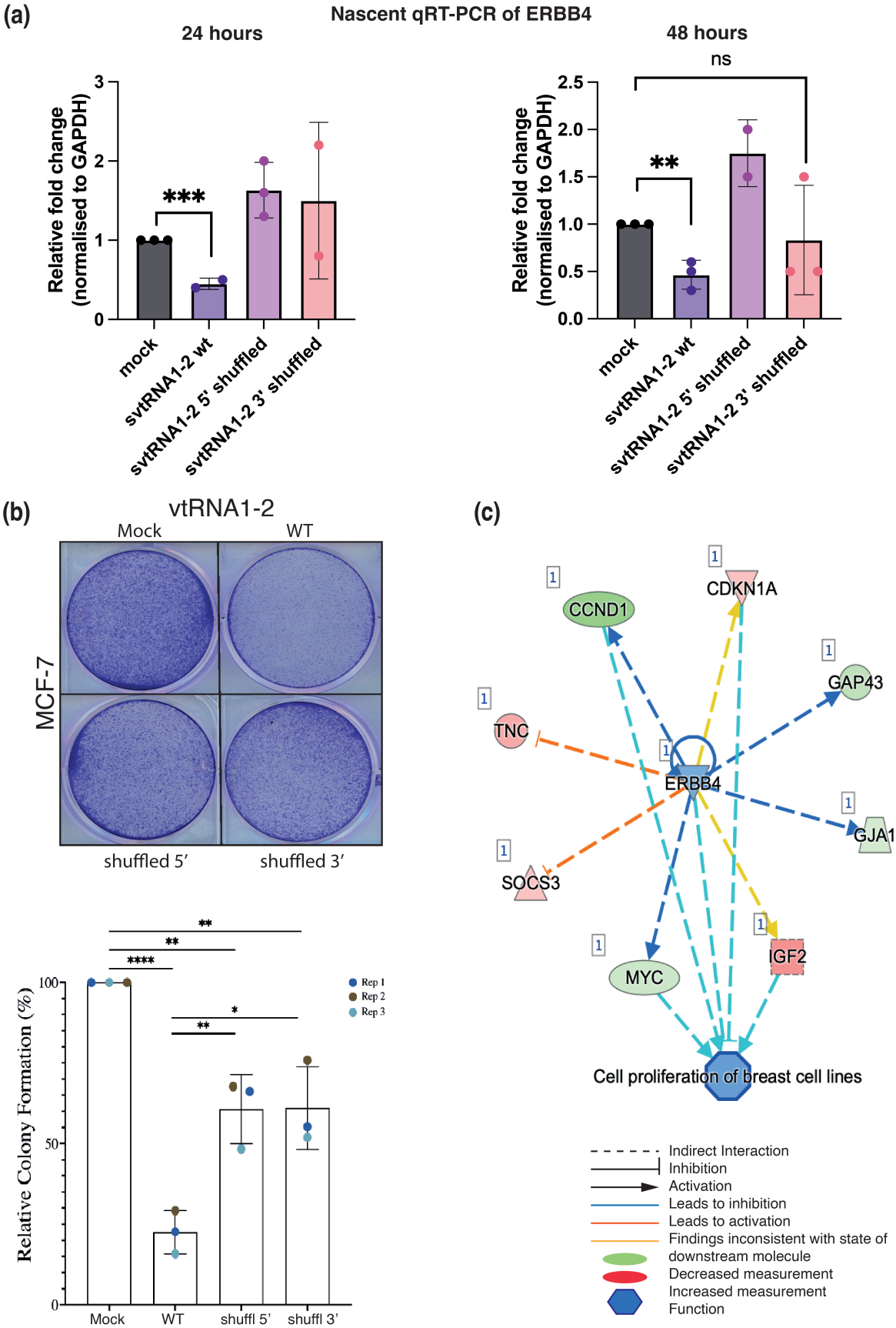


Figure 31: Caption on next page

Figure 31: (previous page) ***(a)*** *qRT-PCR quantification of nascent ERBB4 in MCF7 cells after svtRNA1-2 WT, svtRNA1-2 shuffled 3' and svtRNA shuffled 5' sequence transfection. Samples were isolated after 24 and 48 hours. The values were normalised to the GAPDH. The data represent the mean fold changes from two or three independent experiments (***p-value < 0.001; **p-value < 0.01; n.s not significant).* ***(b)*** *Clonogenic assay of MCF7 WT and cells transfected with 3' and 5' end shuffled svtRNA1-2 sequences. The cells were stained and counted after 10 days of growing (****p-value < 0.0001, ***p-value < 0.001, **p-value < 0.01) by Ruth F. Ketley.* ***(c)*** *Ingenuity pathway analysis (IPA) of differentially expressed genes showing ERBB4 in the regulation of proliferation in breast cancer cells.*

Indeed, the transfection of WT svtRNA1-2 led to an inhibition of proliferation. Furthermore, a partial proliferation phenotype was observed in cells transfected with 5', and 3' shuffled svtRNA1-2, although to a much lesser extent. This could be explained by either general toxicity or indirect effects of sRNA transfection (**Figure 31b**). To further characterise the relation between svtRNA1-2, ERBB4 and cell proliferation, I used Ingenuity pathway analysis (IPA) (Qiagen) with read counts matrix of genes identified in vtRNA1-2 KO and WT samples to perform a differential expression analysis and interpret the results within the context of biological pathways (**Figure 31c**). The provided findings further highlight the involvement of svtRNA1-2 in ERBB4 silencing with decreased cell proliferation. These data suggest that synthetic svtRNA1-2 can lead to specific downregulation of ERBB4 in MCF7, which can, in part, lead to inhibition of proliferation.

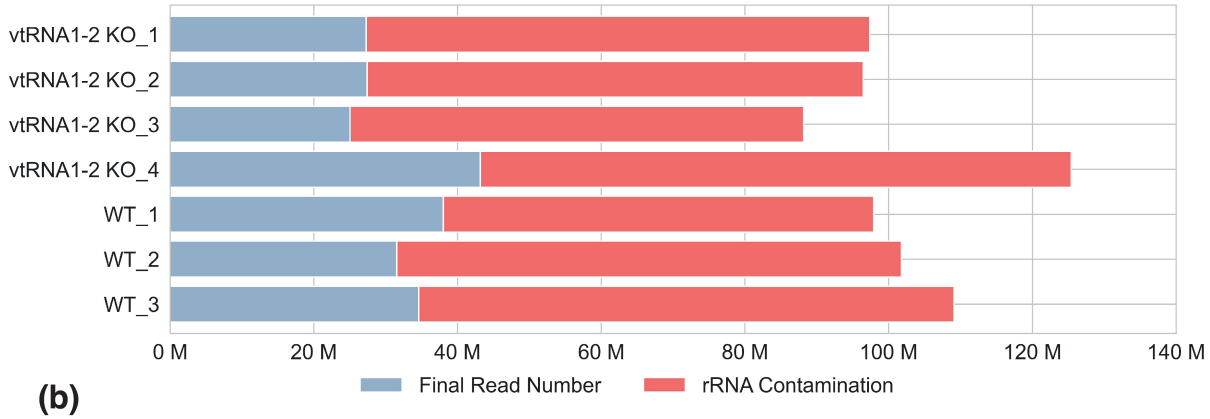
In summary, I showed that synthetic svtRNA1-2 can downregulate ERBB4 expression at nascent RNA level in breast cancer cells. Furthermore, the decreased levels of ERBB4 significantly inhibited the proliferation of MCF7 cells, which was also identified *in silico* from vtRNA1-2 KO chromatin-associated RNA sequencing. These data opens the possibility of using svtRNA fragments for therapeutic purposes.

3.2.6 Steady-state transcriptome profiling of vtRNA1-2 KO cells

The identified svtRNA1-2 nascent RNA targets have an increased expression in vtRNA1-2 KO cells, which should equally be up-regulated at the mRNA level in the cytoplasm. To verify this hypothesis, I isolated the total RNA from vtRNA1-2 KO and WT cells (HEK293T cells) and submitted the samples for paired-end 150 bp sequencing. The initial average read count was around 102 million per sample.

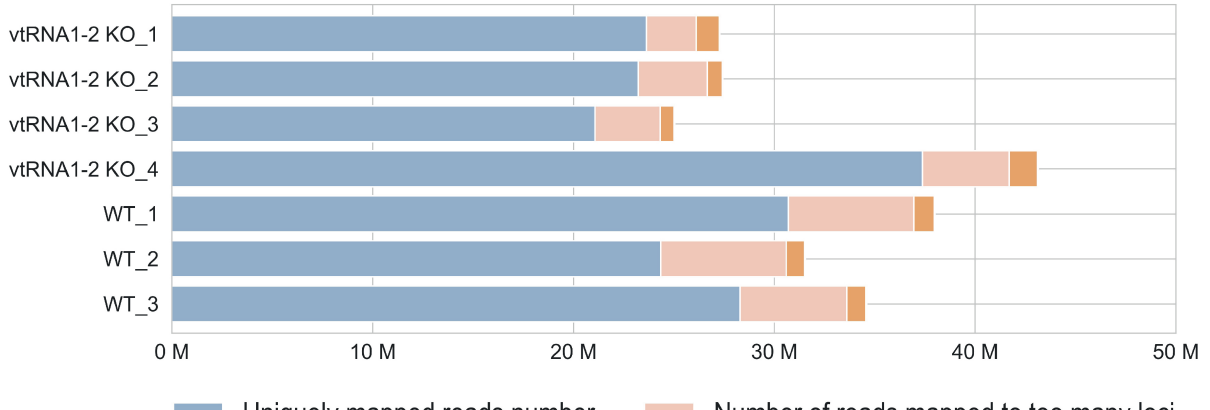
(a)

	vtRNA1-2 KO_1	vtRNA1-2 KO_2	vtRNA1-2 KO_3	vtRNA1-2 KO_4	WT_1	WT_2	WT_3
Total number of reads	97379093	96474548	88161147	125434290	97919780	101789371	109093229
rRNA contamination	71.98%	71.56%	71.60%	65.60%	61.18%	69.00%	68.28%
Number of removed reads	70090939	69033610	63121593	82279812	59911037	70238775	74490748
Final read number	27288154	27440938	25039554	43154478	38008743	31550596	34602481



(b)

	Uniquely mapped reads number	Number of reads mapped to too many loci	Number of reads unmapped: too short	Number of reads unmapped: other
vtRNA1-2 KO_1	23633338	2480164	1151821	22831
vtRNA1-2 KO_2	23222261	3438192	757305	23180
vtRNA1-2 KO_3	21074251	3243908	696311	25084
vtRNA1-2 KO_4	37384402	4319536	1406734	43806
WT_1	30695397	6251424	1026242	35680
WT_2	24352151	6242581	915026	40838
WT_3	28292171	5328091	942457	39762



(c)

	Assigned	Unassigned_NoFeatures	Unassigned_Ambiguity	Total
vtRNA1-2 KO_1	8517175	1771746	13344417	23633338
vtRNA1-2 KO_2	7456540	1692606	14073115	23222261
vtRNA1-2 KO_3	5695739	1547253	13831259	21074251
vtRNA1-2 KO_4	5417490	1886875	30080037	37384402
WT_1	14521183	2593787	13580427	30695397
WT_2	8114677	1506468	14731006	24352151
WT_3	8898906	1908427	17484838	28292171

Figure 32: Caption on next page

Figure 32: (previous page)***(a)*** *The number of reads sequenced for each sample with the subsequent filtration of a large proportion of rRNA contamination.* ***(b)*** *The number of rRNA filtered reads mapped to the human reference genome (hg38).* ***(c)*** *Table summarising final gene-assigned reads counts (column highlighted in green).*

The adapter sequences were removed, and the processed reads were aligned to the human reference genome (hg38). However, the alignment yield of uniquely mapped reads was, on average, only 40%. Further investigation revealed the cause of the suboptimal alignment results, which was ineffective sample rRNA depletion. After filtering the rRNA contaminant, the average sample read count was around 32 million per sample, which meant a loss of approximately 68% of reads (**Figure 32a**). The remaining amount of reads is still sufficient to investigate the differential gene expression. Therefore, I performed the second read alignment using the filtered reads, which resulted in approximately 83% uniquely mapped reads, indicating that rRNA were the reason for the suboptimal alignment (**Figure 32b**). The uniquely mapped reads were further quantified using featureCounts (Liao et al. 2014), which assigned the reads based on their genomic coordinates to the exonic gene features. Despite the optimal alignment results, the number of assigned reads was, on average, 8 million per sample, around 31%, with most reads failing to assign due to the read overlapping two or more features (**Figure 32c**). Nevertheless, to detect only highly expressed genes, I performed the differential expression analysis and found only 564 significantly up-regulated and 328 significantly down-regulated genes (adj. p-value < 0.05 and FC $> 1 \mid < -1$ (**Figure 33a**)).

This outcome was anticipated, as the entered read count had low values. To address the initial hypothesis, I overlapped the 29 predicted targets with significantly up-regulated genes from vtRNA1-2 KO total RNA-seq, which showed partial overlap of 7 genes (**Figure 33b**). Similar partial overlaps were seen for the up and down-regulated genes between chrRNA-seq and total RNA-seq (**Figures 32c and d**). These findings show partial consistency with the obtained results from chrRNA-seq; therefore, it is difficult to comprehensively conclude the impact of vtRNA1-2 KO on mRNA expression level.

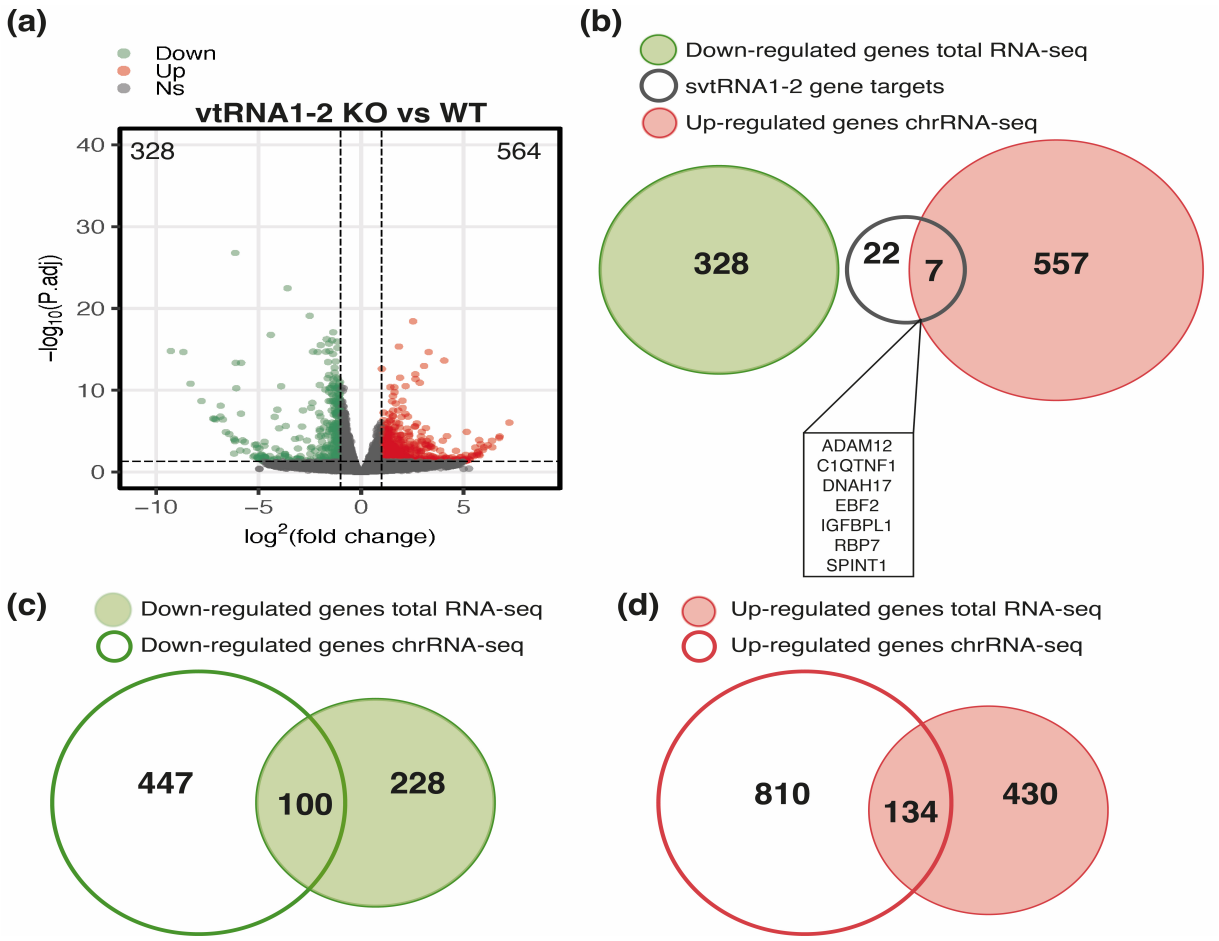


Figure 33: **(a)**Volcano plot highlighting DEGs in vtRNA1-2 KO versus WT. Red and green points identify genes with significantly increased or decreased expression, respectively (adj. p-value < 0.05 and FC $> 1 \mid < -1$). The number of genes that are significantly up or down-regulated is indicated in the plot's top left and right corners.**(b)** Venn diagram illustrating the overlap between 29 svtRNA1-2 targets (grey) with significantly up-regulated genes from total RNA sequencing from HEK293T cells in vtRNA1-2 KO versus WT conditions (red). No overlap was detected for the 29 targets with significantly down-regulated genes.**(c)** Venn diagram showing the overlap between significantly down-regulated genes from chrRNA-seq and total RNA-seq from vtRNA1-2 KO versus WT.**(d)** same as **(c)** but with significantly up-regulated genes.

4

Discussion

4.1 The biogenesis of vault RNA-derived small RNA fragments

In this work, I have identified novel Drosha-independent but Dicer-dependent vault RNA fragment from vault RNA1-2 and already-determined fragments of vault RNA1-1 and vault RNA2-1. I showed that upon Drosha KO in HCT116 cells and KD in HEK293T cells, sRNA-seq reads mapped to the vtRNA, increased, whereas reads mapped to miR-21 were greatly reduced (**Figures 14a and b**). These results indicate Drosha-independent processing of vtRNA, on the other hand Drosha-dependent processing of miR-21. A likely explanation for the increased number of mapped reads could be, that in Drosha's absence or decreased amount, Dicer does not have enough pre-miRNA substrate to process, resulting in more active vtRNA cleavage. The detected 5' end fragments of vtRNA1-1/2/3 upon Drosha KD that were not found in the Drosha KO dataset could potentially indicate insufficient Knock Down of Drosha, which generates more pre-microRNA substrates for Dicer, leaving a higher amount of intact vault RNA. Furthermore, Drosha fCLIP-seq (HEK293T) together with sRNA-seq from HEK293 WT data, demonstrated that vtRNA lacked the Drosha cleavage end-products, including 5' and 3' trailing ends and the pre-miRNA-like sequence, whilst they were present for miRNA processed by Drosha, such as miR-21. This observation is also supported by the fact that none of the vtRNA possesses properties recognised and necessary for Drosha cleavage, for example the clearly defined lower stem region without bulges, clearly defined apical loop or 2-nt 3' overhang after cleavage. Previously published data on Drosha involvement in vault RNA1-1 and vault RNA2-1 processing reinforced the results, demonstrating the fragmentation of vtRNA without Drosha involvement (Persson et al. 2009; J. Wang et al. 2020; Fort et al. 2020; Minones-Moyano et al. 2013).

In the effort to explore the biogenesis of vtRNA fragments, which might further play a role in gene silencing, I analysed sequencing datasets with small RNA isolated from HEK293T cells in siRNA and shRNA-induced Dicer KD and WT conditions. The small

RNA-seq with shRNA-induced Dicer KD contained three replicates per condition; therefore, I used it to investigate the differential expression of sncRNA genes. The obtained reads were subjected to trimming and filtering and subsequently mapped to the human reference genome (hg38). The inspection of mapped read size distribution showed only a mild difference in the peak height at the size range of the most prevalent mature microRNA-like fragments (21-23) (**Figure 17c**). Moreover, I only detected around 29% of miRNA among the down-regulated sncRNA population. The first potential explanation of the mild changes could be a weak shRNA-induced Dicer Knock Down, which is a process dependent on Dicer cleavage of shRNA, interfering with sufficient silencing of the target Dicer mRNA. The second plausible cause for the low number of identified miRNA might be a consequence of the chosen settings for reads alignment, which were more tailored to identify vault RNA. Because the four vtRNA share high sequence similarity of 84%, especially at the 5' end (no mismatch first 28 nt among vtRNA1-1/2/3) and 3' end (only 2 nt mismatches in the last 33 nt between vtRNA1-1 and vtRNA1-3) (**Figure 19**). Therefore, I decided to set the mapping options to only allow 1 mismatch and alignment to max 5 different genomic locations. The reads that did not pass these criteria were not included in the alignment file. Some microRNA can originate from multiple locations in the genome; for instance, miR-1302 has 10 different sites of origin (Chromosome 1, 2, 8, 9, 12, 15, 19, 20). In addition, microRNA are subjected to sequencing errors, genomic variations and A-to-I editing, which might also introduce nucleotide mismatches resulting in their failure to meet the alignment parameters. Both of these scenarios might be responsible for the observed phenotype.

The differential expression analysis between WT and Dicer KD conditions, revealed that vtRNA1-2 and vtRNA2-1 were down-regulated, but vtRNA1-1 showed increased expression levels. In line with my findings, the increased levels of vtRNA upon Dicer KD were also observed by Rybak-Wolf et al. 2014. Although the analysed data suggest that Dicer does not process vtRNA1-1 in HEK293T cells, it is worth noting other contrasting evidence of this effect. Up to this date, the only published research paper presenting evidence of Dicer processing of vault RNA1-1 into smaller fragments, including a fragment from 3' end, in MCF7 cells was done by Persson et al. 2009. In another study, Langenberger

et al. 2013 showed the effect of Dicer depletion on vault RNA in MCF7 cells, with the lowest Log_2Fold Change for vtRNA2-1 λ= -2.12 (this vtRNA was originally annotated as pre-miR-886 but later recognised as vault RNA paralog) followed by vtRNA1-2 λ= -0.76 and least changed was vtRNA1-1 λ= -0.14 (the expression of vtRNA1-3 was not detected). These values are put into perspective when compared with values detected for known microRNA such as miR-21 λ= -0.97 and the most affected miR-483 λ= -3.30 from the analysed dataset.

This discrepancy could reflect the different sensitivity of the experimental techniques and perhaps also Dicer KD efficiency. Furthermore, vtRNA are also differentially expressed in various tissues, suggesting vtRNA1-1 tissue-dependent Dicer processing. The process of vault RNA fragmentation seems to have more regulatory layers. Recent work by Hussain et al. 2013 showed that the processing of vault RNA1-1 is controlled by the m^5C modification on C69 deposited by NSUN2. Interestingly, only the fragment from the 3' end (also identified in my study) showed changes in the expression upon NSUN2 depletion. Moreover, they showed that this fragment associates with Argonaute proteins and could regulate CACNG7 and CACNG8 expression on the mRNA level (similar to miRNA). Yet, no experiments were performed to explore Dicer's involvement in the vtRNA1-1 cleavage mechanism. Future studies will need to address whether Dicer processing of vtRNA1-1 occurs in other cell lines. Taken together, these data support contradictory Dicer roles in vtRNA cleavage, suggesting that its function depends on the cell type and likely the amount of vtRNA and modification content.

Furthermore, the Dicer processing of vtRNA1-2 and vtRNA2-1 into smaller fragments led to the hypothesis that svtRNA might play a role in gene silencing as previously shown for other sncRNA-derived fragments such as tRF (Di Fazio et al. 2022). In pursuit of the hypothesis, I investigated the interaction between the Argonaute proteins and newly identified 3' end fragment of vtRNA1-2 as well as 5' end fragment of vtRNA2-1 and 3' end vtRNA1-1. I identified only 3' end fragments derived from vtRNA1-1/2 associated with AGO proteins and none from 5' svtRNA2-1. Indeed vtRNA1-1 3' end fragment has been previously described to bind AGO proteins (AGO2 and AGO3) (Hussain et al. 2013). In line with the obtained findings, Fort et al. 2020 demonstrated that AGO2

preferentially binds to 3' detected svtRNA2-1 despite the significantly higher expression of 5' svtRNA2-1. Following these observations, another study has also observed increased levels of 5' svtRNA in connection with cancer and poor patient prognosis (J. H. Li et al. 2011). On the contrary, a separate study identified a 5' end fragment of vtRNA2-1 loaded onto AGO proteins (AGO2 and AGO3) but did not show significant downregulation of the predicted targets in SH-SY5Y neuronal cells, suggesting a different functional role than post-transcriptional gene silencing (Minones-Moyano et al. 2013). Thus, these findings support the previous vtRNA observations of cell type-specific processing and also functional relevance.

In recent years, various evidence emerged supporting the presence of RNAi components in both cytoplasm and nucleus (Sarshad et al. 2018; Gagnon et al. 2014; Kalantari et al. 2016a; Langlois et al. 2005), indicating that RNAi might recognise nuclear RNA and control gene expression. In order to investigate the possible mechanism of svtRNA function, I first set to identify the localisation of svtRNA in different cellular compartments. Using computational and experimental validation, I showed that svtRNA1-2 can be found in both the nucleus and cytoplasm (**Figure 22**), suggesting that svtRNA1-2 might act on two levels: nascent transcripts and fully formed messenger transcripts. I have decided to investigate the role of svtRNA1-2 in nascent RNA silencing. This mechanism was described for specific tRNA-derived fragments, which caused the downregulation of target genes on the nascent RNA level (Di Fazio et al. 2022). I reasoned, based on their common characteristics that they might as well share the same mechanism of function. To verify this hypothesis, I first identified the svtRNA1-2 nascent RNA targets using chrRNA-seq from HEK293T cells in WT, AGO2 and Dicer KD conditions. The svtRNA binding sites prediction analysis revealed the highest number of unique targets (227) for svtRNA1-2 with specific binding site distribution and preference for intronic binding. The prediction results of svtRNA targets might be partially attributed to the GC content (svtRNA1-2 with GC content of 56.5%, svtRNA1-1 66.6% and svtRNA1-3 with 74%). For comparison, the average human miRNA GC content is around 52% (**Supplementary Figure 35a**), indicating the plausible role of higher GC content in the reduced target site prediction. I identified the distribution of svtRN1-2 binding sites throughout the transcript length, not

only at the beginning, as shown for svtRNA1-1/3 with the highest prevalence of predicted target regions within intronic RNA (**Figure 24f**). In *drosophila*, the order of the spliced introns follows the direction of the transcription. Recent work by Drexler et al. 2020 revealed that this does not apply to human intron splicing, demonstrating slower splicing with terminal introns spliced out after cleavage and polyadenylation, which opens the opportunity for binding and regulating nascent transcripts with svtRNA1-2 and AGO2 along the sequence. However, it is worth noting that in this study, the nascent transcripts were metabolically labelled by the 4-thiouridine approach, which was shown to include nascent transcripts and other contaminating RNA such as mRNA. Therefore, an approach such as polymerase intact nascent transcript POINT-seq or POINT-nano, which does not rely on base labelling and captures truly nascent RNA, could help resolve the kinetics of the particular intronic regions (Sousa-Luis et al. 2021). Taken together, the obtained results demonstrate that vault RNA are processed into smaller functional sequences independently of Drosha but dependent on Dicer. Moreover, the newly identified svtRNA1-2 bound to Argonaute proteins in the nucleus indicates a functional role in regulating nascent RNA levels via direct sequence-specific target cleavage in the intronic regions.

4.2 The functional role of small vault RNA1-2 fragment

The CRISPR-Cas12a vtRNA1-2 Knock Out (KO) experiment was designed to understand the vtRNA1-2 function. The colony screening revealed a colony with vtRNA1-2 Knock Out. Interestingly, the genomic deletion of vtRNA1-2 resulted in the inhibition of cell growth and proliferation. Such dramatic defects in cell growth and proliferation were not seen previously in vtRNA1-1 and vtRNA1-3 KO cells (Bracher et al. 2020), suggesting that vtRNA1-2 might play a role in essential cellular processes. It is worth noting that this phenotype can be caused by direct and indirect effects of vtRNA1-2/vault complex, intact vtRNA1-2 and/or svtRNA1-2. Recently, independent research groups failed to Knock Out vtRNA2-1/nc886 in HeLa, Du145 and other cell lines due to its interaction with PKR, which plays a role in cell proliferation and apoptosis (Saruuldalai et al. 2022; R. Yang et al. 2022). These results provide an insight into the variability of the vtRNA biological

roles as well as their importance.

After obtaining the vtRNA1-2 KO, the expression changes of ERBB4 on both steady-state and nascent RNA levels were confirmed by qRT-PCR. Using chrRNA-seq, I was able to investigate changes in the nascent transcriptome on a global scale which also confirmed the upregulation of ERBB4 in vtRNA1-2 KO. The GO enrichment analysis of the significantly down-regulated genes showed enrichment for biological processes related to Neurons and Neurogenesis. Interestingly, while this work was in preparation, Wakatsuki et al. 2021 reported that increased expression of vtRNA1-1 upregulates synaptogenesis by activating the MAPK signalling pathway, suggesting vtRNA1-2 might also be involved in similar signalling pathway regulation. The GO analysis of the significantly up-regulated genes showed, on the other hand, enrichment for biological processes related to membrane potential and transport. After the annotation of the 29 svtRNA1-2 targets subcellular localisation, I found most of them in the cellular membrane. This finding was in line with the identified biological processes, indicating that svtRNA1-2 might indeed be partially responsible for the observed proliferation phenotype.

Well-established rules for functional miRNA-target pairing classify target sites into three main categories; canonical sites that pair at both 5' and 3' ends, seed sites, depending on minimal 5' end base pairing, and 3' compensatory sites characterised by weak 5' base pairing and strong compensatory pairing to the 3' end (Alagia et al. 2016). To assess the NRS requirements of the svtRNA1-2 molecule, I analysed the svtRNA1-2 WT pairing conditions between svtRNA1-2 and predicted targets. A large set of predicted targets for svtRNA1-2 WT showed perfect 5' seed complementarity with extended wobble base pairs at the 3' end. Nucleotide conservation analysis of the predicted targets highlighted the presence of a GC-enriched segment at the 5' end that might have additional roles, such as the 5' Terminal Oligo Guanine (TOG motif) described as a characteristic feature of some 5' tRF (Shi et al. 2019). I cannot exclude that svtRNA1-2 might form intermolecular RNA G-quadruplex structures and gain further functions in translational inhibition, protein recognition, or cellular localisation. In addition, svtRNA1-2 central (10-11-12) positions predominantly follow the base pairs scheme Wobble-WC-Wobble. It should be noted that most of the predicted targets have an unpaired loop of 5 nucleotides opposite to positions

11 and 12 of the svtRNA1-2, with a common sequence 5'-ACUAC or AUUAC. It has been demonstrated that AGO2 can accommodate a central loop present on the target gene without losing binding affinity (Becker et al. 2019).

It is tempting to speculate a connection between the identified intronic AGO2 binding sites with a 5 nucleotides bulging consensus motif and the regulation of alternative splicing. Considering, the findings of Ameyar-Zazoua et al. 2012; Allo et al. 2009 which showed, using Knock Down and overexpression experiments, that AGO2 and Dicer can influence alternative exon splicing. Furthermore, it has been shown that siRNA targeting intronic regions could modify the splicing profile of the adjacent exon (Allo et al. 2009). These results further illustrate the role of RNAi components in the regulation of splicing. The chromatin-bound Argonaute proteins showed a recruitment of spliceosome and their binding to the splicing factors including SRSF1, SRSF3, SRSF7 and SRSF10, demonstrating direct interactions between them (Ameyar-Zazoua et al. 2012). Moreover, the identified motif 5'-ACUAC in the loop structure is a binding motif of the SRSF3 splicing factor (Zaharias et al. 2021), which would be blocked for SRSF3 recognition by the AGO2-svtRNA1-2 complex to regulate splicing. Thus, further studies seeking a definitive understanding of the functional roles of svtRNA1-2 should require consideration of alternative splicing modulation.

svtRNA1-2 showed a distinct binding pattern (seed and 3' supplementary region complementarity) to the target sequences. Therefore, two svtRNA1-2 molecules with a shuffled sequence at either the 5' end or 3' supplementary region (shuffled svtRNA1-2 5' and shuffled svtRNA1-2 3', respectively) were designed to further evaluate the minimal pairing requirements between svtRNA1-2 and the targets. Interestingly, the svtRNA1-2 shuffled 5' designed to follow the rules of thermodynamic asymmetry, present in large numbers of miRNA molecules, did not have any common target genes with svtRNA1-2 WT. The complete loss of common targets between WT and 5' and 3' shuffled svtRNA1-2 could have been partially due to the miRanda free energy cut-off. Thus svtRNA1-2 molecule with a swapped 5' and 3' seed sequence was designed to preserve the nucleotide composition and sequence identity while following the rules for designing efficient miRNA molecules. The target prediction analysis revealed an increase in the putative targets and a dramatic

change in the target sequence complementarity localisation. The swapped svtRNA1-2 hybridises within the exons of the predicted genes, in contrast to the svtRNA1-2 WT, which hybridised within introns via loop protrusion. This data sheds light on the importance of the G-rich region at the 5' end for the svtRNA1-2-mediated NRS.

In order to validate the effect of the svtRNA1-2 nascent RNA silencing on the mRNA level, I isolated the total RNA from vtRNA1-2 KO and WT and submitted the samples for sequencing. While processing the sequencing files, I noticed an overexpression of reads belonging to rRNA, which caused the loss of approximately 68% of the reads after filtering. Due to this loss, the resulting number of reads was below optimal. Nevertheless, I quantified the obtained reads and performed a differential expression analysis. As expected, the number of significantly up and down-regulated genes was comparably lower to chrRNA-seq (vtRNA1-2 KO and WT), taking into account the significance levels (total RNA-seq, adj. p-value < 0.05 and $Log_2FC > 1 \mid < -1$) and (chRNA-seq, adj. p-value < 0.001 and $Log_2FC > 1 \mid < -1$). The overlap of 29 gene targets with the up-regulated genes from total RNA sequencing also did not show striking similarity, despite the low significance threshold set for the total RNA-seq genes. Due to the low coverage sequencing data, it is hard to draw a biological conclusion whether the levels of target mRNA are changed due to the svtRNA1-2 silencing or the total low read sample amount. Furthermore, ERBB4 did not even pass the relatively low stringency data filtering stage for the expressed genes, which included the removal of genes with less than 5 reads mapped in at least 2 replicates per condition before running the differential expression analysis using DESeq2 (Love et al. 2014) package in R. This result was not surprising as the ERBB4 expression levels in chrRNA-seq data were Log_2FC 1.4, and the amount of uniquely mapped reads was approximately 100 million, suggesting it would not be detected in the samples with low numbers of assigned reads such as 8 million. To be able to conclude the global impact of svtRNA1-2 nascent RNA silencing on steady-state mRNA, this experiment would have to be repeated. Although the samples were suboptimal, the obtained reads most probably represent highly abundant transcripts in the cells, which might in fact, indicate the role of intact vtRNA1-2 or vtRNA1-2 associated with the vault complex in the cell.

5 Conclusion

My PhD research has focused on the biogenesis of fragments derived from small non-coding RNA called vault RNA and its' regulatory role in mammalian gene expression. Eukaryotic gene expression has multiple layers of regulation. The RNAi process involving small non-coding RNA in regulating RNA activities is one of them. It is becoming increasingly clear that small RNA fragments generated from sncRNA are not just degradation products but maybe also important gene expression regulators. In the second chapter, I investigated the biogenesis of small vault RNA-derived fragments using computational and experimental approaches. I first studied the influence of individual components of RNAi on vault RNA. This study revealed that vtRNA is processed into smaller fragments independently of Drosha but dependent on Dicer. Furthermore, the essential effector part of RNAi AGO2 showed binding to novel svtRNA1-2 and localisation in the nucleus and cytoplasm. I explored the svtRNA1-2 nuclear function using *in silico* analyses of nascent RNA sequencing data obtained from Knock Down experiments (Dicer, AGO2), which revealed a subset of specific plausible svtRNA1-2 gene targets. In the third chapter, I studied the regulatory roles of vtRNA1-2 by generating a stable Knock Out cell line resulting in the loss of intact vtRNA1-2, svtRNA1-2 and vtRNA1-2 bound to vault complex. It emerged that vtRNA1-2 has an essential role in the cell as its absence severely influenced cell proliferation, growth and adhesion. Furthermore, I showed that the expression of ERBB4 is regulated by svtRNA1-2 nascent RNA silencing with a critical preference for target binding to complementary seed and 3' supplementary regions. The effects of svtRNA1-2 were validated using svtRNA1-2 transfection of MCF7 cells (which have increased levels of ERBB4 expression), causing a decrease in cell proliferation. This work resulted from an interaction between my wet and dry lab analyses, where the computationally generated results fuelled a series of hypothesis which I then experimentally validated.

In conclusion, my work provides new insight into vault RNA1-2 cleavage, structure and direct and indirect functions. This is the first comprehensive study focused on vtRNA1-2

that provides a strong foundation for further research.

6

Methods

6.1 The biogenesis of vault RNA-derived small RNA fragments

6.1.1 Creating small non-coding RNA annotation file

Custom small non-coding RNA annotation file was built by merging three annotation files including miRNA from miRbase (GFF3 format) (v22.1) (Kozomara et al. 2019), tRNA from the GENCODE website (GTF format)(release 34) and snoRNA, snRNA, miscRNA, scaRNA, rRNA and vault RNA from GENCODE basic gene annotation (release 34) (hg38) (GTF format) (Frankish et al. 2021). GFF (General Feature Format) and GTF (General Transfer Format) are tab-separated files used to annotate genomic intervals. Their structure is very similar; however, the main differences underlay in formatting and system of the annotation.

6.1.2 Human and Mouse reference genome

The reference genomes for human (hg38) and mouse (m39) were downloaded from the GENCODE website. Three human (hg38) genome indexes were built, the first one using bowtie-build (v1.2.3) (Langmead et al. 2009) with default settings. The second and third indexes were generated by STAR (v2.7.3a)(Dobin et al. 2013) with default settings and w/o basic gene annotation (hereby referred to as genome with annotation and genome without annotation). The genome without annotation was built to align small RNA reads since the annotation file provides information about splice junctions which are unnecessary for the sncRNA mapping.

6.1.3 Identification of Drosha involvement in vault RNA processing

In order to determine the involvement of Drosha in vault RNA processing, I analysed raw data files of two small RNA sequencing datasets from HEK293 (siRNA-induced Drosha KD vs WT) and HCT116 cells (Drosha KO using TALEN and Cas9 vs WT). These data

were retrieved from gene expression omnibus (Edgar et al. 2002; Barrett et al. 2013) containing one biological replicate per condition (WT and Drosha KD) (GSE55333, **Table 1**) and two biological replicates per condition (WT and Drosha KO) (GSE77989, **Table 1**). Sequence adapters TGGAATTCTCGGGTGCCAAGGAACTCCAGTCAC (Illumina TruSeq SmallRNA forward adapter) for both datasets were trimmed from the reads with cutadapt (v2.10) (Marcel 2011). The trimmed reads were then mapped to the human reference genome (hg38) without the annotation file, using STAR (v2.7.3a) with the following options:

```
# STAR alignment
STAR --genomeDir                    #Path to the genome index
     --readFilesIn                  #Input FASTQ file
     --readFilesCommand zcat        #Input compressed FASTQ file
     --outSAMtype SortedByCoordinate #Output sorted BAM file by
   coordinate
     --alignIntronMax 1             #Maximum intron size
     --outFilterMismatchNmax 1      #Number of mismatches
     --outFilterScoreMinOverLread 0  #Threshold for read-length
   normalised alignment score
     --outFilterMatchNminOverLread 0 #Threshold for read-length
   normalised number of matched bases
     --outFilterMultimapNmax 5      #Max number of multiple alignments
   allowed for a read
     --runThreadN 4                 #Number of threads
     --outFileNamePrefix            #Output file name prefix
```

The aforementioned settings were used because STAR is a splice-aware aligner, able to map reads to splice junctions. However, the reads generated from small RNA sequencing do not contain splice junctions. Therefore, for the mapping of FASTQ files, I prohibited splicing with –alignIntronMax 1 option and turned off the filter for the minimal mapped length using –outFilterScoreMinOverLread 0 –outFilterMatchNminOverLread 0 options. The output alignment files were indexed using Samtools (v1.10) function index (H. Li et al. 2009), and their coverage was calculated using bamCoverage of the deeptools package (v2.2.2) (Ramirez et al. 2014) with selected options:

–binSize 1

–normalizeUsing CPM Counts per million (CPM) mapped reads

The bin size was set to 1 to calculate per-base read coverage, which was normalised using the CPM method recommended for RNA-seq. The output files in bigWig format, which is an indexed binary format containing genomic positions and their coverage depth, were then converted to bedGraph using ucsctools (v385) bigWigToBedGraph function. The resulting coverage depth tabulated text file was then loaded into python3 (Van Rossum et al. 1995) to generate coverage plots for vtRNA1-1/2, vtRNA2-1 and miR-21 (WT and Drosha KO). Meanwhile, the bigWig files with normalised read coverage were imported into the IGV (Thorvaldsdottir et al. 2013) for visualisation of the relevant sncRNA (WT and Drosha KD). The reason for different data visualisations is due to the lack of biological replicates of the Dicer KD dataset.

6.1.4 Drosha Formaldehyde Crosslinking, Immunoprecipitation, and sequencing (fCLIP-seq)

This data analysis aims to strengthen the results obtained from Drosha KD and KO vs WT small RNA sequencing, which indicated Drosha-independent vtRNA cleavage. First, I obtained FASTQs of Drosha fCLIP-seq two biological replicates from the data repository (GSE93653, **Table 1**). The raw reads were processed according to the pipeline described in (B. Kim et al. 2019). Briefly, the reads were trimmed using cutadapt (v2.10) to remove adapters 3' TACAGTCCGACGATC and 5' TGGAATTCTCGGGTGCCAAGG. Then combined using pear (v0.9.0) (J. Zhang et al. 2014) with default settings and aligned to the human reference genome with annotation (hg38) using STAR (v2.7.3a), discarding reads mapped to more than 5 locations and with more than 4 mismatches. Next, I indexed the alignment BAMs with Samtools (v1.10) and visualised them in IGV with the above-mentioned small RNA-seq data from HEK293 cells in WT condition (GSE55333, **Table 1**). The small RNA sequencing was used to display end products of the RNAi processing.

6.1.5 Identification of Dicer involvement in vault RNA processing

The samples for small RNA sequencing were prepared in triplicates using HEK293T cells in WT and short-hairpin RNA (shRNA) induced Dicer KD conditions by Dr Matthew Davis in Dr Gullerova's laboratory. The methods for this experiment are reported in Di Fazio et al. 2022.

To investigate the role of Dicer in vault RNA fragmentation, I explored two small RNA sequencing datasets from HEK293T cells in Dicer KD and WT.
First, I used sRNA-seq with shRNA-induced Dicer KD vs WT (GSE126751, **Table 1**). The raw reads were trimmed 3' adapter contaminations AGATCGGAAGAGCACACGTCT-GAACTCCAGTCA (Illumina Universal Adapter) using cutadapt (v2.10) and then filtered based on their length, retaining only those reads of size 15 to 29 nucleotides. This filtration step was carried out based on the sequence length of microRNA, which range from 14 to 27 (**Supplementary Figure 35a**). Next, I mapped these reads to the human reference genome without annotation (hg38) using STAR (v1.2.3), allowing 1 mismatch and alignment to max 5 genomic locations with the same settings as in the '*Identification of Drosha involvement in vault RNA processing*' section.
The mapped reads were then quantified using featureCounts (v2.0.0) (Liao et al. 2014) using the custom annotation file of small non-coding RNA. The output count matrix was then imported into R-Studio (RStudio Team 2020) with R version (4.2.0) as read count matrix. I filtered out sncRNA with less than (minimum 5 reads in 2 biological replicates per condition) and normalised the matrix using DESeq2 (v1.36.0) (Love et al. 2014). To evaluate the effect of Dicer KD, I performed a differential expression analysis and selected significantly up and down-regulated small ncRNA with adj. p-value lower than 0.05 and Fold Change $> 1.5 \mid < -1.5$, respectively. The resulting data visualisation was performed in the R-Studio. In order to examine the read alignment locations, I calculated and visualised the vault RNA per-base coverage in the same way as in the '*Identification of Drosha involvement in vault RNA processing*' section.

The second sRNA-seq dataset of WT and siRNA-induced Dicer KD was downloaded from GEO (GSE55333, **Table 1**) and contained two biological replicates in Dicer KD and one replicate in WT. The raw data files were processed according to instructions described in Rybak-Wolf et al. 2014. The alignment BAMs were indexed and subjected to the normalised coverage depth calculations as described in the '*Identification of Drosha involvement in vault RNA processing*' section.

6.1.6 Dicer PAR-CLIP Sequencing data analysis

To investigate Dicer binding to vtRNA. First, I downloaded FASTQs files from GEO (GSE55333, **Table 1**) in three biological replicates. Then I trimmed the 3' adapter sequences TCTCACGTCGTATGCCGTCTTCTGCTTG, TCTCCATTCGTATGCCGTCTTCTGCTTG, TCTCCCATCGTATGCCGTCTTCTGCTTG and removed reads shorter than 15 bases using cutadapt (v1.8.3) (Marcel 2011). The alignment was performed using Bowtie (v1.2.3) with the same settings as previously mentioned in the '*AGO1-4 PAR-CLIP Sequencing data analysis*' section. The same applied to the PARalyzer (v1.5) run.

6.1.7 vtRNA sequence alignment

The vtRNA sequence alignment was done using MUltiple Sequence Comparison by Log-Expectation (MUSCLE) with default settings in Jalview (Waterhouse et al. 2009), highlighting consensus bases in blue.

6.1.8 PhastCons Conservation Analysis

To determine vtRNA1-2 conservation level, I used phastCons scores which represent from the multiple sequence alignment of 100 vertebrate genomes to the human genome. The files are provided in wigFix format in the UCSC database (hg38.100way.phastCons). The phastCons scores corresponding to vtRNA1-2 were extracted and averaged. The resulting density values of the normalised phastCons scores were plotted using python3.

6.1.9 vtRNA structures prediction and visualisation

In order to obtain a comprehensive understanding of the identified vault RNA fragments with respect to the secondary and tertiary structure, I made use of recently generated immunoprecipitation pull-down selective 2'-hydroxyl acylation followed by Primer Extension (icSHAPE-MaP) data (Luo et al. 2021), from which I retrieved values for vault RNA1-1/2/3 RNA. Chemical probing is a powerful method for characterising RNA structures and improving RNA secondary structure prediction. Next, I used RNAfold (Lorenz et al. 2011) to construct the secondary structures and to generate dot-bracket notation of the assemblies providing input information about vault RNA sequence in FASTA format and icSHAPE values. In addition, I used this tool to highlight the locations of the identified vault RNA1-2 fragment. To construct a 3D structure and produce a PDB file, I employed the RNAComposer Automated RNA Structure 3D Modeling server (Popenda et al. 2012; Antczak et al. 2016). The PDB file was subsequently visualised in Jmol (Jmol development team 2016).

vault RNA structural details			
Transcript Name	**Sequence**	**Dot-Bracket Notation**	**MFE**
vault RNA1-1	GGGCUGGCUUUAGCUCA GCGGUUACUUCGACAGU UCUUUAAUUGAAACAAG CAACCUGUCUGGGUUGU UCGAGACCCGCGGGCGC UCUCCAGUCCUUUU	(((((((.....((((.((((((...((((...(((.......)))....(((((((....))))))))))))).)))))))).....))))))))....	-58.45 kcal/-mol
vault RNA1-2	GGGCUGGCUUUAGCUCA GCGGUUACUUCGAGUAC AUUGUAACCACCUCUCU GGGUGGUUCGAGACCC GCGGGUGCUUUCCAGCU CUUUU	(((((((....(((((.((((....((((..... .(((((((....)))))))))))..)))).)).)))..))))))))....	-53.30 kcal/-mol
vault RNA1-3	GGGCUGGCUUUAGCUCA GCGGUUACUUCGCGUG UCAUCAAACCACCUCUC UGGGUUGUUCGAGACCC GCGGGCGCUCUCCAGCC CUCUU	(((((((.....((((.((((...........(((.((.(((.(((......))).))).)))))))))))). ....))))))))....	-58.21 kcal/-mol
vault RNA2-1	GGGUCGGAGUUAGCUC AAGCGGUUACCUCCUCA UGCCGGACUUUCUAUCU GUCCAUCUCUGUGCUGG GGUUCGAGACCCGCGG GUGCUUACUGACCCUUU UAUGCAAUAAAUUCGGU AUAAUCUGUCA	The icSHAPE data for vault RNA2-1 were not present to generate dot-bracket notation	-37.39 kcal/-mol

6.1.10 Identification of small vault RNA in mouse tissue samples

To investigate the possible expression of murine small vault RNA, I used sRNA-seq data from various mouse tissues, including (spleen, liver, lungs, muscle, brain, and pancreas) (GSE119661, **Table 1**). The data were obtained in triplicates for most of the tissues and then duplicates for (Brain and Heart) and processed according to published methods Kern et al. 2020. Briefly, obtained raw reads were processed and aligned to mouse genome index (m39) using STAR (v2.7.3a), and after the alignment, I generated normalised coverage files in bigWig format using bamCoverage function from deeptools (v2.2.2) and visualised them in IGV.

```
# STAR alignment
STAR --genomeDir                    #Path to the genome index
     --readFilesIn                  #Input FASTQ file
     --readFilesCommand zcat        #Input compressed FASTQ file
     --outSAMtype SortedByCoordinate #Output sorted BAM file by
   coordinate
     --alignIntronMax 1             #Maximum intron size
     --outFilterMatchNmin 16.       #Sum of mates lengths for paired
   -end reads
     --outFilterMismatchNoverLmax 0.05 #Ratio of mismatches
     --outMultimapperOrder Random   #Order of multimapping alignments
     --clip3pAdapterSeq TGGAATTCTC  #Clip adapter sequences from 3p
     --clip3pAdapterMMp 0.1         #Max proportion of mismatches for 3
   p adpater
     --outFilterScoreMinOverLread 0  #Threshold for read-length
   normalised alignment score
     --outFilterMatchNminOverLread 0 #Threshold for read-length
   normalised number of matched bases
     --runThreadN 4                 #Number of threads
     --outFileNamePrefix            #Output file name prefix
```

6.1.11 AGO1-4 PAR-CLIP Sequencing data analysis

Next, I set to investigate the potential binding of svtRNA1-2 to Argonaute proteins. I obtained AGO1-4 photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) datasets (GSE21918, **Table 1**). The experimental technique is useful for identifying the binding sites of cellular RNA-binding proteins (RBPs). The adapter sequence was trimmed using cutadapt (v1.8.3) with the following settings:

```
# Cutadapt adapter sequences trimming
cutadapt -e 0.05  #The maximum error rate
         -m 15    #Discard reads shorter than 15 nt
         -M 29    #Discard reads longer than 29 nt
         -a TCGTATGCCGTCTTCTGCTTGT #Adapter1
         -o file_1_trimmed.fastq.gz #Output file 1
         file_1.fastq.gz #Input file 1
```

Then the reads were mapped to the human reference genome (hg38) by Bowtie (v1.1.2).

```
# Bowtie alignment
bowtie  -v 3 #Number of mismatches allowed
        -m 1 #Report only uniquely mapped alignments
        -a   #Report all valid alignments
        --best --strata #Report alignments in the best alignment "
   stratum"
        -q   #Input fastq file
        -S   #Output sam file
```

The clusters of reads (AGO binding sites) were identified based on T-to-C conversions using PARalyzer (v1.5) (Corcoran et al. 2011) with following settings:

```
BANDWIDTH=3
CONVERSION=T>C
MINIMUM_READ_COUNT_PER_GROUP=5
MINIMUM_READ_COUNT_PER_CLUSTER=5
MINIMUM_READ_COUNT_FOR_KDE=5
MINIMUM_CLUSTER_SIZE=8
MINIMUM_CONVERSION_LOCATIONS_FOR_CLUSTER=1
MINIMUM_CONVERSION_COUNT_FOR_CLUSTER=1
```

```
MINIMUM_READ_COUNT_FOR_CLUSTER_INCLUSION=8
MINIMUM_READ_LENGTH=13
MAXIMUM_NUMBER_OF_NON_CONVERSION_MISMATCHES=0
EXTEND_BY_READ
SAM_FILE=input.sam
GENOME_2BIT_FILE=hg38.2bit
OUTPUT_DISTRIBUTIONS_FILE=output_distribution.csv
OUTPUT_GROUPS_FILE=output_groups.csv
OUTPUT_CLUSTERS_FILE=output_clusters.csv
OUTPUT_READS_FILE=output.sam
```

The output SAM files were converted into BAMs, sorted and indexed using Samtools (v1.10). The sample coverage was calculated using the bamCoverage function from deeptools (v2.2.2) and visualised in IGV.

6.1.12 AGO2/3 RIP-Seq data analysis

The AGO2/3 RIP-seq datasets contained 1 replicate per each AGO sample (GSE55333, **Table 1**). The 3' adapter sequence TCTCGTATCGTATGCCGTCTTCTGCTTG was trimmed using fastx-clipper from fastx toolkit (v0.0.13.2) from Hannon Lab with a quality score of 33 and a read length longer than 15 nt. The reads were mapped to the human reference genome (hg38) using Bowtie (v1.1.2). After the alignment, I generated normalised coverage files in bigWig format using bamCoverage function from deeptools (v2.2.2). Output files were visualised using IGV.

6.1.13 Sequencing data analysis of RNA bound to AGO2 in cytoplasm and nucleus

The raw read files of RNA bound by nuclear and cytoplasmic AGO2 were obtained from the gene expression omnibus repository (GSE66665, **Table 1**). The reads were trimmed using cutadapt (v1.8.3) with the following settings -e 0.05 -q 23 -m 15 -M 30 -a 'A[39]'. The processed reads were then aligned to (hg38) using STAR aligner (v2.7.3a), allowing a maximum of 3 mismatches and retaining uniquely mapped reads. Next, the bamCoverage function from deeptools (v2.2.2) was used to calculate coverage files using

–binSize 1

–normalizeUsing CPM

To create BigWig files, which were then visualised in IGV.

6.1.14 Confocal and bright light microscopy

This experiment has been performed by Dr Adele Alagia in the Dr Gullerova laboratory.

confocal microscopy with fluorescently labelled svtRNA1-2, was performed to experimentally validate the svtRNA1-2 nuclear localisation. svtRNA1-2 5'-CGCGGGUGCUUUCCAGCUCUUUU-3' was labelled with Alexa Fluor® fluorescent dye by FISH Tag RNA Multicolor Kit/RNA Green Kit (Invitrogen) according to manufacturer's instructions. HEK293T cells were grown on a poly-L-lys coverslip and transfected with single-stranded (100nM final concentration) or double-stranded (40nM final concentration) 5'-Cy3-labelled svtRNA1-2 24h after seeding. Cells were fixed at specific time points (4, 6 and 24h) with paraformaldehyde (4% in PBS). Subsequently, the sample was imaged on an FV1200 Olympus laser scanning confocal microscope equipped with an argon laser and a He/Ne 633 nm laser, using 60x/NA1.4 oil immersion objective. Comparative immunofluorescence analyses were performed, maintaining identical acquisition parameters. Nuclear foci were analysed with Cellprofile software (Stirling et al. 2021), and statistical analysis was performed using GraphPad Prism. Bright light microscopy was performed using a 40x lens.

6.1.15 Data analysis of chromatin-associated RNA sequencing in WT, Dicer and AGO2 KD conditions

*The samples for chrRNA sequencing were prepared in duplicates using HEK293T cells in WT, Dicer and AGO2 KD conditions (induced by short-hairpin RNA) by Dr Pong in the laboratory of Dr Gullerova. The methods for this experiment are reported in Di Fazio et al. 2022. The bioinformatics pipeline is summarised in the (**Figure 34**)*

6.1.16 chromatin-associated RNA sequencing data pre-processing

I analysed chromatin-associated RNA sequencing data to identify the potential nuclear targets of svtRNA1-2 (analysis pipeline summarised in a (**Figure 34**)). These sequencing datasets contain nascent transcripts attached to chromatin by RNA polymerase II and many regulatory chromatin-bound long non-coding RNA. First, the adapter sequences were trimmed using cutadapt (v2.10), only keeping reads longer than 10 bases after trimming low-quality reads from 3' and 5' ends and a max error rate of 0.05.

```
# Cutadapt adapter sequences trimming
cutadapt -e 0.05   #The maximum error rate
         -m 10     #Discard reads shorter than 10 nt
         -j 8      #CPU cores
         -q 15,10  #Trim low-quality bases from 3' and 5' end of reads
         -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA #Adapter1
         -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT #Adapter2
         -o file_1_trimmed.fastq.gz  #Output file 1
         -p file_2_trimmed.fastq.gz  #Output file 2
         file_1.fastq.gz file_2.fastq.gz #Input file 1 and Input file 2
```

The processed reads were mapped to the human reference genome (hg38) with annotation (hg38) using STAR (v2.7.3a), discarding reads with multiple mapping genomic locations, leaving uniquely mapped reads.

```
# STAR alignment
STAR --genomeDir #Path to the directory with Genome Index
     --readFilesIn #Input file 1 and input file 2
     fastq_1_trimmed.fastq.gz
     fastq_2_trimmed.fastq.gz
     --outSAMtype BAM SortedByCoordinate #Output parameters
     --sjdbGTFfile gencode.basic.annotation.gtf #Genome annotation file
     --readFilesCommand zcat #Read files from compressed file format
     --outFilterMultimapNmax 1 #Preserve reads mapped to only single
   genomic location
     --runThreadN 8 #Number of used threads
     --outFileNamePrefix #Output file prefix name
```

Reads quantification was done using featureCounts (v2.0.0) with set options for the paired-end library type and counting boundaries of gene coordinates using the genome annotation file. I used the gene coordinates option and not the exon coordinates to acquire counts from the intronic as well as exonic regions. The data was then imported into R-Studio/R version (4.2.0) for statistical analysis.

```
# Quantification
featureCounts -a gencode.basic.annotation.gtf #Genome annotation file
              -T 8 #Number of threads
              -t gene    #Feature type included for read counting
              -g gene_id #Attribute type used to group features
              -s 2 #Reversely stranded
              --verbose  #Provide additional details about this run
              -o  #Output file
              #Input BAM files
```

In the R-Studio/R version (4.2.0), any gene with more than 10 read counts in both replicates per condition was kept, and the rest was filtered out. This step was performed to remove genes with no-expression and noise values. The count matrix was then normalised using DESeq2 (v1.36.0). To identify significantly differentially expressed genes, I used an adj. p-value threshold value of <0.001 and Log_2Fold Change $> 1 \mid < -1$.

6.1.17 GO

The gene ontology enrichment of 227 gene targets was performed using ShinyGO (v0.76) (Ge et al. 2020) with all protein-coding genes as background. GO nodes with less than 2 genes enriched and FDR cutoff score higher than 0.05 were excluded from the analysis. The visualised representation of data was downloaded from the ShinyGO (v0.76) Web interface.

6.1.18 Canonical transcripts annotation file

This step aims to produce an annotation file for only one 'canonical' transcript per gene. I used the GENCODE basic annotation file and parsed it into data frame format. Based on the Ensembl guide, I used the transcript flags to identify transcripts with the highest

quality and relevance. I selected one transcript per gene with the highest number of Tags and created a genome annotation file with only 'canonical transcripts'.

6.1.19 Gene feature annotation file

To investigate the precise locations of predicted binding sites along the transcript length, I constructed an annotation file which included the following information: start, end, chromosome name and feature type of each canonical transcript. To make the script more versatile, I extracted all the data from the GENCODE GTF standard annotation file. First, I calculated the start and end positions of introns per transcript by subtracting the start position from the exons end position, taking into account the transcript strandness. Next, I retrieved the exons and UTRs positions. Furthermore, I converted the genomic transcript coordinates into positions relative to the transcript length; in this way, I was able to map the predicted binding positions back to the transcript feature. In the final step, I merged these files and ensured that the created coordinates were in the correct order based on the feature position and the transcript strandness.

6.1.20 Reference sequences for svtRNA target prediction

After obtaining the list of canonical transcripts per gene from the GENCODE annotation file, I overlapped them with up-regulated protein-coding genes based on their Ensemble gene ID to get canonical transcripts for each significantly up-regulated gene. To create a BED format file, I selected six columns which included chromosome name and transcript start position, which was subtracted by 1 to preserve the BED file 0-based format; the next column contained transcript end position, description data column, transcript ID and strand type. Next, I used bedtools getfasta (v2.29.2)(Quinlan et al. 2010) to retrieve sequences of the up-regulated protein-coding canonical transcripts from GENCODE human reference genome (hg38).

```
#Retrieving sequencing
bedtools getfasta -fi   #Genome FASTA file
                  -bed  #Coordinates BED file
                  -fo   #Output FASTA file
```

```
                    -s    #Force strandedness
                    -name #Use the name field and coordinates for the
   FASTA header
```

6.1.21 Small vault RNA target prediction

To predict RNA targets of svtRNA, I used the miRanda (v3.3a) (Enright et al. 2003) target scanner. This tool computes dynamic-programming alignment and thermodynamics values and provides output information about predicted query sequence and reference target in the human-readable file.

```
#Predicting svtRNA targets
miRanda -sc 150 #Alignment score
        -en -30 #Minimum Free Energy
        -strict #Strict alignment in the seed region (positions 2-8)
```

The Minimum Free Energy (MFE) cut-off value was set to -30 kcal/mol, which is considered sufficient for ideal miRNA-mRNA pairing. For example, the target binding MFE of microRNA let-7 is around -27 kcal/mol. I set the parameter strictly to only consider predictions with perfect sequence complementarity in the seed region (positions 2-8), excluding non-canonical base pairing or gaps in this region.

miRanda svtRNA target prediction	
Query Sequence Name	**Query Sequence**
svtRNA1-1	5'-GACCCGCGGGCGCUCUCCAGUCCUUUU-3'
svtRNA1-2	5'-CGCGGGUGCUUUCCAGCUCUUUU-3'
svtRNA1-3	5'-CGCGGGCGCUCUCCAGCCCUCUU-3'
miR-5587	5'-AUGGUCACCUCCGGGACU-3'

6.1.22 miRanda output processing

The predicted binding locations of the query sequences and the MFE values were retrieved from the miRanda output file. These data were processed into six-columned BED file,

including Target gene ID, Reference sequence start and end positions, Strand, Query sequence name and Transcript ID. All scripts were written in python3.

Annotation of the binding sites was performed using bedtools intersect (v2.29.2)(Quinlan et al. 2010), which overlapped the miRanda output with the custom made feature annotation file '*6.1.19*' to investigate the target sites. I calculated the percentage of binding sites per gene feature taking into account the number of predicted target sites per feature and the length of the feature summing the values together and multiplying the final value by 100 to get the percentage values.

```
#Annotation of predicted target binding sites
bedtools intersect -a   #Feature annotation BED file
                   -b   #Predicted coordinates BED file
                   -wa  #Write the original entry in A for each overlap
                   -wb  #Write the original entry in B for each overlap
                   -loj #Perform a "left outer join"
```

The results were visualised using python3.

6.2 The functional role of small vault RNA1-2

6.2.1 Generating vtRNA1-2 KO cell line

This experiment has been performed by Dr Adele Alagia in Dr Gullerova laboratory.

Cas12a-gRNA ribonucleoprotein complexes containing one sgRNA TTTAGCTCAGCG-GTTACTTCGAGT targeting the VTRNA1-2 gene were transfected in HEK293T using the Alt-R CRISPR-Cas12a system (IDT). Transfected cells were incubated for 48h; then, single cells were sorted into 96-well plates. After two weeks, genomic DNA was extracted from single-cell pellets by phenol:chloroform:isoamyl alcohol (25:24:1) extraction followed by PCR using Phusion High-Fidelity PCR Master Mix (NEB) using primers flanking the vault RNA1-2 gene. PCR products were subjected to agarose gel electrophoresis and visualised by a UV transilluminator. Cells bearing homozygous edited genes were expanded, and deletion of vtRNA1-2 was confirmed by Sanger sequencing.

6.2.2 Cell culture conditions

All cell lines were cultured at 5% CO2 and 37 °C in Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher Scientific).

6.2.3 Proliferation and Colony formation Assays

This experiment has been performed by Dr Adele Alagia in Dr Gullerova laboratory.

WT and vtRNA1-2 KO cells were seeded into 96 well plates at a low density and left to proliferate for 7 days. The MTS assay was carried out using the CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS) (Promega) according to the manufacturer's instructions, and absorbance was measured at 490 nm.
Colony formation assays were carried out to assess the impact of the vtRNA1-2 Knock Out on cell survival. WT and vtRNA1-2 KO cells were seeded into 12 well plates at a density of 1000 cells per well and left for 10 days until visible colonies were formed. Plates were washed in PBS and stained in Crystal Violet staining solution (0.5% crystal violet, 20% ethanol) for 30 minutes. Plates were washed in water, left to dry, and visualised and quantified using the ColonyArea plugin on ImageJ (Schneider et al. 2012).

6.2.4 qRT-PCR validation of svtRNA1-2 predicted target - ERBB4

Total RNA for gene expression analysis was extracted from cultured cells using Trizol and then phenol/chloroform. The RNA was precipitated and then stained by glycoblue. cDNA was prepared by reverse transcription using SuperScript III (1 µg of RNA). The RT-PCR setup was divided into 2 steps. First run was 70°C for 5 mins with the following second step of 55°C for 1 hour and 70°C for 15mins. SYBR green was used to analyse cDNA samples with cycling conditions were 95 °C for 10 minutes and 50 cycles of 95 °C for 15s, 63 °C for 15s and 72°C for the 20s. GAPDH mRNA was used to normalise relative levels of mRNA for tested genes using the 2-ΔΔCt method. All qRT-PCR experiments were

performed using three biological replicates.

Statistical analysis of Real-time PCR (qPCR) data was performed using raw take-off (Ct) values (median of three technical replicates) in GraphPad Prism.

ERBB4 validations qRT-PCR Primer sequences	
Primer Name	**Sequence**
GAPDH_Reverse_Exon	5'-GACGAACATGGGGGCATCAG-3'
GAPDH_Forward	5'-AATTCCATGGCACCGTCAAGG-3'
GAPDH_Intron	5'-CGCCAGACCCTGCACTTTTTAAG-3'
ERBB4_Reverse_Exon	5'-CTGACTGAGAATCGCTGGGC-3'
ERBB4_Forward	5'-GCGACAGGACTTTGGGTCTG-3'
ERBB4_Intron	5'-AGGCTCCCAAACAAAACGCT-3'
PARP12_Reverse_Exon	5'-AAAGCCAGGGGTCGTTCTGA-3'
PARP12_Forward	5'-AACTCATGGCGTTGACCACC-3'
PARP12_Intron	5'-GTGGAACTGGGTTGTTGGCA-3'

6.2.5 vtRNA1-2 KO and WT chromatin-associated RNA samples preparation

This experiment has been performed by Ruth Ketley in Dr Gullerova laboratory.

Chromatin-associated RNA samples for sequencing were prepared according to previous publications (Di Fazio et al. 2022; Nojima et al. 2016) with some minor modifications. Briefly, WT and vtRNA1-2 KO cells were harvested from 15 cm dishes and lysed in HLB+N buffer (10mM Tris-HCl (pH 7.5), 10 mM NaCl, 2.5 mM MgCl2, 0.5% NP-40), and underlaid with HLB+NS buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 2.5 mM MgCl2, 0.5% NP-40, 10% sucrose), before centrifugation at 420g for 5 minutes. Nuclear pellets were then resuspended in NUN1 buffer (20 mM Tris-HCl (pH 7.9), 75 mM NaCl, 0.5 mM EDTA, 50% glycerol). Chromatin was extracted by addition of NUN2 buffer (20

mM HEPES-KOH (pH 7.6), 300 mM NaCl, 0.2 mM EDTA, 7.5 mM MgCl2, 1% NP-40, 1M urea) on ice for 15 minutes with vortexing. Samples were centrifuged at 17,000g for 10 minutes, and chromatin pellets were collected. Chromatin was digested by adding Proteinase K and Turbo DNase (Thermo Fisher). RNA was extracted from the chromatin samples using TRIzol (Invitrogen) and chloroform extraction, followed by the Monarch® Total RNA Miniprep Kit (NEB) according to the manufacturer's instructions and eluted in water.

6.2.6 Fractions validation using Western blot

4x Laemmli buffer (0.2 M Tris–HCl, 8% (w/v) SDS, 40% glycerol, 20% (v/v) β-mercaptoethanol, 0.005% bromophenol blue) was used to treat the whole cell and chromatin extracts following 95^{o}C incubation for 5 min and sonication. Samples separation was performed using mini-PROTEAN® TGX™ gels (Bio-512 Rad Laboratories) followed by transfer onto nitrocellulose membrane (Protran, GE Healthcare) for probing I used Tubulin (55 kDa) and H3 (15 kDa) antibodies. The resulting signals were visualised on Amersham™ Hyperfilm™ ECL(Cytiva) after the application of Pierce™ ECL Western Blotting Substrate (ThermoScientific).

6.2.7 ChrRNA Sequencing data analysis of WT and vtRNA1-2 KO samples

Sample library preparation and sequencing were carried out in The Oxford Genomics Centre, generating paired-end libraries with 150 bases long reads on the Illumina NovaSeq machine in triplicates. First, the adapter sequences were trimmed using cutadapt (v2.10). Removing specific adapter sequences and preserving reads longer than 10 bases after processing with trimmed low-quality reads from 3' and 5' ends and a max error rate of 0.05.

```
# Cutadapt adapter sequences trimming
cutadapt -e 0.05  #Maximum error rate
         -m 10    #Minimal read length
         -j 8     #CPU cores
         -q 15,10 #Trim low-quality ends from 3' and 5' end of reads
```

```
        -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA #Adapter1
        -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT #Adapter2
        -o file_1_trimmed.fastq.gz #Output file 1
        -p file_2_trimmed.fastq.gz #Output file 2
        file_1.fastq.gz file_2.fastq.gz #Input file 1 and Input file 2
```

The trimmed and filtered reads were aligned to the human reference genome (hg38) with annotation using STAR (v2.7.3a) using the same settings as in '*chromatin-associated RNA sequencing data pre-processing*' section. This equally applies to the subsequent reads quantification.

In R-Studio/R version (4.2.0), any gene with more than 10 reads in at least 2 replicates per condition was assigned to the count matrix, which was normalised using DESeq2 (v1.36.0). To identify significantly differentially expressed genes, I used adj. p-value of < 0.001 and Log_2Fold Change > 1 | < -1. The sequences for target prediction and further downstream analyses were performed according to '*chromatin-associated RNA sequencing data pre-processing*' section.

6.2.8 GO and GSEA

The gene ontology enrichment of significantly up and down-regulated genes was performed using ShinyGO (v0.76) with the same settings as in '*GO*' section.

To gain a deeper insight into genome-wide expression profiles, biological processes, and pathways underlying vtRNA1-2 KO, I used Gene Set Enrichment Analysis (GSEA) (v4.1.0) tool (Mootha et al. 2003; Subramanian et al. 2005). I ranked the DESeq2 (v1.36.0) output matrix using the Wald statistics column with all genes above or below adj. p-value. The final list of genes had 22761 entries. GSEA tool was run using the pre-ranked option. The ranked file contained two columns; the first column included 22761 gene names, and the second one Wald statistics values based on which the genes were ranked. GSEA tool requires another file describing a number of samples and a number of compared groups with related condition labels. I performed 2 runs, first using the Hallmark gene sets representing specific, well-defined biological states or processes

and displaying coherent expression and the second run using the KEGG database. The ranked gene list is compared to a gene set, i.e., a list of genes associated with a particular biological process, gene ontology terms, molecular function or pathway. The enrichment score (ES) indicates the degree by which a gene set is over-represented at the extremes of the ranked list. Specific conditions were selected, such as:

Collapse/Remap to gene symbols = No_Collapse

Enrichment statistic = weighted

Max size: exclude larger sets = 500

Min size: exclude smaller sets = 15

The exact same settings were applied to search the KEGG pathways. I used ES and p-value scores to select biological processes and pathways that were found up-regulated in the vtRNA1-2 KO condition.

6.2.9 STRING analysis

STRING (v11.5) (Szklarczyk et al. 2021) was utilised to compare characteristics of svtRNA1-2 predicted targets. I applied the STRING analysis tool (Web version) to understand the direct and indirect interactions among them and their common features and enriched pathways. I used the Ensembl gene IDs of 29 predicted targets as the input parameter. I selected the enriched molecular functions from the constructed network based on PPI value (Interaction Enrichment) and false discovery rate per pathway.

6.2.10 Ingenuity Pathway Analysis (IPA) (Trial version)

I used IPA to obtain a different perspective on the characteristics and functional relevance of the svtRNA1-2 predicted targets. DESeq2 (v1.36.0) output values for significantly (adj. p-value < 0.001) differentially expressed genes in vtRNA1-2 KO were retrieved and submitted to the IPA for further analysis. I selected the Regulator Effects feature to investigate the upstream and downstream regulatory effects of ERBB4 to interpret the phenotypic outcome of vtRNA1-2 KO. Furthermore, I retrieved the findings provided by

IPA about molecular localisation of the target genes within the cell were visualised using python3.

6.2.11 svtRNA-target binding pattern investigation

To understand the mechanism of svtRNA1-2 binding, I generated different *in silico* variants of vtRNA1-2 sequence and used them to search reference sequences (up-regulated genes in vtRNA1-2 KO vs WT) using the miRanda (v3.3a) target scanner with the same settings as described in '*Small vault RNA target prediction*' section. The 5' and 3' ends were shuffled (randomised) while considering the sequence composition.

miRanda scan for binding pattern	
Query vault RNA	**Query Sequence**
svtRNA1-1 wild type	5'-GACCCGCGGGCGCUCUCCAGUCCUUUU-3'
svtRNA1-2 wild type	5'-CGCGGGUGCUUUCCAGCUCUUUU-3'
svtRNA1-2 swapped	5'-CUCCAGCUCUUGCGGGUGCUUUU-3'
svtRNA1-2 shuffled 5'	5'-CAAUUUGAAACUCCAGCUCUUUU-3'
svtRNA1-2 shuffled 3'	5'-CGCGGGUGCUUAAACUACCUUUU-3'

6.2.12 Target validation of ERBB4 in MCF7 cells

This experiment has been performed by Ruth Ketley in Dr Gullerova laboratory.

MCF-7 cells were reverse-transfected in 0.01% poly-L-Lys coated 6-well plates at a density of 1.2 × 106 cells per well with 100 nM of single-stranded svtRNA1-2 or 60 nM of double-stranded svtRNA1-2, using 9 µL of Lipofectamine RNAiMAX reagent. Following the manufacturer's instructions, 24h after transfection, cells were split, and parallel cultures were maintained for 24h without further treatments. qRT-PCR was performed according to protocol as in '*qRT-PCR validation of svtRNA1-2 predicted target - ERBB4*' section.

6.2.13 Total RNA sequencing sample preparation

This experiment has been performed partly by Dr Adele Alagia and myself in Dr Gullerova laboratory.

Total RNA from HEK293T cells (vtRNA1-2 KO and WT) was extracted using TRIzol (Life Technologies) and then purified by spin column purification (RNeasy mini kit; Qiagen), according to the manufacturer's instructions. RNA was quantified using a Nanodrop (Thermo Fisher), and its integrity was assessed on a Bioanalyzer (Agilent) using the RNA 6000 Nano kit. The RIN scores were at least nine for samples used in the total RNA-seq, with three biological replicates for WT and four replicates for vtRNA1-2 KO. The libraries were constructed using the TruSeq Stranded Total RNA with Ribo-Zero Gold rRNA removal kit (Illumina). Library purity, fragment size, and concentration were assessed using the Bioanalyzer Agilent High Sensitivity DNA kit. Prepared libraries with a peak area (300 bp) were individually barcoded, pooled, and sequenced on an Illumina NovaSeq platform.

6.2.14 Total RNA sequencing data analysis

In the downstream analysis of total RNA-seq, first I trimmed Illumina TruSeq adapter sequences together with reads shorter than 10 bases. Moreover, reads with trimmed low-quality from 3' and 5' ends and with a max error rate of 0.05 were also removed using cutadapt (v2.10).

```
# Cutadapt adapter sequences trimming
cutadapt -e 0.05  #Maximum error rate
         -m 10    #Discard reads shorter than 10 bases
         -j 8     #CPU cores
         -q 15,10 #Trim low-quality ends from 3' and 5' end of reads
         -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA #Adapter1
         -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT #Adapter2
         -o file_1_trimmed.fastq.gz #Output file 1
         -p file_2_trimmed.fastq.gz #Output file2
         file_1.fastq.gz file_2.fastq.gz #Input file 1 and 2
```

The obtained reads were then mapped to the human reference genome with annotation (hg38) using STAR (v2.7.3a) with the settings previously described for chRNA-seq (vtRNA1-2 KO vs WT). After the first alignment, I noticed low number of uniquely mapped reads and a high percentage of multi-mapping reads. Further investigation revealed rRNA sample contamination. Therefore, I obtained the rRNA sequence in FASTA format and used bbduk.sh funtion from BBTools (Bushnell 2016) to remove the rRNA contamination. On average, 70% of the reads aligned to the reference rRNA sequence and therefore were removed from the FASTQ files. After the filtration, I mapped the reads again to the human reference genome (hg38) and gained on average 80% of uniquely mapped reads. These reads were quantified with featureCounts (v2.0.0) from the Subread package, counting reads that fall into the boundaries of exon coordinates from the annotation file. However, the percentage of assigned reads was on average only 40%. The final count matrix was then imported into R-Studio/R version (4.2.0) for statistical analysis.

```
# Reads quantification
featureCounts -a gencode.annotation.gtf #Genome annotation file
              -T 8       #Number of threads
              -t exon    #Feature type included for read counting
              -g gene_id #Attribute type used to group features
              -s 2       #Reversely stranded library
              --verbose  #Provide additional details about this run
              -o         #Output file
              #Input alignment files
```

In R-Studio/R version (4.2.0)), any gene with less than 5 counts in less than 2 replicates per condition was removed. The count matrix was then normalised using DESeq2 (v1.36.0). To identify significantly differentially expressed genes, I used adj. p-value threshold of <0.05 and Log_2Fold Change $> 1 \mid < -1$. The significance thresholds were adapted to the low end-number of reads. All the subsequent data plotting was performed in R (v4.2.0).

Study	Technique	GEO ID	Cell type	Publication
Drosha KO vs WT	small RNA-seq	GSE77989	HCT116	Y. K. Kim et al. 2016
Drosha KD vs WT	small RNA-seq	GSE55333	HEK293	Rybak-Wolf et al. 2014
Drosha	fCLIP	GSE93653	HEK293T	B. Kim et al. 2019
Dicer KD vs WT	small RNA-seq	GSE126751	HEK293	Di Fazio et al. 2022
Dicer KD vs WT	small RNA-seq	GSE55333	HEK293	Rybak-Wolf et al. 2014
WT	small RNA-seq	GSE119661	Multiple mouse tissues	Kern et al. 2020
AGO1-4	PAR-CLIP	GSE21918	HEK293	Macias et al. 2012
AGO2/3	RIP-seq	GSE55333	HEK293	Rybak-Wolf et al. 2014
nuclear and cytoplasmic AGO2	RIP-seq	GSE66665	MCF-7	Fernandez-Valverde et al.
Dicer	PAR-CLIP	GSE55333	HEK293	Rybak-Wolf et al. 2014
AGO2/Dicer KD vs WT	chrRNA-seq	GSE126751	HEK293	Di Fazio et al. 2022
vtRNA1-2 KO vs WT	chrRNA-seq	GSE198298	HEK293	Terenova et al.
vtRNA1-2 KO vs WT	total RNA-seq	This work	HEK293	Terenova et al.

Table 1: Analysed datasets

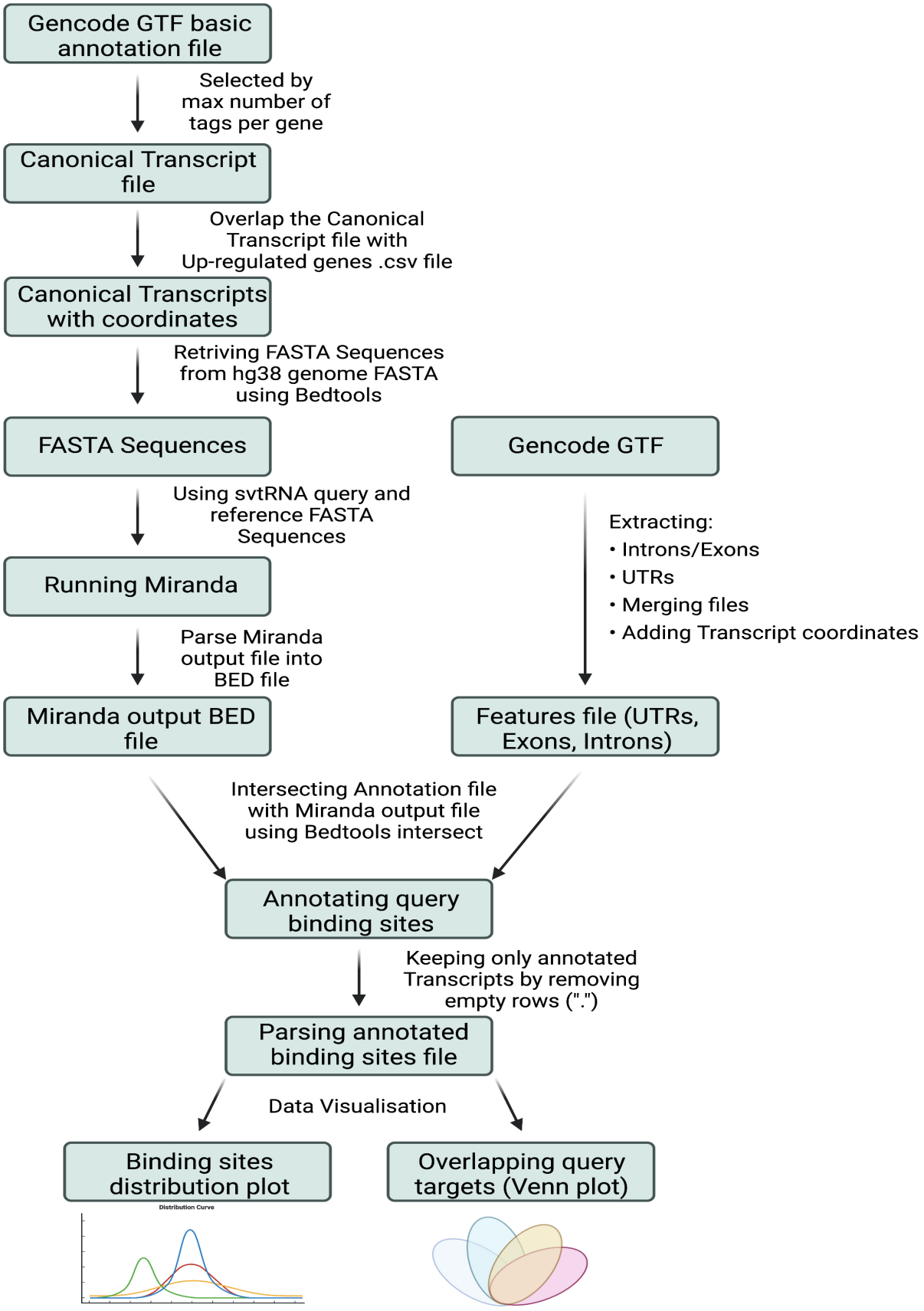


Figure 34: *Schematic representation of chrRNA-sequencing data analysis pipeline.*

References

Agris, P. F., H. Koh, and D. Soll (1973). "The effect of growth temperatures on the in vivo ribose methylation of Bacillus stearothermophilus transfer RNA". In: *Arch Biochem Biophys* 154.1, pp. 277–82. ISSN: 0003-9861 (Print) 0003-9861 (Linking). DOI: 10.1016/0003-9861(73)90058-1. URL: https://www.ncbi.nlm.nih.gov/pubmed/4689778.

Alagia, A. and R. Eritja (2016). "siRNA and RNAi optimization". In: *Wiley Interdiscip Rev RNA* 7.3, pp. 316–29. ISSN: 1757-7012 (Electronic) 1757-7004 (Linking). DOI: 10.1002/wrna.1337. URL: https://www.ncbi.nlm.nih.gov/pubmed/26840434.

Alexandrov, A., M. R. Martzen, and E. M. Phizicky (2002). "Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA". In: *RNA* 8.10, pp. 1253–66. ISSN: 1355-8382 (Print) 1355-8382 (Linking). DOI: 10.1017/s1355838202024019. URL: https://www.ncbi.nlm.nih.gov/pubmed/12403464.

Alexandrov, A. et al. (2006). "Rapid tRNA decay can result from lack of nonessential modifications". In: *Mol Cell* 21.1, pp. 87–96. ISSN: 1097-2765 (Print) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2005.10.036. URL: https://www.ncbi.nlm.nih.gov/pubmed/16387656.

Allo, M. et al. (2009). "Control of alternative splicing through siRNA-mediated transcriptional gene silencing". In: *Nat Struct Mol Biol* 16.7, pp. 717–24. ISSN: 1545-9985 (Electronic) 1545-9985 (Linking). DOI: 10.1038/nsmb.1620. URL: https://www.ncbi.nlm.nih.gov/pubmed/19543290.

Ameyar-Zazoua, M. et al. (2012). "Argonaute proteins couple chromatin silencing to alternative splicing". In: *Nat Struct Mol Biol* 19.10, pp. 998–1004. ISSN: 1545-9985 (Electronic) 1545-9985 (Linking). DOI: 10.1038/nsmb.2373. URL: https://www.ncbi.nlm.nih.gov/pubmed/22961379.

Amort, M. et al. (2015). "Expression of the vault RNA protects cells from undergoing apoptosis". In: *Nat Commun* 6, p. 7030. ISSN: 2041-1723 (Electronic) 2041-1723 (Linking). DOI: 10.1038/ncomms8030. URL: https://www.ncbi.nlm.nih.gov/pubmed/25952297.

Angelova, M. T. et al. (2020). "tRNA 2'-O-methylation by a duo of TRM7/FTSJ1 proteins modulates small RNA silencing in Drosophila". In: *Nucleic Acids Res* 48.4, pp. 2050–2072. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkaa002. URL: https://www.ncbi.nlm.nih.gov/pubmed/31943105.

Antczak, M. et al. (2016). "New functionality of RNAComposer: an application to shape the axis of miR160 precursor structure". In: *Acta Biochim Pol* 63.4, pp. 737–744. ISSN: 1734-154X (Electronic) 0001-527X (Linking). DOI: 10.18388/abp.2016_1329. URL: https://www.ncbi.nlm.nih.gov/pubmed/27741327.

Arteaga, C. L. and J. A. Engelman (2014). "ERBB receptors: from oncogene discovery to basic science to mechanism-based cancer therapeutics". In: *Cancer Cell* 25.3, pp. 282–303. ISSN: 1878-3686 (Electronic) 1535-6108 (Linking). DOI: 10.1016/j.ccr.2014.02.025. URL: https://www.ncbi.nlm.nih.gov/pubmed/24651011.

Arts, G. J., M. Fornerod, and I. W. Mattaj (1998). "Identification of a nuclear export receptor for tRNA". In: *Curr Biol* 8.6, pp. 305–14. ISSN: 0960-9822 (Print) 0960-9822 (Linking). DOI: 10.1016/s0960-9822(98)70130-7. URL: https://www.ncbi.nlm.nih.gov/pubmed/9512417.

Babiarz, J. E. et al. (2008). "Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs". In: *Genes Dev* 22.20, pp. 2773–85. ISSN: 0890-9369 (Print) 0890-9369 (Linking). DOI: 10.1101/gad.1705308. URL: https://www.ncbi.nlm.nih.gov/pubmed/18923076.

Bachmann, M. et al. (1989). "The La antigen shuttles between the nucleus and the cytoplasm in CV-1 cells". In: *Mol Cell Biochem* 85.2, pp. 103–14. ISSN: 0300-8177 (Print) 0300-8177 (Linking). DOI: 10.1007/BF00577106. URL: https://www.ncbi.nlm.nih.gov/pubmed/2657384.

Barrett, T. et al. (2013). "NCBI GEO: archive for functional genomics data sets–update". In: *Nucleic Acids Res* 41.Database issue, pp. D991–5. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gks1193. URL: https://www.ncbi.nlm.nih.gov/pubmed/23193258.

Bartel, D. P. (2018). "Metazoan MicroRNAs". In: *Cell* 173.1, pp. 20–51. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2018.03.006. URL: https://www.ncbi.nlm.nih.gov/pubmed/29570994.

Becker, W. R. et al. (2019). "High-Throughput Analysis Reveals Rules for Target RNA Binding and Cleavage by AGO2". In: *Mol Cell* 75.4, 741–755 e11. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2019.06.012. URL: https://www.ncbi.nlm.nih.gov/pubmed/31324449.

Behm-Ansmant, I. et al. (2003). "The Saccharomyces cerevisiae U2 snRNA:pseudouridine-synthase Pus7p is a novel multisite-multisubstrate RNA:Psi-synthase also acting on tRNAs". In: *RNA* 9.11, pp. 1371–82. ISSN: 1355-8382 (Print) 1355-8382 (Linking). DOI: 10.1261/rna.5520403. URL: https://www.ncbi.nlm.nih.gov/pubmed/14561887.

Behm-Ansmant, I. et al. (2006). "mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes". In: *Genes Dev* 20.14, pp. 1885–98. ISSN: 0890-9369 (Print) 0890-9369 (Linking). DOI: 10.1101/gad.1424106. URL: https://www.ncbi.nlm.nih.gov/pubmed/16815998.

Berger, W. et al. (2009). "Vaults and the major vault protein: novel roles in signal pathway regulation and immunity". In: *Cell Mol Life Sci* 66.1, pp. 43–61. ISSN: 1420-9071 (Electronic) 1420-682X (Linking). DOI: 10.1007/s00018-008-8364-z. URL: https://www.ncbi.nlm.nih.gov/pubmed/18759128.

Blanco, S. et al. (2014). "Aberrant methylation of tRNAs links cellular stress to neurodevelopmental disorders". In: *EMBO J* 33.18, pp. 2020–39. ISSN: 1460-2075 (Electronic) 0261-4189 (Linking). DOI: 10.15252/embj.201489282. URL: https://www.ncbi.nlm.nih.gov/pubmed/25063673.

Bracher, L. et al. (2020). "Human vtRNA1-1 Levels Modulate Signaling Pathways and Regulate Apoptosis in Human Cancer Cells". In: *Biomolecules* 10.4. ISSN: 2218-273X (Electronic) 2218-273X (Linking). DOI: 10.3390/biom10040614. URL: https://www.ncbi.nlm.nih.gov/pubmed/32316166.

Brancati, G. and H. Grosshans (2018). "An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity". In: *Nucleic Acids Res* 46.7,

pp. 3259–3269. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gky201. URL: https://www.ncbi.nlm.nih.gov/pubmed/29897601.

Broughton, J. P. et al. (2016). "Pairing beyond the Seed Supports MicroRNA Targeting Specificity". In: *Mol Cell* 64.2, pp. 320–333. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2016.09.004. URL: https://www.ncbi.nlm.nih.gov/pubmed/27720646.

Buscher, M., R. Horos, and M. W. Hentze (2020). "'High vault-age': non-coding RNA control of autophagy". In: *Open Biol* 10.2, p. 190307. ISSN: 2046-2441 (Electronic) 2046-2441 (Linking). DOI: 10.1098/rsob.190307. URL: https://www.ncbi.nlm.nih.gov/pubmed/32070232.

Buscher, M. et al. (2022). "Vault RNA1-1 riboregulates the autophagic function of p62 by binding to lysine 7 and arginine 21, both of which are critical for p62 oligomerization". In: *RNA* 28.5, pp. 742–755. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.079129.122. URL: https://www.ncbi.nlm.nih.gov/pubmed/35210358.

Bushnell, B. (2016). "BBTools". In: *https://jgi.doe.gov/data-and-tools/bbtools*.

Cao, J. et al. (2013). "DNA methylation-mediated repression of miR-886-3p predicts poor outcome of human small cell lung cancer". In: *Cancer Res* 73.11, pp. 3326–35. ISSN: 1538-7445 (Electronic) 0008-5472 (Linking). DOI: 10.1158/0008-5472.CAN-12-3055. URL: https://www.ncbi.nlm.nih.gov/pubmed/23592755.

Castanotto, D. et al. (2005). "Short hairpin RNA-directed cytosine (CpG) methylation of the RASSF1A gene promoter in HeLa cells". In: *Mol Ther* 12.1, pp. 179–83. ISSN: 1525-0016 (Print) 1525-0016 (Linking). DOI: 10.1016/j.ymthe.2005.03.003. URL: https://www.ncbi.nlm.nih.gov/pubmed/15963934.

Chatterjee, K. et al. (2018). "tRNA dynamics between the nucleus, cytoplasm and mitochondrial surface: Location, location, location". In: *Biochim Biophys Acta Gene Regul Mech* 1861.4, pp. 373–386. ISSN: 1874-9399 (Print) 1874-9399 (Linking). DOI: 10.1016/j.bbagrm.2017.11.007. URL: https://www.ncbi.nlm.nih.gov/pubmed/29191733.

Cheloufi, S. et al. (2010). "A dicer-independent miRNA biogenesis pathway that requires Ago catalysis". In: *Nature* 465.7298, pp. 584–9. ISSN: 1476-4687 (Electronic) 0028-0836

(Linking). DOI: 10.1038/nature09092. URL: https://www.ncbi.nlm.nih.gov/pubmed/20424607.

Chen, J. et al. (2018). "Vault RNAs partially induces drug resistance of human tumor cells MCF-7 by binding to the RNA/DNA-binding protein PSF and inducing oncogene GAGE6". In: *PLoS One* 13.1, e0191325. ISSN: 1932-6203 (Electronic) 1932-6203 (Linking). DOI: 10.1371/journal.pone.0191325. URL: https://www.ncbi.nlm.nih.gov/pubmed/29346433.

Chugani, D. C., L. H. Rome, and N. L. Kedersha (1993). "Evidence that vault ribonucleoprotein particles localize to the nuclear pore complex". In: *J Cell Sci* 106 (Pt 1), pp. 23–9. ISSN: 0021-9533 (Print) 0021-9533 (Linking). URL: https://www.ncbi.nlm.nih.gov/pubmed/8270627.

Ciechanowska, K., M. Pokornowska, and A. Kurzynska-Kokorniak (2021). "Genetic Insight into the Domain Structure and Functions of Dicer-Type Ribonucleases". In: *Int J Mol Sci* 22.2. ISSN: 1422-0067 (Electronic) 1422-0067 (Linking). DOI: 10.3390/ijms22020616. URL: https://www.ncbi.nlm.nih.gov/pubmed/33435485.

Cole, C. et al. (2009). "Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs". In: *RNA* 15.12, pp. 2147–60. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.1738409. URL: https://www.ncbi.nlm.nih.gov/pubmed/19850906.

Cook, A. G. et al. (2009). "Structures of the tRNA export factor in the nuclear and cytosolic states". In: *Nature* 461.7260, pp. 60–5. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature08394. URL: https://www.ncbi.nlm.nih.gov/pubmed/19680239.

Corcoran, D. L. et al. (2011). "PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data". In: *Genome Biol* 12.8, R79. ISSN: 1474-760X (Electronic) 1474-7596 (Linking). DOI: 10.1186/gb-2011-12-8-r79. URL: https://www.ncbi.nlm.nih.gov/pubmed/21851591.

Cursons, J. et al. (2018). "Combinatorial Targeting by MicroRNAs Co-ordinates Post-transcriptional Control of EMT". In: *Cell Syst* 7.1, 77–91 e7. ISSN: 2405-4712 (Print)

2405-4712 (Linking). DOI: 10.1016/j.cels.2018.05.019. URL: https://www.ncbi.nlm.nih.gov/pubmed/30007539.

Davis, D. R. (1995). "Stabilization of RNA stacking by pseudouridine". In: *Nucleic Acids Res* 23.24, pp. 5020–6. ISSN: 0305-1048 (Print) 0305-1048 (Linking). DOI: 10.1093/nar/23.24.5020. URL: https://www.ncbi.nlm.nih.gov/pubmed/8559660.

Deng, L. et al. (2022). "Long Non-coding RNA ANRIL Downregulation Alleviates Neuroinflammation in an Ischemia Stroke Model via Modulation of the miR-671-5p/NF-kappaB Pathway". In: *Neurochem Res* 47.7, pp. 2002–2015. ISSN: 1573-6903 (Electronic) 0364-3190 (Linking). DOI: 10.1007/s11064-022-03585-1. URL: https://www.ncbi.nlm.nih.gov/pubmed/35359242.

Dey, P., M. A. Soyer, and B. K. Dey (2022). "MicroRNA-24-3p promotes skeletal muscle differentiation and regeneration by regulating HMGA1". In: *Cell Mol Life Sci* 79.3, p. 170. ISSN: 1420-9071 (Electronic) 1420-682X (Linking). DOI: 10.1007/s00018-022-04168-7. URL: https://www.ncbi.nlm.nih.gov/pubmed/35238991.

Di Fazio, A. et al. (2022). "Dicer dependent tRNA derived small RNAs promote nascent RNA silencing". In: *Nucleic Acids Res* 50.3, pp. 1734–1752. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkac022. URL: https://www.ncbi.nlm.nih.gov/pubmed/35048990.

Ding, K. et al. (2018). "Solution Structures of Engineered Vault Particles". In: *Structure* 26.4, 619–626 e3. ISSN: 1878-4186 (Electronic) 0969-2126 (Linking). DOI: 10.1016/j.str.2018.02.014. URL: https://www.ncbi.nlm.nih.gov/pubmed/29551289.

Dobin, A. et al. (2013). "STAR: ultrafast universal RNA-seq aligner". In: *Bioinformatics* 29.1, pp. 15–21. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking). DOI: 10.1093/bioinformatics/bts635. URL: https://www.ncbi.nlm.nih.gov/pubmed/23104886.

Doyle, M. et al. (2013). "The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal". In: *RNA* 19.9, pp. 1238–52. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.039255.113. URL: https://www.ncbi.nlm.nih.gov/pubmed/23882114.

Drexler, H. L., K. Choquet, and L. S. Churchman (2020). "Splicing Kinetics and Coordination Revealed by Direct Nascent RNA Sequencing through Nanopores". In: *Mol Cell*

77.5, 985–998 e8. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2019.11.017. URL: https://www.ncbi.nlm.nih.gov/pubmed/31839405.

Durant, P. C. and D. R. Davis (1999). "Stabilization of the anticodon stem-loop of tRNALys,3 by an A+-C base-pair and by pseudouridine". In: *J Mol Biol* 285.1, pp. 115–31. ISSN: 0022-2836 (Print) 0022-2836 (Linking). DOI: 10.1006/jmbi.1998.2297. URL: https://www.ncbi.nlm.nih.gov/pubmed/9878393.

Edgar, R., M. Domrachev, and A. E. Lash (2002). "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository". In: *Nucleic Acids Res* 30.1, pp. 207–10. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/30.1.207. URL: https://www.ncbi.nlm.nih.gov/pubmed/11752295.

Elkordy, A. et al. (2018). "Stress-induced tRNA cleavage and tiRNA generation in rat neuronal PC12 cells". In: *J Neurochem* 146.5, pp. 560–569. ISSN: 1471-4159 (Electronic) 0022-3042 (Linking). DOI: 10.1111/jnc.14321. URL: https://www.ncbi.nlm.nih.gov/pubmed/29431851.

Enright, A. J. et al. (2003). "MicroRNA targets in Drosophila". In: *Genome Biol* 5.1, R1. ISSN: 1474-760X (Electronic) 1474-7596 (Linking). DOI: 10.1186/gb-2003-5-1-r1. URL: https://www.ncbi.nlm.nih.gov/pubmed/14709173.

Feder, M. et al. (2003). "Molecular phylogenetics of the RrmJ/fibrillarin superfamily of ribose 2'-O-methyltransferases". In: *Gene* 302.1-2, pp. 129–38. ISSN: 0378-1119 (Print) 0378-1119 (Linking). DOI: 10.1016/s0378-1119(02)01097-1. URL: https://www.ncbi.nlm.nih.gov/pubmed/12527203.

Ferro, I. et al. (2022). "The human vault RNA enhances tumorigenesis and chemoresistance through the lysosome in hepatocellular carcinoma". In: *Autophagy* 18.1, pp. 191–203. ISSN: 1554-8635 (Electronic) 1554-8627 (Linking). DOI: 10.1080/15548627.2021.1922983. URL: https://www.ncbi.nlm.nih.gov/pubmed/33960270.

Fok, V., K. Friend, and J. A. Steitz (2006). "Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes nucleocytoplasmic shuttling". In: *J Cell Biol* 173.3, pp. 319–25. ISSN: 0021-9525 (Print) 0021-9525 (Linking). DOI: 10.1083/jcb.200601026. URL: https://www.ncbi.nlm.nih.gov/pubmed/16682524.

Fort, R. S. and M. A. Duhagon (2021). "Pan-cancer chromatin analysis of the human vtRNA genes uncovers their association with cancer biology". In: *F1000Res* 10, p. 182. ISSN: 2046-1402 (Electronic) 2046-1402 (Linking). DOI: 10.12688/f1000research.28510.2. URL: https://www.ncbi.nlm.nih.gov/pubmed/34354812.

Fort, R. S. et al. (2018). "Nc886 is epigenetically repressed in prostate cancer and acts as a tumor suppressor through the inhibition of cell growth". In: *BMC Cancer* 18.1, p. 127. ISSN: 1471-2407 (Electronic) 1471-2407 (Linking). DOI: 10.1186/s12885-018-4049-7. URL: https://www.ncbi.nlm.nih.gov/pubmed/29394925.

Fort, R. S. et al. (2020). "vtRNA2-1/nc886 Produces a Small RNA That Contributes to Its Tumor Suppression Action through the microRNA Pathway in Prostate Cancer". In: *Noncoding RNA* 6.1. ISSN: 2311-553X (Electronic) 2311-553X (Linking). DOI: 10.3390/ncrna6010007. URL: https://www.ncbi.nlm.nih.gov/pubmed/32093270.

Frank, F., N. Sonenberg, and B. Nagar (2010). "Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2". In: *Nature* 465.7299, pp. 818–22. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature09039. URL: https://www.ncbi.nlm.nih.gov/pubmed/20505670.

Frankish, A. et al. (2021). "Gencode 2021". In: *Nucleic Acids Res* 49.D1, pp. D916–D923. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkaa1087. URL: https://www.ncbi.nlm.nih.gov/pubmed/33270111.

Freude, K. et al. (2004). "Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation". In: *Am J Hum Genet* 75.2, pp. 305–9. ISSN: 0002-9297 (Print) 0002-9297 (Linking). DOI: 10.1086/422507. URL: https://www.ncbi.nlm.nih.gov/pubmed/15162322.

Fu, H. et al. (2009). "Stress induces tRNA cleavage by angiogenin in mammalian cells". In: *FEBS Lett* 583.2, pp. 437–42. ISSN: 1873-3468 (Electronic) 0014-5793 (Linking). DOI: 10.1016/j.febslet.2008.12.043. URL: https://www.ncbi.nlm.nih.gov/pubmed/19114040.

Gagnon, K. T. et al. (2014). "RNAi factors are present and active in human cell nuclei". In: *Cell Rep* 6.1, pp. 211–21. ISSN: 2211-1247 (Electronic). DOI: 10.1016/j.celrep.2013.12.013. URL: https://www.ncbi.nlm.nih.gov/pubmed/24388755.

Galbiati, E. et al. (2018). "A fast and straightforward procedure for vault nanoparticle purification and the characterization of its endocytic uptake". In: *Biochim Biophys Acta Gen Subj* 1862.10, pp. 2254–2260. ISSN: 0304-4165 (Print) 0304-4165 (Linking). DOI: 10.1016/j.bbagen.2018.07.018. URL: https://www.ncbi.nlm.nih.gov/pubmed/30036602.

Gallo, S. et al. (2022). "Small but Powerful: The Human Vault RNAs as Multifaceted Modulators of Pro-Survival Characteristics and Tumorigenesis". In: *Cancers (Basel)* 14.11. ISSN: 2072-6694 (Print) 2072-6694 (Linking). DOI: 10.3390/cancers14112787. URL: https://www.ncbi.nlm.nih.gov/pubmed/35681764.

Ge, S. X., D. Jung, and R. Yao (2020). "ShinyGO: a graphical gene-set enrichment tool for animals and plants". In: *Bioinformatics* 36.8, pp. 2628–2629. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking). DOI: 10.1093/bioinformatics/btz931. URL: https://www.ncbi.nlm.nih.gov/pubmed/31882993.

Gebetsberger, J. et al. (2012). "tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in Haloferax volcanii". In: *Archaea* 2012, p. 260909. ISSN: 1472-3654 (Electronic) 1472-3646 (Linking). DOI: 10.1155/2012/260909. URL: https://www.ncbi.nlm.nih.gov/pubmed/23326205.

Giraldez, A. J. et al. (2006). "Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs". In: *Science* 312.5770, pp. 75–9. ISSN: 1095-9203 (Electronic) 0036-8075 (Linking). DOI: 10.1126/science.1122689. URL: https://www.ncbi.nlm.nih.gov/pubmed/16484454.

Gong, M. et al. (2021). "DDX21 interacts with nuclear AGO2 and regulates the alternative splicing of SMN2". In: *Biosci Biotechnol Biochem* 85.2, pp. 272–279. ISSN: 1347-6947 (Electronic) 0916-8451 (Linking). DOI: 10.1093/bbb/zbaa029. URL: https://www.ncbi.nlm.nih.gov/pubmed/33604619.

Gopinath, S. C., R. Wadhwa, and P. K. Kumar (2010). "Expression of noncoding vault RNA in human malignant cells and its importance in mitoxantrone resistance". In: *Mol Cancer Res* 8.11, pp. 1536–46. ISSN: 1557-3125 (Electronic) 1541-7786 (Linking). DOI: 10.1158/1541-7786.MCR-10-0242. URL: https://www.ncbi.nlm.nih.gov/pubmed/20881010.

Gopinath, S. C. et al. (2005). "Human vault-associated non-coding RNAs bind to mitoxantrone, a chemotherapeutic compound". In: *Nucleic Acids Res* 33.15, pp. 4874–81. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gki809. URL: https://www.ncbi.nlm.nih.gov/pubmed/16150923.

Grimson, A. et al. (2007). "MicroRNA targeting specificity in mammals: determinants beyond seed pairing". In: *Mol Cell* 27.1, pp. 91–105. ISSN: 1097-2765 (Print) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2007.06.017. URL: https://www.ncbi.nlm.nih.gov/pubmed/17612493.

Gu, J. et al. (2012). "Potential roles of N-glycosylation in cell adhesion". In: *Glycoconj J* 29.8-9, pp. 599–607. ISSN: 1573-4986 (Electronic) 0282-0080 (Linking). DOI: 10.1007/s10719-012-9386-1. URL: https://www.ncbi.nlm.nih.gov/pubmed/22565826.

Guzzi, N. et al. (2018). "Pseudouridylation of tRNA-Derived Fragments Steers Translational Control in Stem Cells". In: *Cell* 173.5, 1204–1216 e26. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2018.03.008. URL: https://www.ncbi.nlm.nih.gov/pubmed/29628141.

Ha, M. and V. N. Kim (2014). "Regulation of microRNA biogenesis". In: *Nat Rev Mol Cell Biol* 15.8, pp. 509–24. ISSN: 1471-0080 (Electronic) 1471-0072 (Linking). DOI: 10.1038/nrm3838. URL: https://www.ncbi.nlm.nih.gov/pubmed/25027649.

Hahne, J. C., A. Lampis, and N. Valeri (2021). "Vault RNAs: hidden gems in RNA and protein regulation". In: *Cell Mol Life Sci* 78.4, pp. 1487–1499. ISSN: 1420-9071 (Electronic) 1420-682X (Linking). DOI: 10.1007/s00018-020-03675-9. URL: https://www.ncbi.nlm.nih.gov/pubmed/33063126.

Hamill, D. R. and K. A. Suprenant (1997). "Characterization of the sea urchin major vault protein: a possible role for vault ribonucleoprotein particles in nucleocytoplasmic transport". In: *Dev Biol* 190.1, pp. 117–28. ISSN: 0012-1606 (Print) 0012-1606 (Linking). DOI: 10.1006/dbio.1997.8676. URL: https://www.ncbi.nlm.nih.gov/pubmed/9331335.

Han, T. et al. (2009). "26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans". In: *Proc Natl Acad Sci U S A* 106.44, pp. 18674–9.

ISSN: 1091-6490 (Electronic) 0027-8424 (Linking). DOI: 10.1073/pnas.0906378106. URL: https://www.ncbi.nlm.nih.gov/pubmed/19846761.

Hasler, D. et al. (2016). "The Lupus Autoantigen La Prevents Mis-channeling of tRNA Fragments into the Human MicroRNA Pathway". In: *Mol Cell* 63.1, pp. 110–24. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2016.05.026. URL: https://www.ncbi.nlm.nih.gov/pubmed/27345152.

Haussecker, D. et al. (2010). "Human tRNA-derived small RNAs in the global regulation of RNA silencing". In: *RNA* 16.4, pp. 673–95. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.2000810. URL: https://www.ncbi.nlm.nih.gov/pubmed/20181738.

Hawkins, P. G. et al. (2009). "Promoter targeted small RNAs induce long-term transcriptional gene silencing in human cells". In: *Nucleic Acids Res* 37.9, pp. 2984–95. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkp127. URL: https://www.ncbi.nlm.nih.gov/pubmed/19304753.

Hellmuth, K. et al. (1998). "Yeast Los1p has properties of an exportin-like nucleocytoplasmic transport factor for tRNA". In: *Mol Cell Biol* 18.11, pp. 6374–86. ISSN: 0270-7306 (Print) 0270-7306 (Linking). DOI: 10.1128/MCB.18.11.6374. URL: https://www.ncbi.nlm.nih.gov/pubmed/9774653.

Helwak, A. et al. (2013). "Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding". In: *Cell* 153.3, pp. 654–65. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2013.03.043. URL: https://www.ncbi.nlm.nih.gov/pubmed/23622248.

Herrmann, C. et al. (1996). "The major vault protein (MVP100) is contained in cholinergic nerve terminals of electric ray electric organ". In: *J Biol Chem* 271.23, pp. 13908–15. ISSN: 0021-9258 (Print) 0021-9258 (Linking). DOI: 10.1074/jbc.271.23.13908. URL: https://www.ncbi.nlm.nih.gov/pubmed/8662815.

Hoffer, P. et al. (2011). "Posttranscriptional gene silencing in nuclei". In: *Proc Natl Acad Sci U S A* 108.1, pp. 409–14. ISSN: 1091-6490 (Electronic) 0027-8424 (Linking). DOI: 10.1073/pnas.1009805108. URL: https://www.ncbi.nlm.nih.gov/pubmed/21173264.

Horos, R. et al. (2019). "The Small Non-coding Vault RNA1-1 Acts as a Riboregulator of Autophagy". In: *Cell* 176.5, 1054–1067 e12. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2019.01.030. URL: https://www.ncbi.nlm.nih.gov/pubmed/30773316.

Hu, H. Y. et al. (2009). "Sequence features associated with microRNA strand selection in humans and flies". In: *BMC Genomics* 10, p. 413. ISSN: 1471-2164 (Electronic) 1471-2164 (Linking). DOI: 10.1186/1471-2164-10-413. URL: https://www.ncbi.nlm.nih.gov/pubmed/19732433.

Huang, B. et al. (2017). "tRF/miR-1280 Suppresses Stem Cell-like Cells and Metastasis in Colorectal Cancer". In: *Cancer Res* 77.12, pp. 3194–3206. ISSN: 1538-7445 (Electronic) 0008-5472 (Linking). DOI: 10.1158/0008-5472.CAN-16-3146. URL: https://www.ncbi.nlm.nih.gov/pubmed/28446464.

Huh, W. K. et al. (2003). "Global analysis of protein localization in budding yeast". In: *Nature* 425.6959, pp. 686–91. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature02026. URL: https://www.ncbi.nlm.nih.gov/pubmed/14562095.

Hussain, S. et al. (2013). "NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs". In: *Cell Rep* 4.2, pp. 255–61. ISSN: 2211-1247 (Electronic). DOI: 10.1016/j.celrep.2013.06.029. URL: https://www.ncbi.nlm.nih.gov/pubmed/23871666.

Ivanov, P. et al. (2011). "Angiogenin-induced tRNA fragments inhibit translation initiation". In: *Mol Cell* 43.4, pp. 613–23. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2011.06.022. URL: https://www.ncbi.nlm.nih.gov/pubmed/21855800.

Jackman, J. E. et al. (2003). "Identification of the yeast gene encoding the tRNA m1G methyltransferase responsible for modification at position 9". In: *RNA* 9.5, pp. 574–85. ISSN: 1355-8382 (Print) 1355-8382 (Linking). DOI: 10.1261/rna.5070303. URL: https://www.ncbi.nlm.nih.gov/pubmed/12702816.

Janowski, B. A. et al. (2005). "Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs". In: *Nat Chem Biol* 1.4, pp. 216–22. ISSN:

1552-4450 (Print) 1552-4450 (Linking). DOI: 10.1038/nchembio725. URL: https://www.ncbi.nlm.nih.gov/pubmed/16408038.

Janowski, B. A. et al. (2006). "Involvement of AGO1 and AGO2 in mammalian transcriptional silencing". In: *Nat Struct Mol Biol* 13.9, pp. 787–92. ISSN: 1545-9993 (Print) 1545-9985 (Linking). DOI: 10.1038/nsmb1140. URL: https://www.ncbi.nlm.nih.gov/pubmed/16936728.

Jensen, L. R. et al. (2019). "A mouse model for intellectual disability caused by mutations in the X-linked 2'Omethyltransferase Ftsj1 gene". In: *Biochim Biophys Acta Mol Basis Dis* 1865.9, pp. 2083–2093. ISSN: 1879-260X (Electronic) 0925-4439 (Linking). DOI: 10.1016/j.bbadis.2018.12.011. URL: https://www.ncbi.nlm.nih.gov/pubmed/30557699.

Jeppesen, D. K. et al. (2019). "Reassessment of Exosome Composition". In: *Cell* 177.2, 428–445 e18. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2019.02.029. URL: https://www.ncbi.nlm.nih.gov/pubmed/30951670.

Jmol development team (Oct. 7, 2016). *Jmol*. Version 14.6.4. URL: http://jmol.sourceforge.net/.

Jo, M. H. et al. (2015). "Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs". In: *Mol Cell* 59.1, pp. 117–24. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2015.04.027. URL: https://www.ncbi.nlm.nih.gov/pubmed/26140367.

Kalantari, R., C. M. Chiang, and D. R. Corey (2016a). "Regulation of mammalian transcription and splicing by Nuclear RNAi". In: *Nucleic Acids Res* 44.2, pp. 524–37. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkv1305. URL: https://www.ncbi.nlm.nih.gov/pubmed/26612865.

Kalantari, R. et al. (2016b). "Stable association of RNAi machinery is conserved between the cytoplasm and nucleus of human cells". In: *RNA* 22.7, pp. 1085–98. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.056499.116. URL: https://www.ncbi.nlm.nih.gov/pubmed/27198507.

Kar, U. K. et al. (2011). "Novel CCL21-vault nanocapsule intratumoral delivery inhibits lung cancer growth". In: *PLoS One* 6.5, e18758. ISSN: 1932-6203 (Electronic) 1932-6203

(Linking). DOI: 10.1371/journal.pone.0018758. URL: https://www.ncbi.nlm.nih.gov/pubmed/21559281.

Kawahara, Y. et al. (2007). "RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex". In: *EMBO Rep* 8.8, pp. 763–9. ISSN: 1469-221X (Print) 1469-221X (Linking). DOI: 10.1038/sj.embor.7401011. URL: https://www.ncbi.nlm.nih.gov/pubmed/17599088.

Kawamata, T., H. Seitz, and Y. Tomari (2009). "Structural determinants of miRNAs for RISC loading and slicer-independent unwinding". In: *Nat Struct Mol Biol* 16.9, pp. 953–60. ISSN: 1545-9985 (Electronic) 1545-9985 (Linking). DOI: 10.1038/nsmb.1630. URL: https://www.ncbi.nlm.nih.gov/pubmed/19684602.

Keam, S. P. and G. Hutvagner (2015). "tRNA-Derived Fragments (tRFs): Emerging New Roles for an Ancient RNA in the Regulation of Gene Expression". In: *Life (Basel)* 5.4, pp. 1638–51. ISSN: 2075-1729 (Print) 2075-1729 (Linking). DOI: 10.3390/life5041638. URL: https://www.ncbi.nlm.nih.gov/pubmed/26703738.

Kedersha, N. L. and L. H. Rome (1986). "Isolation and characterization of a novel ribonucleoprotein particle: large structures contain a single species of small RNA". In: *J Cell Biol* 103.3, pp. 699–709. ISSN: 0021-9525 (Print) 0021-9525 (Linking). DOI: 10.1083/jcb.103.3.699. URL: https://www.ncbi.nlm.nih.gov/pubmed/2943744.

Kedersha, N. L. et al. (1990). In: *J Cell Biol* 110.4, pp. 895–901. ISSN: 0021-9525 (Print) 0021-9525 (Linking). DOI: 10.1083/jcb.110.4.895. URL: https://www.ncbi.nlm.nih.gov/pubmed/1691193.

Kedersha, N. L. et al. (1991). "Vaults. III. Vault ribonucleoprotein particles open into flower-like structures with octagonal symmetry". In: *J Cell Biol* 112.2, pp. 225–35. ISSN: 0021-9525 (Print) 0021-9525 (Linking). DOI: 10.1083/jcb.112.2.225. URL: https://www.ncbi.nlm.nih.gov/pubmed/1988458.

Kern, F. et al. (2020). "miRSwitch: detecting microRNA arm shift and switch events". In: *Nucleic Acids Res* 48.W1, W268–W274. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkaa323. URL: https://www.ncbi.nlm.nih.gov/pubmed/32356893.

Kertesz, M. et al. (2007). "The role of site accessibility in microRNA target recognition". In: *Nat Genet* 39.10, pp. 1278–84. ISSN: 1546-1718 (Electronic) 1061-4036 (Linking). DOI: 10.1038/ng2135. URL: https://www.ncbi.nlm.nih.gov/pubmed/17893677.

Kickhoefer, V. A. et al. (1993). "Vault ribonucleoprotein particles from rat and bullfrog contain a related small RNA that is transcribed by RNA polymerase III". In: *J Biol Chem* 268.11, pp. 7868–73. ISSN: 0021-9258 (Print) 0021-9258 (Linking). URL: https://www.ncbi.nlm.nih.gov/pubmed/7681830.

Kickhoefer, V. A. et al. (1998). "Vaults are up-regulated in multidrug-resistant cancer cell lines". In: *J Biol Chem* 273.15, pp. 8971–4. ISSN: 0021-9258 (Print) 0021-9258 (Linking). DOI: 10.1074/jbc.273.15.8971. URL: https://www.ncbi.nlm.nih.gov/pubmed/9535882.

Kickhoefer, V. A. et al. (1999a). "The 193-kD vault protein, VPARP, is a novel poly(ADP-ribose) polymerase". In: *J Cell Biol* 146.5, pp. 917–28. ISSN: 0021-9525 (Print) 0021-9525 (Linking). DOI: 10.1083/jcb.146.5.917. URL: https://www.ncbi.nlm.nih.gov/pubmed/10477748.

Kickhoefer, V. A. et al. (1999b). "Vaults and telomerase share a common subunit, TEP1". In: *J Biol Chem* 274.46, pp. 32712–7. ISSN: 0021-9258 (Print) 0021-9258 (Linking). DOI: 10.1074/jbc.274.46.32712. URL: https://www.ncbi.nlm.nih.gov/pubmed/10551828.

Kickhoefer, V. A. et al. (2001). "The Telomerase/vault-associated protein TEP1 is required for vault RNA stability and its association with the vault particle". In: *J Cell Biol* 152.1, pp. 157–64. ISSN: 0021-9525 (Print) 0021-9525 (Linking). DOI: 10.1083/jcb.152.1.157. URL: https://www.ncbi.nlm.nih.gov/pubmed/11149928.

Kickhoefer, V. A. et al. (2002). "The La RNA-binding protein interacts with the vault RNA and is a vault-associated protein". In: *J Biol Chem* 277.43, pp. 41282–6. ISSN: 0021-9258 (Print) 0021-9258 (Linking). DOI: 10.1074/jbc.M206980200. URL: https://www.ncbi.nlm.nih.gov/pubmed/12196535.

Kickhoefer, V. A. et al. (2009). "Targeting vault nanoparticles to specific cell surface receptors". In: *ACS Nano* 3.1, pp. 27–36. ISSN: 1936-086X (Electronic) 1936-0851 (Linking). DOI: 10.1021/nn800638x. URL: https://www.ncbi.nlm.nih.gov/pubmed/19206245.

Kim, B. and V. N. Kim (2019). "fCLIP-seq for transcriptomic footprinting of dsRNA-binding proteins: Lessons from DROSHA". In: *Methods* 152, pp. 3–11. ISSN: 1095-9130 (Electronic) 1046-2023 (Linking). DOI: 10.1016/j.ymeth.2018.06.004. URL: https://www.ncbi.nlm.nih.gov/pubmed/29902563.

Kim, D. H. et al. (2006). "Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells". In: *Nat Struct Mol Biol* 13.9, pp. 793–7. ISSN: 1545-9993 (Print) 1545-9985 (Linking). DOI: 10.1038/nsmb1142. URL: https://www.ncbi.nlm.nih.gov/pubmed/16936726.

Kim, H. K. et al. ((2017)). "A transfer-RNA-derived small RNA regulates ribosome biogenesis". In: *Nature* 552.7683, pp. 57–62. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature25005. URL: https://www.ncbi.nlm.nih.gov/pubmed/29186115.

Kim, Y. K., B. Kim, and V. N. Kim (2016). "Re-evaluation of the roles of DROSHA, Export in 5, and DICER in microRNA biogenesis". In: *Proc Natl Acad Sci U S A* 113.13, E1881–9. ISSN: 1091-6490 (Electronic) 0027-8424 (Linking). DOI: 10.1073/pnas.1602532113. URL: https://www.ncbi.nlm.nih.gov/pubmed/26976605.

Kong, L. B. et al. (1999). "Structure of the vault, a ubiquitous celular component". In: *Structure* 7.4, pp. 371–9. ISSN: 0969-2126 (Print) 0969-2126 (Linking). DOI: 10.1016/s0969-2126(99)80050-1. URL: https://www.ncbi.nlm.nih.gov/pubmed/10196123.

Kowalak, J. A. et al. (1994). "The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles". In: *Biochemistry* 33.25, pp. 7869–76. ISSN: 0006-2960 (Print) 0006-2960 (Linking). DOI: 10.1021/bi00191a014. URL: https://www.ncbi.nlm.nih.gov/pubmed/7516708.

Kozomara, A., M. Birgaoanu, and S. Griffiths-Jones (2019). "miRBase: from microRNA sequences to function". In: *Nucleic Acids Research* 47.D1, pp. D155–D162. ISSN: 0305-1048. DOI: 10.1093/nar/gky1141. URL: %3CGo%20to%20ISI%3E://WOS:000462587400026.

Kumar, P. et al. (2014). "Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets". In: *BMC Biol* 12, p. 78. ISSN: 1741-7007 (Electronic) 1741-7007 (Linking). DOI:

10.1186/s12915-014-0078-0. URL: https://www.ncbi.nlm.nih.gov/pubmed/25270025.

Kumar, P. et al. (2015). "tRFdb: a database for transfer RNA fragments". In: *Nucleic Acids Res* 43.Database issue, pp. D141–5. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gku1138. URL: https://www.ncbi.nlm.nih.gov/pubmed/25392422.

Kuscu, C. et al. (2018). "tRNA fragments (tRFs) guide Ago to regulate gene expression post-transcriptionally in a Dicer-independent manner". In: *RNA* 24.8, pp. 1093–1105. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.066126.118. URL: https://www.ncbi.nlm.nih.gov/pubmed/29844106.

Kwak, P. B. and Y. Tomari (2012). "The N domain of Argonaute drives duplex unwinding during RISC assembly". In: *Nat Struct Mol Biol* 19.2, pp. 145–51. ISSN: 1545-9985 (Electronic) 1545-9985 (Linking). DOI: 10.1038/nsmb.2232. URL: https://www.ncbi.nlm.nih.gov/pubmed/22233755.

Lal, A. et al. (2009). "miR-24 Inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to "seedless" 3'UTR microRNA recognition elements". In: *Mol Cell* 35.5, pp. 610–25. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2009.08.020. URL: https://www.ncbi.nlm.nih.gov/pubmed/19748357.

Lal, A. et al. (2011). "Capture of microRNA-bound mRNAs identifies the tumor suppressor miR-34a as a regulator of growth factor signaling". In: *PLoS Genet* 7.11, e1002363. ISSN: 1553-7404 (Electronic) 1553-7390 (Linking). DOI: 10.1371/journal.pgen.1002363. URL: https://www.ncbi.nlm.nih.gov/pubmed/22102825.

Landgraf, P. et al. (2007). "A mammalian microRNA expression atlas based on small RNA library sequencing". In: *Cell* 129.7, pp. 1401–14. ISSN: 0092-8674 (Print) 0092-8674 (Linking). DOI: 10.1016/j.cell.2007.04.040. URL: https://www.ncbi.nlm.nih.gov/pubmed/17604727.

Langenberger, D. et al. (2013). "Dicer-processed small RNAs: rules and exceptions". In: *J Exp Zool B Mol Dev Evol* 320.1, pp. 35–46. ISSN: 1552-5015 (Electronic) 1552-5007

(Linking). DOI: 10.1002/jez.b.22481. URL: https://www.ncbi.nlm.nih.gov/pubmed/23165937.

Langlois, M. A. et al. (2005). "Cytoplasmic and nuclear retained DMPK mRNAs are targets for RNA interference in myotonic dystrophy cells". In: *J Biol Chem* 280.17, pp. 16949–54. ISSN: 0021-9258 (Print) 0021-9258 (Linking). DOI: 10.1074/jbc.M501591200. URL: https://www.ncbi.nlm.nih.gov/pubmed/15722335.

Langmead, B. et al. (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome". In: *Genome Biol* 10.3, R25. ISSN: 1474-760X (Electronic) 1474-7596 (Linking). DOI: 10.1186/gb-2009-10-3-r25. URL: https://www.ncbi.nlm.nih.gov/pubmed/19261174.

Lau, N. C. et al. (2006). "Characterization of the piRNA complex from rat testes". In: *Science* 313.5785, pp. 363–7. ISSN: 1095-9203 (Electronic) 0036-8075 (Linking). DOI: 10.1126/science.1130164. URL: https://www.ncbi.nlm.nih.gov/pubmed/16778019.

Lau, P. W. et al. (2012). "The molecular architecture of human Dicer". In: *Nat Struct Mol Biol* 19.4, pp. 436–40. ISSN: 1545-9985 (Electronic) 1545-9985 (Linking). DOI: 10.1038/nsmb.2268. URL: https://www.ncbi.nlm.nih.gov/pubmed/22426548.

Lee, E. K. et al. (2016). "nc886, a non-coding RNA and suppressor of PKR, exerts an oncogenic function in thyroid cancer". In: *Oncotarget* 7.46, pp. 75000–75012. ISSN: 1949-2553 (Electronic) 1949-2553 (Linking). DOI: 10.18632/oncotarget.11852. URL: https://www.ncbi.nlm.nih.gov/pubmed/27612419.

Lee, H. S. et al. (2014). "Epigenetic silencing of the non-coding RNA nc886 provokes oncogenes during human esophageal tumorigenesis". In: *Oncotarget* 5.11, pp. 3472–81. ISSN: 1949-2553 (Electronic) 1949-2553 (Linking). DOI: 10.18632/oncotarget.1927. URL: https://www.ncbi.nlm.nih.gov/pubmed/25004084.

Lee, K. et al. (2011). "Precursor miR-886, a novel noncoding RNA repressed in cancer, associates with PKR and modulates its activity". In: *RNA* 17.6, pp. 1076–89. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.2701111. URL: https://www.ncbi.nlm.nih.gov/pubmed/21518807.

Lee, K. S. et al. (2014). "nc886, a non-coding RNA of anti-proliferative role, is suppressed by CpG DNA methylation in human gastric cancer". In: *Oncotarget* 5.11, pp. 3944–55. ISSN: 1949-2553 (Electronic) 1949-2553 (Linking). DOI: 10.18632/oncotarget.2047. URL: https://www.ncbi.nlm.nih.gov/pubmed/25003254.

Lee, Y. S. (2015). "A Novel Type of Non-coding RNA, nc886, Implicated in Tumor Sensing and Suppression". In: *Genomics Inform* 13.2, pp. 26–30. ISSN: 1598-866X (Print) 1598-866X (Linking). DOI: 10.5808/GI.2015.13.2.26. URL: https://www.ncbi.nlm.nih.gov/pubmed/26175659.

Lei, J. et al. (2017). "Non-coding RNA 886 promotes renal cell carcinoma growth and metastasis through the Janus kinase 2/signal transducer and activator of transcription 3 signaling pathway". In: *Mol Med Rep* 16.4, pp. 4273–4278. ISSN: 1791-3004 (Electronic) 1791-2997 (Linking). DOI: 10.3892/mmr.2017.7093. URL: https://www.ncbi.nlm.nih.gov/pubmed/28765891.

Leucci, E. et al. (2013). "microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus". In: *Sci Rep* 3, p. 2535. ISSN: 2045-2322 (Electronic) 2045-2322 (Linking). DOI: 10.1038/srep02535. URL: https://www.ncbi.nlm.nih.gov/pubmed/23985560.

Leuschner, P. J. et al. (2006). "Cleavage of the siRNA passenger strand during RISC assembly in human cells". In: *EMBO Rep* 7.3, pp. 314–20. ISSN: 1469-221X (Print) 1469-221X (Linking). DOI: 10.1038/sj.embor.7400637. URL: https://www.ncbi.nlm.nih.gov/pubmed/16439995.

Li, F. et al. (2015). "Robust expression of vault RNAs induced by influenza A virus plays a critical role in suppression of PKR-mediated innate immunity". In: *Nucleic Acids Res* 43.21, pp. 10321–37. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkv1078. URL: https://www.ncbi.nlm.nih.gov/pubmed/26490959.

Li, H. et al. (2009). "The Sequence Alignment/Map format and SAMtools". In: *Bioinformatics* 25.16, pp. 2078–9. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking). DOI: 10.1093/bioinformatics/btp352. URL: https://www.ncbi.nlm.nih.gov/pubmed/19505943.

Li, J. H. et al. (2011). "MicroRNA miR-886-5p inhibits apoptosis by down-regulating Bax expression in human cervical carcinoma cells". In: *Gynecol Oncol* 120.1, pp. 145–51. ISSN: 1095-6859 (Electronic) 0090-8258 (Linking). DOI: 10.1016/j.ygyno.2010.09.009. URL: https://www.ncbi.nlm.nih.gov/pubmed/20947150.

Liao, Y., G. K. Smyth, and W. Shi (2014). "featureCounts: an efficient general purpose program for assigning sequence reads to genomic features". In: *Bioinformatics* 30.7, pp. 923–30. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking). DOI: 10.1093/bioinformatics/btt656. URL: https://www.ncbi.nlm.nih.gov/pubmed/24227677.

Liu, F. et al. (2016). "ALKBH1-Mediated tRNA Demethylation Regulates Translation". In: *Cell* 167.3, 816–828 e16. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2016.09.038. URL: https://www.ncbi.nlm.nih.gov/pubmed/27745969.

Liu, J. et al. (2004). "Argonaute2 is the catalytic engine of mammalian RNAi". In: *Science* 305.5689, pp. 1437–41. ISSN: 1095-9203 (Electronic) 0036-8075 (Linking). DOI: 10.1126/science.1102513. URL: https://www.ncbi.nlm.nih.gov/pubmed/15284456.

Liu, Y. et al. (2000). "Telomerase-associated protein TEP1 is not essential for telomerase activity or telomere length maintenance in vivo". In: *Mol Cell Biol* 20.21, pp. 8178–84. ISSN: 0270-7306 (Print) 0270-7306 (Linking). DOI: 10.1128/MCB.20.21.8178-8184.2000. URL: https://www.ncbi.nlm.nih.gov/pubmed/11027287.

Lloret, M. et al. (2008). "MVP expression is related to IGF1-R in cervical carcinoma patients treated by radiochemotherapy". In: *Gynecol Oncol* 110.3, pp. 304–7. ISSN: 1095-6859 (Electronic) 0090-8258 (Linking). DOI: 10.1016/j.ygyno.2008.04.034. URL: https://www.ncbi.nlm.nih.gov/pubmed/18599112.

Lorenz, R. et al. (2011). "ViennaRNA Package 2.0". In: *Algorithms Mol Biol* 6, p. 26. ISSN: 1748-7188 (Electronic) 1748-7188 (Linking). DOI: 10.1186/1748-7188-6-26. URL: https://www.ncbi.nlm.nih.gov/pubmed/22115189.

Love, M. I., W. Huber, and S. Anders (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". In: *Genome Biol* 15.12, p. 550. ISSN: 1474-760X (Electronic) 1474-7596 (Linking). DOI: 10.1186/s13059-014-0550-8. URL: https://www.ncbi.nlm.nih.gov/pubmed/25516281.

Luo, Q. J. et al. (2021). "RNA structure probing reveals the structural basis of Dicer binding and cleavage". In: *Nat Commun* 12.1, p. 3397. ISSN: 2041-1723 (Electronic) 2041-1723 (Linking). DOI: 10.1038/s41467-021-23607-w. URL: https://www.ncbi.nlm.nih.gov/pubmed/34099665.

Ma, J. B., K. Ye, and D. J. Patel (2004). "Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain". In: *Nature* 429.6989, pp. 318–322. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature02519. URL: https://www.ncbi.nlm.nih.gov/pubmed/15152257.

Macias, S. et al. (2012). "DGCR8 HITS-CLIP reveals novel functions for the Microprocessor". In: *Nat Struct Mol Biol* 19.8, pp. 760–6. ISSN: 1545-9985 (Electronic) 1545-9985 (Linking). DOI: 10.1038/nsmb.2344. URL: https://www.ncbi.nlm.nih.gov/pubmed/22796965.

Marcel, Martin (2011). "Cutadapt removes adapter sequences from high-throughput sequencing reads". In: *EMBnet.journal* 17, pp. 10–12. DOI: http://dx.doi.org/10.14806/ej.17.1.200. URL: http://journal.embnet.org/index.php/embnetjournal/article/view/200.

Mashima, T. et al. (2008). "Interactions between antitumor drugs and vault RNA". In: *Nucleic Acids Symp Ser (Oxf)* 52, pp. 217–8. ISSN: 1746-8272 (Electronic) 0261-3166 (Linking). DOI: 10.1093/nass/nrn110. URL: https://www.ncbi.nlm.nih.gov/pubmed/18776331.

Matzke, M. A. et al. (1989). "Reversible methylation and inactivation of marker genes in sequentially transformed tobacco plants". In: *EMBO J* 8.3, pp. 643–9. ISSN: 0261-4189 (Print) 0261-4189 (Linking). DOI: 10.1002/j.1460-2075.1989.tb03421.x. URL: https://www.ncbi.nlm.nih.gov/pubmed/16453872.

Maute, R. L. et al. (2013). "tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma". In: *Proc Natl Acad Sci U S A* 110.4, pp. 1404–9. ISSN: 1091-6490 (Electronic) 0027-8424 (Linking). DOI: 10.1073/pnas.1206761110. URL: https://www.ncbi.nlm.nih.gov/pubmed/23297232.

McCloskey, J. A. et al. (2001). "Post-transcriptional modification in archaeal tRNAs: identities and phylogenetic relations of nucleotides from mesophilic and hyperthermophilic Methanococcales". In: *Nucleic Acids Res* 29.22, pp. 4699–706. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/29.22.4699. URL: https://www.ncbi.nlm.nih.gov/pubmed/11713320.

McLaurin, D. M. et al. (2020). "Molecular determinants that govern scaRNA processing by Drosha/DGCR8". In: *Biol Open* 9.10. ISSN: 2046-6390 (Electronic) 2046-6390 (Linking). DOI: 10.1242/bio.054619. URL: https://www.ncbi.nlm.nih.gov/pubmed/33037012.

Meister, G. et al. (2004). "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs". In: *Mol Cell* 15.2, pp. 185–97. ISSN: 1097-2765 (Print) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2004.07.007. URL: https://www.ncbi.nlm.nih.gov/pubmed/15260970.

Meng, Y. et al. (2013). "Introns targeted by plant microRNAs: a possible novel mechanism of gene regulation". In: *Rice (N Y)* 6.1, p. 8. ISSN: 1939-8425 (Print) 1939-8425 (Linking). DOI: 10.1186/1939-8433-6-8. URL: https://www.ncbi.nlm.nih.gov/pubmed/24280590.

Mi, S. et al. (2008). "Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide". In: *Cell* 133.1, pp. 116–27. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2008.02.034. URL: https://www.ncbi.nlm.nih.gov/pubmed/18342361.

Mikyas, Y. et al. (2004). "Cryoelectron microscopy imaging of recombinant and tissue derived vaults: localization of the MVP N termini and VPARP". In: *J Mol Biol* 344.1, pp. 91–105. ISSN: 0022-2836 (Print) 0022-2836 (Linking). DOI: 10.1016/j.jmb.2004.09.021. URL: https://www.ncbi.nlm.nih.gov/pubmed/15504404.

Minones-Moyano, E. et al. (2013). "Upregulation of a small vault RNA (svtRNA2-1a) is an early event in Parkinson disease and induces neuronal dysfunction". In: *RNA Biol* 10.7, pp. 1093–106. ISSN: 1555-8584 (Electronic) 1547-6286 (Linking). DOI: 10.4161/rna.24813. URL: https://www.ncbi.nlm.nih.gov/pubmed/23673382.

Moore, C. B. et al. (2010). "Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown". In: *Methods Mol Biol* 629, pp. 141–58. ISSN: 1940-6029 (Electronic) 1064-3745 (Linking). DOI: 10.1007/978-1-60761-657-3_10. URL: https://www.ncbi.nlm.nih.gov/pubmed/20387148.

Mootha, V. K. et al. (2003). "PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes". In: *Nat Genet* 34.3, pp. 267–73. ISSN: 1061-4036 (Print) 1061-4036 (Linking). DOI: 10.1038/ng1180. URL: https://www.ncbi.nlm.nih.gov/pubmed/12808457.

Morris, K. V. (2009). "Long antisense non-coding RNAs function to direct epigenetic complexes that regulate transcription in human cells". In: *Epigenetics* 4.5, pp. 296–301. ISSN: 1559-2308 (Electronic) 1559-2294 (Linking). DOI: 10.4161/epi.4.5.9282. URL: https://www.ncbi.nlm.nih.gov/pubmed/19633414.

Motorin, Y. and H. Grosjean (1999). "Multisite-specific tRNA:m5C-methyltransferase (Trm4) in yeast Saccharomyces cerevisiae: identification of the gene and substrate specificity of the enzyme". In: *RNA* 5.8, pp. 1105–18. ISSN: 1355-8382 (Print) 1355-8382 (Linking). DOI: 10.1017/s1355838299982201. URL: https://www.ncbi.nlm.nih.gov/pubmed/10445884.

Mrazek, J. et al. (2007). "Subtractive hybridization identifies novel differentially expressed ncRNA species in EBV-infected human B cells". In: *Nucleic Acids Res* 35.10, e73. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkm244. URL: https://www.ncbi.nlm.nih.gov/pubmed/17478510.

Muller, M., F. Fazi, and C. Ciaudo (2019). "Argonaute Proteins: From Structure to Function in Development and Pathological Cell Fate Determination". In: *Front Cell Dev Biol* 7, p. 360. ISSN: 2296-634X (Print) 2296-634X (Linking). DOI: 10.3389/fcell.2019.00360. URL: https://www.ncbi.nlm.nih.gov/pubmed/32039195.

Muramatsu, T. et al. (1988). "Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification". In: *Nature* 336.6195, pp. 179–81. ISSN: 0028-0836 (Print) 0028-0836 (Linking). DOI: 10.1038/336179a0. URL: https://www.ncbi.nlm.nih.gov/pubmed/3054566.

Nakanishi, K. (2016). "Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins?" In: *Wiley Interdiscip Rev RNA* 7.5, pp. 637–60. ISSN: 1757-7012 (Electronic) 1757-7004 (Linking). DOI: 10.1002/wrna.1356. URL: https://www.ncbi.nlm.nih.gov/pubmed/27184117.

— (2022). "Anatomy of four human Argonaute proteins". In: *Nucleic Acids Res*. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkac519. URL: https://www.ncbi.nlm.nih.gov/pubmed/35736234.

Nandy, C. et al. (2009). "Epstein-barr virus-induced expression of a novel human vault RNA". In: *J Mol Biol* 388.4, pp. 776–84. ISSN: 1089-8638 (Electronic) 0022-2836 (Linking). DOI: 10.1016/j.jmb.2009.03.031. URL: https://www.ncbi.nlm.nih.gov/pubmed/19298825.

Nguyen, T. D. et al. (2022). "Secondary structure RNA elements control the cleavage activity of DICER". In: *Nat Commun* 13.1, p. 2138. ISSN: 2041-1723 (Electronic) 2041-1723 (Linking). DOI: 10.1038/s41467-022-29822-3. URL: https://www.ncbi.nlm.nih.gov/pubmed/35440644.

Nobles, K. N. et al. (2002). "Highly conserved modified nucleosides influence Mg2+-dependent tRNA folding". In: *Nucleic Acids Res* 30.21, pp. 4751–60. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkf595. URL: https://www.ncbi.nlm.nih.gov/pubmed/12409466.

Noh, K. H. et al. (2022). "Novel cancer stem cell marker MVP enhances temozolomide-resistance in glioblastoma". In: *Transl Oncol* 15.1, p. 101255. ISSN: 1936-5233 (Print) 1936-5233 (Linking). DOI: 10.1016/j.tranon.2021.101255. URL: https://www.ncbi.nlm.nih.gov/pubmed/34742152.

Nojima, T. et al. (2016). "Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide". In: *Nat Protoc* 11.3, pp. 413–28. ISSN: 1750-2799 (Electronic) 1750-2799 (Linking). DOI: 10.1038/nprot.2016.012. URL: https://www.ncbi.nlm.nih.gov/pubmed/26844429.

Okazaki, S. et al. (2016). "Development of an ErbB4 monoclonal antibody that blocks neuregulin-1-induced ErbB4 activation in cancer cells". In: *Biochem Biophys Res Commun* 470.1, pp. 239–244. ISSN: 1090-2104 (Electronic) 0006-291X (Linking). DOI: 10.

1016/j.bbrc.2016.01.045. URL: https://www.ncbi.nlm.nih.gov/pubmed/26780728.

Paddison, P. J. et al. (2002). "Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells". In: *Genes Dev* 16.8, pp. 948–58. ISSN: 0890-9369 (Print) 0890-9369 (Linking). DOI: 10.1101/gad.981002. URL: https://www.ncbi.nlm.nih.gov/pubmed/11959843.

Pan, B., J. Yu, and X. Liu (2022). "Upregulation of miR-886 indicates poor prognosis and promotes tumour progression of prostate cancer". In: *Andrologia* 54.1, e14296. ISSN: 1439-0272 (Electronic) 0303-4569 (Linking). DOI: 10.1111/and.14296. URL: https://www.ncbi.nlm.nih.gov/pubmed/34787343.

Park, J. H. and C. Shin (2015). "Slicer-independent mechanism drives small-RNA strand separation during human RISC assembly". In: *Nucleic Acids Res* 43.19, pp. 9418–33. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkv937. URL: https://www.ncbi.nlm.nih.gov/pubmed/26384428.

Park, M. S. et al. (2017). "Human Argonaute3 has slicer activity". In: *Nucleic Acids Res* 45.20, pp. 11867–11877. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkx916. URL: https://www.ncbi.nlm.nih.gov/pubmed/29040713.

Park, M. S. et al. (2020). "Human Argonaute2 and Argonaute3 are catalytically activated by different lengths of guide RNA". In: *Proc Natl Acad Sci U S A* 117.46, pp. 28576–28578. ISSN: 1091-6490 (Electronic) 0027-8424 (Linking). DOI: 10.1073/pnas.2015026117. URL: https://www.ncbi.nlm.nih.gov/pubmed/33122430.

Parker, J. S., S. M. Roe, and D. Barford (2004). "Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity". In: *EMBO J* 23.24, pp. 4727–37. ISSN: 0261-4189 (Print) 0261-4189 (Linking). DOI: 10.1038/sj.emboj.7600488. URL: https://www.ncbi.nlm.nih.gov/pubmed/15565169.

— (2005). "Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex". In: *Nature* 434.7033, pp. 663–6. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature03462. URL: https://www.ncbi.nlm.nih.gov/pubmed/15800628.

Persson, H. et al. (2009). "The non-coding RNA of the multidrug resistance-linked vault particle encodes multiple regulatory small RNAs". In: *Nat Cell Biol* 11.10, pp. 1268–71. ISSN: 1476-4679 (Electronic) 1465-7392 (Linking). DOI: 10.1038/ncb1972. URL: https://www.ncbi.nlm.nih.gov/pubmed/19749744.

Phizicky, E. M. and J. D. Alfonzo (2010). "Do all modifications benefit all tRNAs?" In: *FEBS Lett* 584.2, pp. 265–71. ISSN: 1873-3468 (Electronic) 0014-5793 (Linking). DOI: 10.1016/j.febslet.2009.11.049. URL: https://www.ncbi.nlm.nih.gov/pubmed/19931536.

Piskounova, E. et al. (2011). "Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms". In: *Cell* 147.5, pp. 1066–79. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2011.10.039. URL: https://www.ncbi.nlm.nih.gov/pubmed/22118463.

Plotnikova, O., A. Baranova, and M. Skoblov (2019). "Comprehensive Analysis of Human microRNA-mRNA Interactome". In: *Front Genet* 10, p. 933. ISSN: 1664-8021 (Print) 1664-8021 (Linking). DOI: 10.3389/fgene.2019.00933. URL: https://www.ncbi.nlm.nih.gov/pubmed/31649721.

Poderycki, M. J. et al. (2005). "The p80 homology region of TEP1 is sufficient for its association with the telomerase and vault RNAs, and the vault particle". In: *Nucleic Acids Res* 33.3, pp. 893–902. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gki234. URL: https://www.ncbi.nlm.nih.gov/pubmed/15701761.

Poderycki, M. J. et al. (2006). "The vault exterior shell is a dynamic structure that allows incorporation of vault-associated proteins into its interior". In: *Biochemistry* 45.39, pp. 12184–93. ISSN: 0006-2960 (Print) 0006-2960 (Linking). DOI: 10.1021/bi0610552. URL: https://www.ncbi.nlm.nih.gov/pubmed/17002318.

Popenda, M. et al. (2012). "Automated 3D structure composition for large RNAs". In: *Nucleic Acids Res* 40.14, e112. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gks339. URL: https://www.ncbi.nlm.nih.gov/pubmed/22539264.

Quinlan, A. R. and I. M. Hall (2010). "BEDTools: a flexible suite of utilities for comparing genomic features". In: *Bioinformatics* 26.6, pp. 841–2. ISSN: 1367-4811 (Electronic)

1367-4803 (Linking). DOI: 10.1093/bioinformatics/btq033. URL: https://www.ncbi.nlm.nih.gov/pubmed/20110278.

Ramberg, H. et al. (2021). "Proteomic analyses identify major vault protein as a prognostic biomarker for fatal prostate cancer". In: *Carcinogenesis* 42.5, pp. 685–693. ISSN: 1460-2180 (Electronic) 0143-3334 (Linking). DOI: 10.1093/carcin/bgab015. URL: https://www.ncbi.nlm.nih.gov/pubmed/33609362.

Ramirez, F. et al. (2014). "deepTools: a flexible platform for exploring deep-sequencing data". In: *Nucleic Acids Res* 42.Web Server issue, W187–91. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gku365. URL: https://www.ncbi.nlm.nih.gov/pubmed/24799436.

Reifur, L. et al. (2012). "Distinct subcellular localization of tRNA-derived fragments in the infective metacyclic forms of Trypanosoma cruzi". In: *Mem Inst Oswaldo Cruz* 107.6, pp. 816–9. ISSN: 1678-8060 (Electronic) 0074-0276 (Linking). DOI: 10.1590/s0074-02762012000600018. URL: https://www.ncbi.nlm.nih.gov/pubmed/22990974.

Ren, Y. F. et al. (2012). "Dicer-dependent biogenesis of small RNAs derived from 7SL RNA". In: *PLoS One* 7.7, e40705. ISSN: 1932-6203 (Electronic) 1932-6203 (Linking). DOI: 10.1371/journal.pone.0040705. URL: https://www.ncbi.nlm.nih.gov/pubmed/22808238.

Rome, L., N. Kedersha, and D. Chugani (1991). "Unlocking vaults: organelles in search of a function". In: *Trends Cell Biol* 1.2-3, pp. 47–50. ISSN: 0962-8924 (Print) 0962-8924 (Linking). DOI: 10.1016/0962-8924(91)90088-q. URL: https://www.ncbi.nlm.nih.gov/pubmed/14731565.

RStudio Team (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA. URL: http://www.rstudio.com/.

Ruby, J. G., C. H. Jan, and D. P. Bartel (2007). "Intronic microRNA precursors that bypass Drosha processing". In: *Nature* 448.7149, pp. 83–6. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature05983. URL: https://www.ncbi.nlm.nih.gov/pubmed/17589500.

Rybak-Wolf, A. et al. (2014). "A variety of dicer substrates in human and C. elegans". In: *Cell* 159.5, pp. 1153–1167. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI:

10.1016/j.cell.2014.10.040. URL: https://www.ncbi.nlm.nih.gov/pubmed/25416952.

Saikia, M. et al. (2010). "Genome-wide analysis of N1-methyl-adenosine modification in human tRNAs". In: *RNA* 16.7, pp. 1317–27. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.2057810. URL: https://www.ncbi.nlm.nih.gov/pubmed/20484468.

Sajini, A. A. et al. (2019). "Loss of 5-methylcytosine alters the biogenesis of vault-derived small RNAs to coordinate epidermal differentiation". In: *Nat Commun* 10.1, p. 2550. ISSN: 2041-1723 (Electronic) 2041-1723 (Linking). DOI: 10.1038/s41467-019-10020-7. URL: https://www.ncbi.nlm.nih.gov/pubmed/31186410.

Sarshad, A. A. et al. (2018). "Argonaute-miRNA Complexes Silence Target mRNAs in the Nucleus of Mammalian Stem Cells". In: *Mol Cell* 71.6, 1040–1050 e8. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2018.07.020. URL: https://www.ncbi.nlm.nih.gov/pubmed/30146314.

Saruuldalai, E. et al. (2022). "A host non-coding RNA, nc886, plays a pro-viral role by promoting virus trafficking to the nucleus". In: *Mol Ther Oncolytics* 24, pp. 683–694. ISSN: 2372-7705 (Print) 2372-7705 (Linking). DOI: 10.1016/j.omto.2022.02.018. URL: https://www.ncbi.nlm.nih.gov/pubmed/35284627.

Sasaki, T. et al. (2003). "Identification of eight members of the Argonaute family in the human genome". In: *Genomics* 82.3, pp. 323–30. ISSN: 0888-7543 (Print) 0888-7543 (Linking). DOI: 10.1016/s0888-7543(03)00129-0. URL: https://www.ncbi.nlm.nih.gov/pubmed/12906857.

Scheffer, G. L. et al. (1995). "The drug resistance-related protein LRP is the human major vault protein". In: *Nat Med* 1.6, pp. 578–82. ISSN: 1078-8956 (Print) 1078-8956 (Linking). DOI: 10.1038/nm0695-578. URL: https://www.ncbi.nlm.nih.gov/pubmed/7585126.

Scheper, R. J. et al. (1993). "Overexpression of a M(r) 110,000 vesicular protein in non-P-glycoprotein-mediated multidrug resistance". In: *Cancer Res* 53.7, pp. 1475–9. ISSN: 0008-5472 (Print) 0008-5472 (Linking). URL: https://www.ncbi.nlm.nih.gov/pubmed/7680954.

Schirle, N. T., J. Sheu-Gruttadauria, and I. J. MacRae (2014). "Structural basis for microRNA targeting". In: *Science* 346.6209, pp. 608–13. ISSN: 1095-9203 (Electronic) 0036-8075 (Linking). DOI: 10.1126/science.1258040. URL: https://www.ncbi.nlm.nih.gov/pubmed/25359968.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri (2012). "NIH Image to ImageJ: 25 years of image analysis". In: *Nat Methods* 9.7, pp. 671–5. ISSN: 1548-7105 (Electronic) 1548-7091 (Linking). DOI: 10.1038/nmeth.2089. URL: https://www.ncbi.nlm.nih.gov/pubmed/22930834.

Shi, J. et al. (2019). "tsRNAs: The Swiss Army Knife for Translational Regulation". In: *Trends Biochem Sci* 44.3, pp. 185–189. ISSN: 0968-0004 (Print) 0968-0004 (Linking). DOI: 10.1016/j.tibs.2018.09.007. URL: https://www.ncbi.nlm.nih.gov/pubmed/30297206.

Shurtleff, M. J. et al. (2017). "Broad role for YBX1 in defining the small noncoding RNA composition of exosomes". In: *Proc Natl Acad Sci U S A* 114.43, E8987–E8995. ISSN: 1091-6490 (Electronic) 0027-8424 (Linking). DOI: 10.1073/pnas.1712108114. URL: https://www.ncbi.nlm.nih.gov/pubmed/29073095.

Slesina, M. et al. (2005). "Nuclear localization of the major vault protein in U373 cells". In: *Cell Tissue Res* 321.1, pp. 97–104. ISSN: 0302-766X (Print) 0302-766X (Linking). DOI: 10.1007/s00441-005-1086-8. URL: https://www.ncbi.nlm.nih.gov/pubmed/15902504.

Slesina, M. et al. (2006). "Movement of vault particles visualized by GFP-tagged major vault protein". In: *Cell Tissue Res* 324.3, pp. 403–10. ISSN: 0302-766X (Print) 0302-766X (Linking). DOI: 10.1007/s00441-006-0158-8. URL: https://www.ncbi.nlm.nih.gov/pubmed/16505994.

Sobala, A. and G. Hutvagner (2013). "Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells". In: *RNA Biol* 10.4, pp. 553–63. ISSN: 1555-8584 (Electronic) 1547-6286 (Linking). DOI: 10.4161/rna.24285. URL: https://www.ncbi.nlm.nih.gov/pubmed/23563448.

Sousa-Luis, R. et al. (2021). "POINT technology illuminates the processing of polymerase-associated intact nascent transcripts". In: *Mol Cell* 81.9, 1935–1950 e6. ISSN: 1097-4164

(Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2021.02.034. URL: https://www.ncbi.nlm.nih.gov/pubmed/33735606.

Srinivasan, S. et al. (2019). "Small RNA Sequencing across Diverse Biofluids Identifies Optimal Methods for exRNA Isolation". In: *Cell* 177.2, 446–462 e16. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2019.03.024. URL: https://www.ncbi.nlm.nih.gov/pubmed/30951671.

Stadler, P. F. et al. (2009). "Evolution of vault RNAs". In: *Mol Biol Evol* 26.9, pp. 1975–91. ISSN: 1537-1719 (Electronic) 0737-4038 (Linking). DOI: 10.1093/molbev/msp112. URL: https://www.ncbi.nlm.nih.gov/pubmed/19491402.

Stephen, A. G. et al. (2001). "Assembly of vault-like particles in insect cells expressing only the major vault protein". In: *J Biol Chem* 276.26, pp. 23217–20. ISSN: 0021-9258 (Print) 0021-9258 (Linking). DOI: 10.1074/jbc.C100226200. URL: https://www.ncbi.nlm.nih.gov/pubmed/11349122.

Stirling, D. R. et al. (2021). "CellProfiler 4: improvements in speed, utility and usability". In: *BMC Bioinformatics* 22.1, p. 433. ISSN: 1471-2105 (Electronic) 1471-2105 (Linking). DOI: 10.1186/s12859-021-04344-9. URL: https://www.ncbi.nlm.nih.gov/pubmed/34507520.

Su, X. Y. et al. (2020). "$\mathrm{Circ}_0002232 Acts as a Potential Biomarker for AML and Reveals a Potential ceRNA$ $92a - 3p/PTEN$". In: *Cancer Manag Res* 12, pp. 11871–11881. ISSN: 1179-1322 (Print) 1179-1322 (Linking). DOI: 10.2147/CMAR.S278499. URL: https://www.ncbi.nlm.nih.gov/pubmed/33239917.

Su, Z. et al. (2019). "Angiogenin generates specific stress-induced tRNA halves and is not involved in tRF-3-mediated gene silencing". In: *J Biol Chem* 294.45, pp. 16930–16941. ISSN: 1083-351X (Electronic) 0021-9258 (Linking). DOI: 10.1074/jbc.RA119.009272. URL: https://www.ncbi.nlm.nih.gov/pubmed/31582561.

Su, Z. et al. (2020). "Noncanonical Roles of tRNAs: tRNA Fragments and Beyond". In: *Annu Rev Genet* 54, pp. 47–69. ISSN: 1545-2948 (Electronic) 0066-4197 (Linking). DOI: 10.1146/annurev-genet-022620-101840. URL: https://www.ncbi.nlm.nih.gov/pubmed/32841070.

Subramanian, A. et al. (2005). "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles". In: *Proc Natl Acad Sci U S A* 102.43, pp. 15545–50. ISSN: 0027-8424 (Print) 0027-8424 (Linking). DOI: 10.1073/pnas.0506580102. URL: https://www.ncbi.nlm.nih.gov/pubmed/16199517.

Szklarczyk, D. et al. (2019). "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets". In: *Nucleic Acids Res* 47.D1, pp. D607–D613. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gky1131. URL: https://www.ncbi.nlm.nih.gov/pubmed/30476243.

Szklarczyk, D. et al. (2021). "The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets". In: *Nucleic Acids Res* 49.D1, pp. D605–D612. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkaa1074. URL: https://www.ncbi.nlm.nih.gov/pubmed/33237311.

Takano, K. et al. (2008). "A loss-of-function mutation in the FTSJ1 gene causes non-syndromic X-linked mental retardation in a Japanese family". In: *Am J Med Genet B Neuropsychiatr Genet* 147B.4, pp. 479–84. ISSN: 1552-485X (Electronic) 1552-4841 (Linking). DOI: 10.1002/ajmg.b.30638. URL: https://www.ncbi.nlm.nih.gov/pubmed/18081026.

Telonis, A. G. et al. (2015). "Dissecting tRNA-derived fragment complexities using personalized transcriptomes reveals novel fragment classes and unexpected dependencies". In: *Oncotarget* 6.28, pp. 24797–822. ISSN: 1949-2553 (Electronic) 1949-2553 (Linking). DOI: 10.18632/oncotarget.4695. URL: https://www.ncbi.nlm.nih.gov/pubmed/26325506.

Teng, Y. et al. (2017). "MVP-mediated exosomal sorting of miR-193a promotes colon cancer progression". In: *Nat Commun* 8, p. 14448. ISSN: 2041-1723 (Electronic) 2041-1723 (Linking). DOI: 10.1038/ncomms14448. URL: https://www.ncbi.nlm.nih.gov/pubmed/28211508.

Thomas, E. et al. (2019). "Novel Missense Variants in ADAT3 as a Cause of Syndromic Intellectual Disability". In: *J Pediatr Genet* 8.4, pp. 244–251. ISSN: 2146-4596 (Print)

2146-460X (Linking). DOI: 10.1055/s-0039-1693151. URL: https://www.ncbi.nlm.nih.gov/pubmed/31687266.

Thorvaldsdottir, H., J. T. Robinson, and J. P. Mesirov (2013). "Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration". In: *Brief Bioinform* 14.2, pp. 178–92. ISSN: 1477-4054 (Electronic) 1467-5463 (Linking). DOI: 10.1093/bib/bbs017. URL: https://www.ncbi.nlm.nih.gov/pubmed/22517427.

Torres, A. G. et al. (2014). "A-to-I editing on tRNAs: biochemical, biological and evolutionary implications". In: *FEBS Lett* 588.23, pp. 4279–86. ISSN: 1873-3468 (Electronic) 0014-5793 (Linking). DOI: 10.1016/j.febslet.2014.09.025. URL: https://www.ncbi.nlm.nih.gov/pubmed/25263703.

Van Rossum, Guido and Fred L Drake Jr (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

Vare, V. Y. et al. (2017). "Chemical and Conformational Diversity of Modified Nucleosides Affects tRNA Structure and Function". In: *Biomolecules* 7.1. ISSN: 2218-273X (Electronic) 2218-273X (Linking). DOI: 10.3390/biom7010029. URL: https://www.ncbi.nlm.nih.gov/pubmed/28300792.

Vergani-Junior, C. A. et al. (2021). "DICER: structure, function, and regulation". In: *Biophys Rev* 13.6, pp. 1081–1090. ISSN: 1867-2450 (Print) 1867-2450 (Linking). DOI: 10.1007/s12551-021-00902-w. URL: https://www.ncbi.nlm.nih.gov/pubmed/35059029.

Vilalta, A. et al. (1994). "The rat vault RNA gene contains a unique RNA polymerase III promoter composed of both external and internal elements that function synergistically". In: *J Biol Chem* 269.47, pp. 29752–9. ISSN: 0021-9258 (Print) 0021-9258 (Linking). URL: https://www.ncbi.nlm.nih.gov/pubmed/7525587.

Viphakone, N. et al. (2019). "Co-transcriptional Loading of RNA Export Factors Shapes the Human Transcriptome". In: *Mol Cell* 75.2, 310–323 e8. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2019.04.034. URL: https://www.ncbi.nlm.nih.gov/pubmed/31104896.

Wajahat, M., C. P. Bracken, and A. Orang (2021). "Emerging Functions for snoRNAs and snoRNA-Derived Fragments". In: *Int J Mol Sci* 22.19. ISSN: 1422-0067 (Electronic)

1422-0067 (Linking). DOI: 10.3390/ijms221910193. URL: https://www.ncbi.nlm.nih.gov/pubmed/34638533.

Wakatsuki, S., M. Ohno, and T. Araki (2021). "Human vault RNA1-1, but not vault RNA2-1, modulates synaptogenesis". In: *Commun Integr Biol* 14.1, pp. 61–65. ISSN: 1942-0889 (Print) 1942-0889 (Linking). DOI: 10.1080/19420889.2021.1909229. URL: https://www.ncbi.nlm.nih.gov/pubmed/33948134.

Wang, H. W. et al. (2009). "Structural insights into RNA processing by the human RISC-loading complex". In: *Nat Struct Mol Biol* 16.11, pp. 1148–53. ISSN: 1545-9985 (Electronic) 1545-9985 (Linking). DOI: 10.1038/nsmb.1673. URL: https://www.ncbi.nlm.nih.gov/pubmed/19820710.

Wang, J. et al. (2020). "XPO5 promotes primary miRNA processing independently of RanGTP". In: *Nat Commun* 11.1, p. 1845. ISSN: 2041-1723 (Electronic) 2041-1723 (Linking). DOI: 10.1038/s41467-020-15598-x. URL: https://www.ncbi.nlm.nih.gov/pubmed/32296071.

Wang, Y. et al. (2008). "Structure of the guide-strand-containing argonaute silencing complex". In: *Nature* 456.7219, pp. 209–13. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature07315. URL: https://www.ncbi.nlm.nih.gov/pubmed/18754009.

Warda, A. S. et al. (2017). "Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs". In: *EMBO Rep* 18.11, pp. 2004–2014. ISSN: 1469-3178 (Electronic) 1469-221X (Linking). DOI: 10.15252/embr.201744940. URL: https://www.ncbi.nlm.nih.gov/pubmed/29051200.

Watanabe, K. et al. (1976). "Heat-induced stability of tRNA from an extreme thermophile, Thermus thermophilus". In: *Biochem Biophys Res Commun* 72.3, pp. 1137–44. ISSN: 0006-291X (Print) 0006-291X (Linking). DOI: 10.1016/s0006-291x(76)80250-1. URL: https://www.ncbi.nlm.nih.gov/pubmed/985514.

Waterhouse, A. M. et al. (2009). "Jalview Version 2–a multiple sequence alignment editor and analysis workbench". In: *Bioinformatics* 25.9, pp. 1189–91. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking). DOI: 10.1093/bioinformatics/btp033. URL: https://www.ncbi.nlm.nih.gov/pubmed/19151095.

Wei, J. W. et al. (2017). "Non-coding RNAs as regulators in epigenetics (Review)". In: *Oncol Rep* 37.1, pp. 3–9. ISSN: 1791-2431 (Electronic) 1021-335X (Linking). DOI: 10.3892/or.2016.5236. URL: https://www.ncbi.nlm.nih.gov/pubmed/27841002.

Wei, Z. et al. (2017). "Coding and noncoding landscape of extracellular RNA released by human glioma stem cells". In: *Nat Commun* 8.1, p. 1145. ISSN: 2041-1723 (Electronic) 2041-1723 (Linking). DOI: 10.1038/s41467-017-01196-x. URL: https://www.ncbi.nlm.nih.gov/pubmed/29074968.

Whipple, J. M. et al. (2011). "The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA". In: *Genes Dev* 25.11, pp. 1173–84. ISSN: 1549-5477 (Electronic) 0890-9369 (Linking). DOI: 10.1101/gad.2050711.10.1101/gad.2050711. URL: https://www.ncbi.nlm.nih.gov/pubmed/21632824.

Wu, L., J. Fan, and J. G. Belasco (2006). "MicroRNAs direct rapid deadenylation of mRNA". In: *Proc Natl Acad Sci U S A* 103.11, pp. 4034–9. ISSN: 0027-8424 (Print) 0027-8424 (Linking). DOI: 10.1073/pnas.0510928103. URL: https://www.ncbi.nlm.nih.gov/pubmed/16495412.

Xiao, Y. and I. J. MacRae (2020). "Robust differential microRNA targeting driven by supplementary interactions in vitro". In: *RNA* 26.2, pp. 162–174. ISSN: 1469-9001 (Electronic) 1355-8382 (Linking). DOI: 10.1261/rna.072264.119. URL: https://www.ncbi.nlm.nih.gov/pubmed/31732536.

Xiao, Y. S. et al. (2019). "Major vault protein is a direct target of Notch1 signaling and contributes to chemoresistance in triple-negative breast cancer cells". In: *Cancer Lett* 440-441, pp. 156–167. ISSN: 1872-7980 (Electronic) 0304-3835 (Linking). DOI: 10.1016/j.canlet.2018.09.031. URL: https://www.ncbi.nlm.nih.gov/pubmed/30336197.

Xie, M. et al. (2013). "Mammalian 5'-capped microRNA precursors that generate a single microRNA". In: *Cell* 155.7, pp. 1568–80. ISSN: 1097-4172 (Electronic) 0092-8674 (Linking). DOI: 10.1016/j.cell.2013.11.027. URL: https://www.ncbi.nlm.nih.gov/pubmed/24360278.

Xie, Y. et al. (2020). "Action mechanisms and research methods of tRNA-derived small RNAs". In: *Signal Transduct Target Ther* 5.1, p. 109. ISSN: 2059-3635 (Electronic)

2059-3635 (Linking). DOI: 10.1038/s41392-020-00217-4. URL: https://www.ncbi.nlm.nih.gov/pubmed/32606362.

Xing, F., M. R. Martzen, and E. M. Phizicky (2002). "A conserved family of Saccharomyces cerevisiae synthases effects dihydrouridine modification of tRNA". In: *RNA* 8.3, pp. 370–81. ISSN: 1355-8382 (Print) 1355-8382 (Linking). DOI: 10.1017/s1355838202029825. URL: https://www.ncbi.nlm.nih.gov/pubmed/12003496.

Xing, F. et al. (2004). "The specificities of four yeast dihydrouridine synthases for cytoplasmic tRNAs". In: *J Biol Chem* 279.17, pp. 17850–60. ISSN: 0021-9258 (Print) 0021-9258 (Linking). DOI: 10.1074/jbc.M401221200. URL: https://www.ncbi.nlm.nih.gov/pubmed/14970222.

Xiong, Y. et al. (2011). "MiR-886-3p regulates cell proliferation and migration, and is dysregulated in familial non-medullary thyroid cancer". In: *PLoS One* 6.10, e24717. ISSN: 1932-6203 (Electronic) 1932-6203 (Linking). DOI: 10.1371/journal.pone.0024717. URL: https://www.ncbi.nlm.nih.gov/pubmed/21998631.

Yamasaki, S. et al. (2009). "Angiogenin cleaves tRNA and promotes stress-induced translational repression". In: *J Cell Biol* 185.1, pp. 35–42. ISSN: 1540-8140 (Electronic) 0021-9525 (Linking). DOI: 10.1083/jcb.200811106. URL: https://www.ncbi.nlm.nih.gov/pubmed/19332886.

Yan, Y. et al. (2018). "The sequence features that define efficient and specific hAGO2-dependent miRNA silencing guides". In: *Nucleic Acids Res* 46.16, pp. 8181–8196. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gky546. URL: https://www.ncbi.nlm.nih.gov/pubmed/30239883.

Yang, D. et al. (2002). "Short RNA duplexes produced by hydrolysis with Escherichia coli RNase III mediate effective RNA interference in mammalian cells". In: *Proc Natl Acad Sci U S A* 99.15, pp. 9942–7. ISSN: 0027-8424 (Print) 0027-8424 (Linking). DOI: 10.1073/pnas.152327299. URL: https://www.ncbi.nlm.nih.gov/pubmed/12096193.

Yang, J. et al. (2010). "Vaults are dynamically unconstrained cytoplasmic nanoparticles capable of half vault exchange". In: *ACS Nano* 4.12, pp. 7229–40. ISSN: 1936-086X (Electronic) 1936-0851 (Linking). DOI: 10.1021/nn102051r. URL: https://www.ncbi.nlm.nih.gov/pubmed/21121616.

Yang, J. S. and E. C. Lai (2011). "Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants". In: *Mol Cell* 43.6, pp. 892–903. ISSN: 1097-4164 (Electronic) 1097-2765 (Linking). DOI: 10.1016/j.molcel.2011.07.024. URL: https://www.ncbi.nlm.nih.gov/pubmed/21925378.

Yang, J. S. et al. (2010). "Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis". In: *Proc Natl Acad Sci U S A* 107.34, pp. 15163–8. ISSN: 1091-6490 (Electronic) 0027-8424 (Linking). DOI: 10.1073/pnas.1006432107. URL: https://www.ncbi.nlm.nih.gov/pubmed/20699384.

Yang, R. et al. (2022). "Downregulation of nc886 contributes to prostate cancer cell invasion and TGFbeta1-induced EMT". In: *Genes Dis* 9.4, pp. 1086–1098. ISSN: 2352-3042 (Electronic) 2352-3042 (Linking). DOI: 10.1016/j.gendis.2020.12.010. URL: https://www.ncbi.nlm.nih.gov/pubmed/35685460.

Yarian, C. S. et al. (1999). "Structural and functional roles of the N1- and N3-protons of psi at tRNA's position 39". In: *Nucleic Acids Res* 27.17, pp. 3543–9. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/27.17.3543. URL: https://www.ncbi.nlm.nih.gov/pubmed/10446245.

Ye, W. et al. (2008). "The effect of central loops in miRNA:MRE duplexes on the efficiency of miRNA-mediated gene regulation". In: *PLoS One* 3.3, e1719. ISSN: 1932-6203 (Electronic) 1932-6203 (Linking). DOI: 10.1371/journal.pone.0001719. URL: https://www.ncbi.nlm.nih.gov/pubmed/18320040.

Yeung, M. L. et al. (2009). "Pyrosequencing of small non-coding RNAs in HIV-1 infected cells: evidence for the processing of a viral-cellular double-stranded RNA hybrid". In: *Nucleic Acids Res* 37.19, pp. 6575–86. ISSN: 1362-4962 (Electronic) 0305-1048 (Linking). DOI: 10.1093/nar/gkp707. URL: https://www.ncbi.nlm.nih.gov/pubmed/19729508.

Yu, M. C. et al. (2020). "Differential hypermethylation of the VTRNA2-1 promoter in hepatocellular carcinoma as a prognostic factor: Tumor marker prognostic study". In: *Int J Surg* 79, pp. 282–289. ISSN: 1743-9159 (Electronic) 1743-9159 (Linking). DOI: 10.1016/j.ijsu.2020.05.016. URL: https://www.ncbi.nlm.nih.gov/pubmed/32417463.

Yu, W. et al. (2008). "Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA". In: *Nature* 451.7175, pp. 202–6. ISSN: 1476-4687 (Electronic) 0028-0836 (Linking). DOI: 10.1038/nature06468. URL: https://www.ncbi.nlm.nih.gov/pubmed/18185590.

Yuan, S. et al. (2014). "Methylation by NSun2 represses the levels and function of microRNA 125b". In: *Mol Cell Biol* 34.19, pp. 3630–41. ISSN: 1098-5549 (Electronic) 0270-7306 (Linking). DOI: 10.1128/MCB.00243-14. URL: https://www.ncbi.nlm.nih.gov/pubmed/25047833.

Zaharias, S. et al. (2021). "Intrinsically disordered electronegative clusters improve stability and binding specificity of RNA-binding proteins". In: *J Biol Chem* 297.2, p. 100945. ISSN: 1083-351X (Electronic) 0021-9258 (Linking). DOI: 10.1016/j.jbc.2021.100945. URL: https://www.ncbi.nlm.nih.gov/pubmed/34246632.

Zasloff, M. (1983). "tRNA transport from the nucleus in a eukaryotic cell: carrier-mediated translocation process". In: *Proc Natl Acad Sci U S A* 80.21, pp. 6436–40. ISSN: 0027-8424 (Print) 0027-8424 (Linking). DOI: 10.1073/pnas.80.21.6436. URL: https://www.ncbi.nlm.nih.gov/pubmed/6579529.

Zhang, J. et al. (2014). "PEAR: a fast and accurate Illumina Paired-End reAd mergeR". In: *Bioinformatics* 30.5, pp. 614–20. ISSN: 1367-4811 (Electronic) 1367-4803 (Linking). DOI: 10.1093/bioinformatics/btt593. URL: https://www.ncbi.nlm.nih.gov/pubmed/24142950.

Zhang, K. et al. (2020). "An intellectual disability-associated missense variant in TRMT1 impairs tRNA modification and reconstitution of enzymatic activity". In: *Hum Mutat* 41.3, pp. 600–607. ISSN: 1098-1004 (Electronic) 1059-7794 (Linking). DOI: 10.1002/humu.23976. URL: https://www.ncbi.nlm.nih.gov/pubmed/31898845.

Zhao, Y. N. et al. (2016). "Association of single nucleotide polymorphisms in the MVP gene with platinum resistance and survival in patients with epithelial ovarian cancer". In: *Oncol Lett* 11.4, pp. 2925–2933. ISSN: 1792-1074 (Print) 1792-1074 (Linking). DOI: 10.3892/ol.2016.4311. URL: https://www.ncbi.nlm.nih.gov/pubmed/27073578.

Zon, A. van et al. (2001). "Multiple human vault RNAs. Expression and association with the vault complex". In: *J Biol Chem* 276.40, pp. 37715–21. ISSN: 0021-9258 (Print)

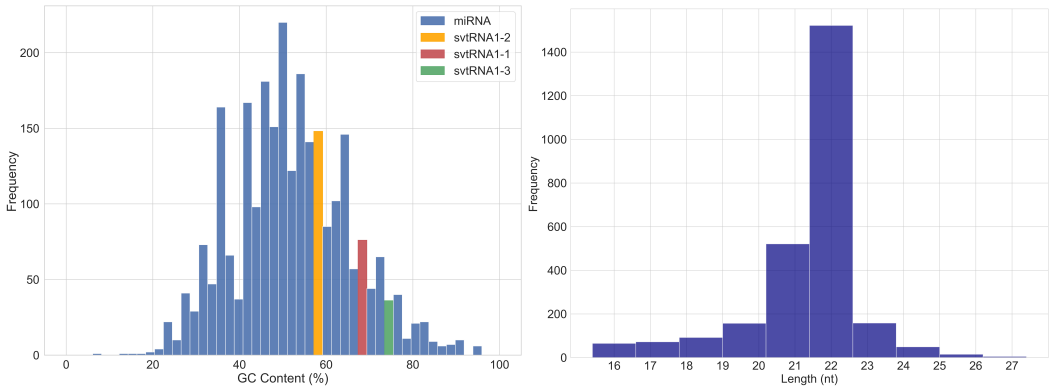
0021-9258 (Linking). DOI: 10.1074/jbc.M106055200. URL: https://www.ncbi.nlm.nih.gov/pubmed/11479319.

Zon, A. van et al. (2003). "The formation of vault-tubes: a dynamic interaction between vaults and vault PARP". In: *J Cell Sci* 116.Pt 21, pp. 4391–400. ISSN: 0021-9533 (Print) 0021-9533 (Linking). DOI: 10.1242/jcs.00749. URL: https://www.ncbi.nlm.nih.gov/pubmed/13130096.

7

Appendix

7.1 Supplementary Figures



(a) **(b)**

Figure 35: ***(a)*** *The histogram shows the GC content percentage distribution in mature microRNA (blue) and svtRNA (yellow, red, green).* ***(b)*** *The distribution of all mature miRNA sequence lengths reported in miRBase release 22.1 (Kozomara et al. 2019) with the most prevalent length of 22 nucleotides.*

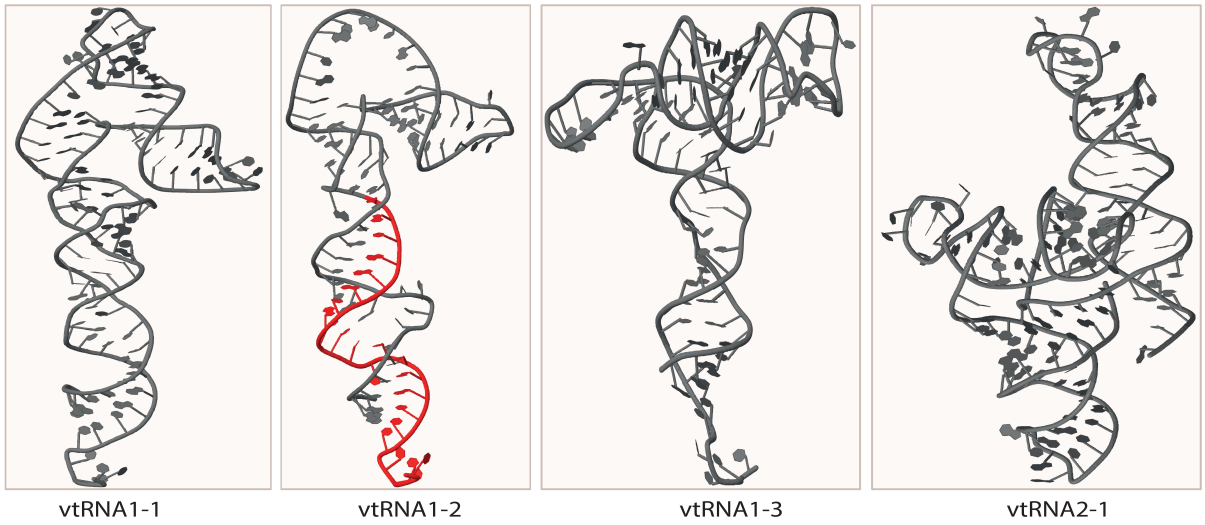


Figure 36: *In silico predicted vault RNA tertiary structures.*

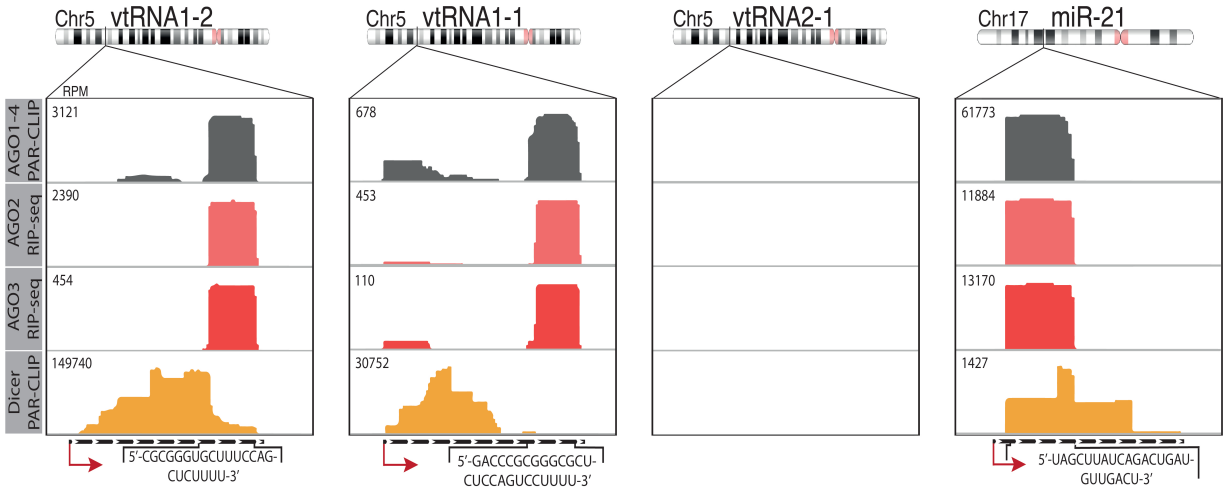


Figure 37: *The IGV tracks depicting vtRNA1-1/2, vtRNA2-1 and miR-21 association of Dicer with each noncoding RNA, compared with normalised AGO1-4 PAR-CLIP and AGO2/3 RIP sequencing. Chromosome ideograms are used to depict locations of selected vtRNA and microRNA genes. The y-axis shows the top range of normalised read count as RPM. Note: no reads were detected for vtRNA2-1. The bottom black line indicates gene annotation and the arrows indicate the orientation of the encoded RNA. The identified small RNA fragment sequences are expanded below.*

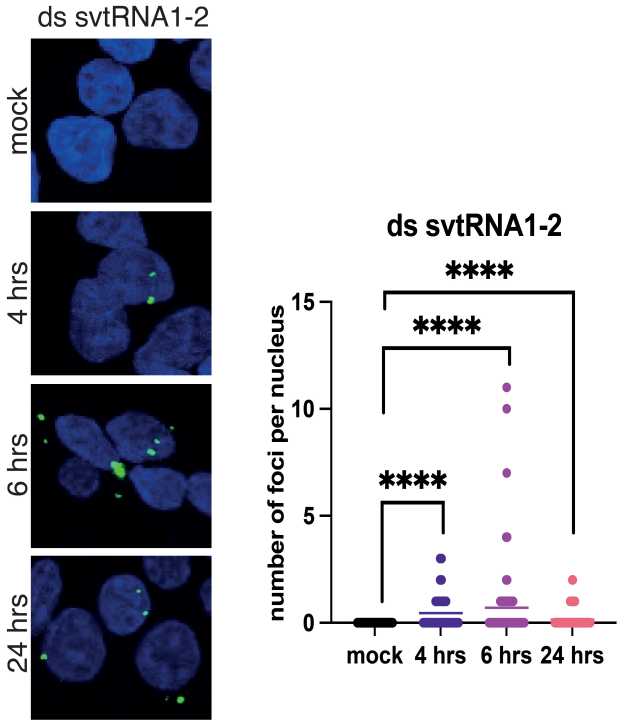


Figure 38: *Confocal microscopy images show nuclear localisation of double-stranded (ds) svtRNA1-2 (green) at different time points. DAPI was used to stain the nuclei of the cells. Images were quantified using Cell profiler software (Stirling et al. 2021). Significance was calculated using t-test (****p-value < 0.0001; ***p-value < 0.001)*

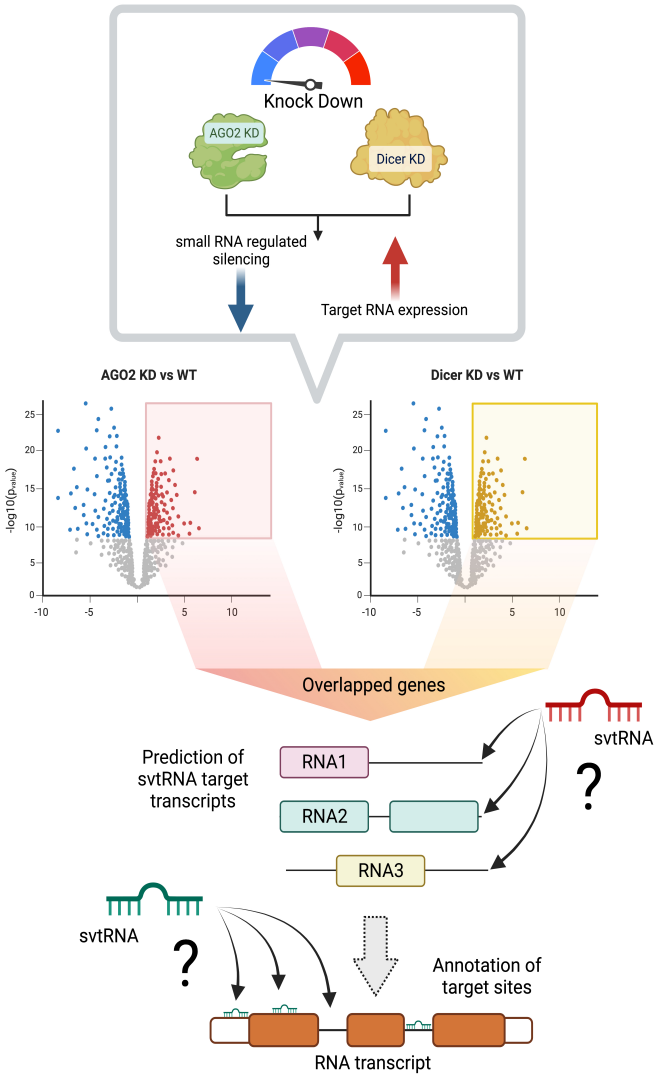


Figure 39: *Computational workflow overview of svtRNA1-2 nascent RNA target predictions.*

7.2 Supplementary Tables

Table 2: Top 20 Gene Ontology terms for 227 predicted svtRNA1-2 target genes

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	GO ID	Genes
0.003544896	4	15	27.02	Cyclic nucleotide catabolic proc.	GO:0009214	PDE10A PDE4D PDE1A PDE7B
0.005578508	4	17	23.84	CGMP metabolic proc.	GO:0046068	PDE10A PDE1A GUCY2D GUCY1A2
0.000281457	7	42	16.89	Cyclic nucleotide metabolic proc.	GO:0009187	PDE10A PDE4D PDE1A ADCY7 GUCY2D GUCY1A2 PDE7B
0.008257965	5	39	12.99	Purine ribonucleotide catabolic proc.	GO:0009154	PDE10A PDE4D PDE1A AMPD3 PDE7B
0.002856957	6	49	12.41	Negative chemotaxis	GO:0050919	SEMA5B LGR6 SEMA3E SLIT3 FLRT2 PLXNA4
0.012071872	5	44	11.51	Ribonucleotide catabolic proc.	GO:0009261	PDE10A PDE4D PDE1A AMPD3 PDE7B
0.006852873	6	61	9.97	Glomerulus development	GO:0032835	MEF2C ENPEP KIRREL3 COL4A3 PDGFD BASP1
0.000363117	9	92	9.91	Reg. of neurotransmitter secretion	GO:0046928	RIMS1 MEF2C PPFIA2 MCTP2 SNCA PPP1R9A NLGN1 MCTP1 RIMS2
0.000281457	10	111	9.13	Reg. of neurotransmitter transport	GO:0051588	RIMS1 MEF2C NOS1 PPFIA2 MCTP2 SNCA PPP1R9A NLGN1 MCTP1 RIMS2
0.012071872	6	71	8.56	CAMP-mediated signaling	GO:0019933	PDE10A PDE4D EPHA5 AKAP6 PDE7B RIMS2
0.001658102	10	146	6.94	Nephron development	GO:0072006	MEF2C ENPEP ADAMTS16 KLHL3 GPC3 KIRREL3 COL4A3 PDGFD BASP1 DCHS2
0.001658102	10	155	6.54	Neurotransmitter secretion	GO:0007269	RIMS1 MEF2C CADPS2 PPFIA2 MCTP2 SNCA PPP1R9A NLGN1 MCTP1 RIMS2
0.001658102	10	155	6.54	Signal release from synapse	GO:0099643	RIMS1 MEF2C CADPS2 PPFIA2 MCTP2 SNCA PPP1R9A NLGN1 MCTP1 RIMS2
0.001658102	12	231	5.26	Neurotransmitter transport	GO:0006836	RIMS1 MEF2C CADPS2 NOS1 PPFIA2 MCTP2 SNCA SLC22A3 PPP1R9A NLGN1 MCTP1 RIMS2
0.001730937	12	235	5.17	Reg. of neurotransmitter levels	GO:0001505	RIMS1 MEF2C CADPS2 NOS1 PPFIA2 MCTP2 SNCA SLC22A3 PPP1R9A NLGN1 MCTP1 RIMS2

Table 2 continued from previous page

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	GO ID	Genes
0.005677176	**13**	**324**	**4.07**	**Reg. of developmental growth**	**GO:0048638**	**RIMS1 MEF2C SEMA5B STAT5A KIAA0319 AKAP6 TRIM46 TSHR SEMA3E RIMS2 BASP1 ERBB4 PLXNA4**
0.012071872	12	312	3.90	Renal system development	GO:0072001	MEF2C ALDH1A2 ENPEP LGR5 ADAMTS16 KLHL3 GPC3 KIRREL3 COL4A3 PDGFD BASP1 DCHS2
0.002297249	**20**	**640**	**3.17**	**Developmental growth**	**GO:0048589**	**RIMS1 MEF2C SEMA5B CPQ ITGA4 STAT5A LGR6 TEC KIAA0319 RSPO2 AKAP6 TRIM46 TSHR SEMA3E RIMS2 BASP1 ERBB4 SLIT3 DMD PLXNA4**
0.008842246	**24**	**985**	**2.47**	**Growth**	**GO:0040007**	**RIMS1 MEF2C SEMA5B CPQ ITGA4 PAPPA2 STAT5A LGR6 TEC IGFBPL1 KIAA0319 GPC3 RSPO2 AKAP6 CDA TRIM46 TSHR SEMA3E RIMS2 BASP1 ERBB4 SLIT3 DMD PLXNA4**
0.006151573	**34**	**1624**	**2.12**	**Generation of neurons**	**GO:0048699**	**CNTN1 PRKCH WNT8B RIMS1 MEF2C ATP8B1 SEMA5B RPGRIP1 PHACTR1 ITGA4 MYB RUNX2 ALDH1A2 LGR6 SOX5 KIAA0319 PPFIA2 EPHA5 RSPO2 KIRREL3 IGSF10 PTPRK PPP1R9A TRIM46 TSHR SPINT1 NLGN1 SEMA3E RIMS2 ERBB4 SLIT3 FLRT2 DMD PLXNA4**

Table 3: 227 predicted svtRNA1-2 target genes

Gene name	Gene ID	Gene Type	L2FC	Description
PTGR1	ENSG00000106853	Protein Coding	13.22027807	AGO2_Dicer_Intersection
MCTP2	ENSG00000140563	Protein Coding	13.12754148	AGO2_Dicer_Intersection
ZNF43	ENSG00000198521	Protein Coding	12.92779747	AGO2_Dicer_Intersection
RPSAP58	ENSG00000288920	Protein Coding	12.1078459	AGO2_Dicer_Intersection
ZNF682	ENSG00000197124	Protein Coding	10.12524757	AGO2_Dicer_Intersection
ENSG00000283515	ENSG00000283515	Protein Coding	8.594989274	AGO2_Dicer_Intersection
ERBB4	**ENSG00000178568**	**Protein Coding**	**8.370910562**	**AGO2_Dicer_Intersection**
GRIA3	ENSG00000125675	Protein Coding	8.139908026	AGO2_Dicer_Intersection
DDX60L	ENSG00000181381	Protein Coding	7.96952771	AGO2_Dicer_Intersection
ZNF331	ENSG00000130844	Protein Coding	7.475267722	AGO2_Dicer_Intersection
ZNF354C	ENSG00000177932	Protein Coding	7.220158853	AGO2_Dicer_Intersection
SLC22A3	ENSG00000146477	Protein Coding	7.139701078	AGO2_Dicer_Intersection
ZNF671	ENSG00000083814	Protein Coding	6.414348728	AGO2_Dicer_Intersection
FLRT2	ENSG00000185070	Protein Coding	6.108626489	AGO2_Dicer_Intersection
ADAMTS16	ENSG00000145536	Protein Coding	5.59231154	AGO2_Dicer_Intersection
PPFIA2	ENSG00000139220	Protein Coding	5.470054463	AGO2_Dicer_Intersection
SCIN	ENSG00000006747	Protein Coding	5.303116411	AGO2_Dicer_Intersection
BASP1	ENSG00000176788	Protein Coding	5.22345946	AGO2_Dicer_Intersection
ZNF91	ENSG00000167232	Protein Coding	5.214817726	AGO2_Dicer_Intersection
VWA2	ENSG00000165816	Protein Coding	5.164072179	AGO2_Dicer_Intersection
PGR	ENSG00000082175	Protein Coding	5.1465595	AGO2_Dicer_Intersection
GRM3	ENSG00000198822	Protein Coding	4.980763321	AGO2_Dicer_Intersection
NAALAD2	ENSG00000077616	Protein Coding	4.971144852	AGO2_Dicer_Intersection
EPHA5	ENSG00000145242	Protein Coding	4.750125155	AGO2_Dicer_Intersection
SEMA3E	ENSG00000170381	Protein Coding	4.65566828	AGO2_Dicer_Intersection
ENSG00000267173	ENSG00000267173	Protein Coding	4.568270654	AGO2_Dicer_Intersection
ZNF808	ENSG00000198482	Protein Coding	4.533344221	AGO2_Dicer_Intersection
HACD4	ENSG00000188921	Protein Coding	4.521126551	AGO2_Dicer_Intersection
MYOM2	ENSG00000036448	Protein Coding	4.50957794	AGO2_Dicer_Intersection
NDST4	ENSG00000138653	Protein Coding	4.361709387	AGO2_Dicer_Intersection
ENSG00000270149	ENSG00000270149	Protein Coding	4.352078811	AGO2_Dicer_Intersection
PPP2R2B	ENSG00000156475	Protein Coding	4.015771274	AGO2_Dicer_Intersection

Table 3 continued from previous page

Gene name	Gene ID	Gene Type	L2FC	Description
SKAP1	ENSG00000141293	Protein Coding	3.892498752	AGO2_Dicer_Intersection
RBMS3	ENSG00000144642	Protein Coding	3.854423287	AGO2_Dicer_Intersection
SPINT1	ENSG00000166145	Protein Coding	3.827547998	AGO2_Dicer_Intersection
CDCP1	ENSG00000163814	Protein Coding	3.793828838	AGO2_Dicer_Intersection
PTPRR	ENSG00000153233	Protein Coding	3.733118086	AGO2_Dicer_Intersection
RBP7	ENSG00000162444	Protein Coding	3.671008993	AGO2_Dicer_Intersection
ABCA9	ENSG00000154258	Protein Coding	3.646243847	AGO2_Dicer_Intersection
PHACTR1	ENSG00000112137	Protein Coding	3.632694402	AGO2_Dicer_Intersection
TXLNB	ENSG00000164440	Protein Coding	3.604362424	AGO2_Dicer_Intersection
ACOXL	ENSG00000153093	Protein Coding	3.570224779	AGO2_Dicer_Intersection
ATP2C2	ENSG00000064270	Protein Coding	3.551813423	AGO2_Dicer_Intersection
GUCY1A2	ENSG00000152402	Protein Coding	3.478199705	AGO2_Dicer_Intersection
IGFBPL1	ENSG00000137142	Protein Coding	3.456070236	AGO2_Dicer_Intersection
C4orf45	ENSG00000164123	Protein Coding	3.4532601	AGO2_Dicer_Intersection
STAT5A	ENSG00000126561	Protein Coding	3.308231963	AGO2_Dicer_Intersection
CASP10	ENSG00000003400	Protein Coding	3.305113367	AGO2_Dicer_Intersection
KCNH1	ENSG00000143473	Protein Coding	3.288576426	AGO2_Dicer_Intersection
PDE7B	ENSG00000171408	Protein Coding	3.280689109	AGO2_Dicer_Intersection
SQOR	ENSG00000137767	Protein Coding	3.226881374	AGO2_Dicer_Intersection
ENSG00000289695	ENSG00000289695	Protein Coding	3.186693608	AGO2_Dicer_Intersection
ADAMTS18	ENSG00000140873	Protein Coding	3.170376259	AGO2_Dicer_Intersection
ENSG00000267360	ENSG00000267360	Protein Coding	3.120309724	AGO2_Dicer_Intersection
CCDC149	ENSG00000181982	Protein Coding	3.110233866	AGO2_Dicer_Intersection
CCDC33	ENSG00000140481	Protein Coding	3.095160173	AGO2_Dicer_Intersection
COL14A1	ENSG00000187955	Protein Coding	3.09230751	AGO2_Dicer_Intersection
NCKAP5	ENSG00000176771	Protein Coding	3.025744549	AGO2_Dicer_Intersection
ENSG00000267022	ENSG00000267022	Protein Coding	3.002518724	AGO2_Dicer_Intersection
COL4A3	ENSG00000169031	Protein Coding	2.934104774	AGO2_Dicer_Intersection
MCTP1	ENSG00000175471	Protein Coding	2.924005056	AGO2_Dicer_Intersection
GALNT5	ENSG00000136542	Protein Coding	2.895955721	AGO2_Dicer_Intersection
PDE1A	ENSG00000115252	Protein Coding	2.84083012	AGO2_Dicer_Intersection
RSPO2	ENSG00000147655	Protein Coding	2.816712999	AGO2_Dicer_Intersection
DCHS2	ENSG00000197410	Protein Coding	2.802784806	AGO2_Dicer_Intersection

Table 3 continued from previous page

Gene name	Gene ID	Gene Type	L2FC	Description
THSD7A	ENSG00000005108	Protein Coding	2.745083064	AGO2_Dicer_Intersection
KIRREL3	ENSG00000149571	Protein Coding	2.742324582	AGO2_Dicer_Intersection
CDK15	ENSG00000138395	Protein Coding	2.714488214	AGO2_Dicer_Intersection
COL8A1	ENSG00000144810	Protein Coding	2.692354099	AGO2_Dicer_Intersection
EFCAB13	ENSG00000178852	Protein Coding	2.670057678	AGO2_Dicer_Intersection
KLHL15	ENSG00000174010	Protein Coding	2.650588142	AGO2_Dicer_Intersection
RUNX2	ENSG00000124813	Protein Coding	2.634217661	AGO2_Dicer_Intersection
MTMR7	ENSG00000003987	Protein Coding	2.627734705	AGO2_Dicer_Intersection
MLH3	ENSG00000119684	Protein Coding	2.602787899	AGO2_Dicer_Intersection
LGR5	ENSG00000139292	Protein Coding	2.590103252	AGO2_Dicer_Intersection
ENSG00000269237	ENSG00000269237	Protein Coding	2.529780084	AGO2_Dicer_Intersection
CR1L	ENSG00000197721	Protein Coding	2.517391877	AGO2_Dicer_Intersection
CNTN1	ENSG00000018236	Protein Coding	2.514062113	AGO2_Dicer_Intersection
PRICKLE2	ENSG00000163637	Protein Coding	2.511403016	AGO2_Dicer_Intersection
MRO	ENSG00000134042	Protein Coding	2.435060445	AGO2_Dicer_Intersection
IGSF10	ENSG00000152580	Protein Coding	2.43503714	AGO2_Dicer_Intersection
VEPH1	ENSG00000197415	Protein Coding	2.426881378	AGO2_Dicer_Intersection
SLC14A2	ENSG00000132874	Protein Coding	2.419601827	AGO2_Dicer_Intersection
STXBP6	ENSG00000168952	Protein Coding	2.349480963	AGO2_Dicer_Intersection
PRKCH	ENSG00000027075	Protein Coding	2.335694979	AGO2_Dicer_Intersection
BEST1	ENSG00000167995	Protein Coding	2.301288317	AGO2_Dicer_Intersection
SOX5	ENSG00000134532	Protein Coding	2.275730271	AGO2_Dicer_Intersection
PCED1B	ENSG00000179715	Protein Coding	2.273316789	AGO2_Dicer_Intersection
ITGA4	ENSG00000115232	Protein Coding	2.272282338	AGO2_Dicer_Intersection
ABCD2	ENSG00000173208	Protein Coding	2.242988574	AGO2_Dicer_Intersection
SEMA5B	ENSG00000082684	Protein Coding	2.221833247	AGO2_Dicer_Intersection
ENSG00000267552	ENSG00000267552	Protein Coding	2.212386646	AGO2_Dicer_Intersection
ANKS1B	ENSG00000185046	Protein Coding	2.210559568	AGO2_Dicer_Intersection
HPSE2	ENSG00000172987	Protein Coding	2.204645222	AGO2_Dicer_Intersection
TICAM1	ENSG00000127666	Protein Coding	2.204055317	AGO2_Dicer_Intersection
VAT1L	ENSG00000171724	Protein Coding	2.194224466	AGO2_Dicer_Intersection
DNAH17	ENSG00000187775	Protein Coding	2.183004114	AGO2_Dicer_Intersection
KIAA0319	ENSG00000137261	Protein Coding	2.179393439	AGO2_Dicer_Intersection

Table 3 continued from previous page

Gene name	Gene ID	Gene Type	L2FC	Description
SLC35F4	ENSG00000151812	Protein Coding	2.165841609	AGO2_Dicer_Intersection
L3MBTL4	ENSG00000154655	Protein Coding	2.165299665	AGO2_Dicer_Intersection
CHSY3	ENSG00000198108	Protein Coding	2.158306168	AGO2_Dicer_Intersection
ADAM12	ENSG00000148848	Protein Coding	2.148262783	AGO2_Dicer_Intersection
DMD	ENSG00000198947	Protein Coding	2.135023215	AGO2_Dicer_Intersection
ATP8B1	ENSG00000081923	Protein Coding	2.125699382	AGO2_Dicer_Intersection
CGNL1	ENSG00000128849	Protein Coding	2.124952781	AGO2_Dicer_Intersection
CADPS2	ENSG00000081803	Protein Coding	2.102060702	AGO2_Dicer_Intersection
ZNF420	ENSG00000197050	Protein Coding	2.091244078	AGO2_Dicer_Intersection
ZNF571	ENSG00000180479	Protein Coding	2.076724931	AGO2_Dicer_Intersection
SCN2A	ENSG00000136531	Protein Coding	2.072834992	AGO2_Dicer_Intersection
ALDH1A2	ENSG00000128918	Protein Coding	2.07032594	AGO2_Dicer_Intersection
PAPPA2	ENSG00000116183	Protein Coding	2.056821209	AGO2_Dicer_Intersection
GPC3	ENSG00000147257	Protein Coding	2.051377889	AGO2_Dicer_Intersection
LTBP2	ENSG00000119681	Protein Coding	2.046016916	AGO2_Dicer_Intersection
PDE4D	ENSG00000113448	Protein Coding	2.025533038	AGO2_Dicer_Intersection
CDA	ENSG00000158825	Protein Coding	2.017454113	AGO2_Dicer_Intersection
PLCL1	ENSG00000115896	Protein Coding	1.989347787	AGO2_Dicer_Intersection
ZNF607	ENSG00000198182	Protein Coding	1.95150901	AGO2_Dicer_Intersection
LAMA4	ENSG00000112769	Protein Coding	1.922079516	AGO2_Dicer_Intersection
NOS1	ENSG00000089250	Protein Coding	1.915491055	AGO2_Dicer_Intersection
TBX4	ENSG00000121075	Protein Coding	1.905696221	AGO2_Dicer_Intersection
MAN1C1	ENSG00000117643	Protein Coding	1.89588807	AGO2_Dicer_Intersection
ZNF780B	ENSG00000128000	Protein Coding	1.890006435	AGO2_Dicer_Intersection
PROCR	ENSG00000101000	Protein Coding	1.887626922	AGO2_Dicer_Intersection
ZNF385B	ENSG00000144331	Protein Coding	1.879007347	AGO2_Dicer_Intersection
ZNF45	ENSG00000124459	Protein Coding	1.877705282	AGO2_Dicer_Intersection
RIMS1	ENSG00000079841	Protein Coding	1.876536865	AGO2_Dicer_Intersection
MVP	ENSG00000013364	Protein Coding	1.846145561	AGO2_Dicer_Intersection
ATP6V1C2	ENSG00000143882	Protein Coding	1.833687687	AGO2_Dicer_Intersection
ADGRG2	ENSG00000173698	Protein Coding	1.825751676	AGO2_Dicer_Intersection
WNT8B	ENSG00000075290	Protein Coding	1.80501672	AGO2_Dicer_Intersection
PLXNA4	ENSG00000221866	Protein Coding	1.798961149	AGO2_Dicer_Intersection

Table 3 continued from previous page

Gene name	Gene ID	Gene Type	L2FC	Description
GUCY2D	ENSG00000132518	Protein Coding	1.787081121	AGO2_Dicer_Intersection
RIMS2	ENSG00000176406	Protein Coding	1.781294617	AGO2_Dicer_Intersection
MELTF	ENSG00000163975	Protein Coding	1.770939397	AGO2_Dicer_Intersection
LGR6	ENSG00000133067	Protein Coding	1.75701586	AGO2_Dicer_Intersection
PLEKHG1	ENSG00000120278	Protein Coding	1.756977793	AGO2_Dicer_Intersection
GSTO2	ENSG00000065621	Protein Coding	1.716360866	AGO2_Dicer_Intersection
SVIL	ENSG00000197321	Protein Coding	1.703172795	AGO2_Dicer_Intersection
MEF2C	ENSG00000081189	Protein Coding	1.702359851	AGO2_Dicer_Intersection
MYB	ENSG00000118513	Protein Coding	1.697501656	AGO2_Dicer_Intersection
LRP1B	ENSG00000168702	Protein Coding	1.693877568	AGO2_Dicer_Intersection
CERKL	ENSG00000188452	Protein Coding	1.693230302	AGO2_Dicer_Intersection
TSHR	ENSG00000165409	Protein Coding	1.676418419	AGO2_Dicer_Intersection
GLIS3	ENSG00000107249	Protein Coding	1.660418755	AGO2_Dicer_Intersection
C1QTNF1	ENSG00000173918	Protein Coding	1.648324585	AGO2_Dicer_Intersection
AMPD3	ENSG00000133805	Protein Coding	1.645443938	AGO2_Dicer_Intersection
RNF144B	ENSG00000137393	Protein Coding	1.643093964	AGO2_Dicer_Intersection
UBE2L6	ENSG00000156587	Protein Coding	1.639113188	AGO2_Dicer_Intersection
EBF2	ENSG00000221818	Protein Coding	1.589876779	AGO2_Dicer_Intersection
JUP	ENSG00000173801	Protein Coding	1.588132545	AGO2_Dicer_Intersection
RPGRIP1	ENSG00000092200	Protein Coding	1.57671924	AGO2_Dicer_Intersection
ALOXE3	ENSG00000179148	Protein Coding	1.570422837	AGO2_Dicer_Intersection
CDH3	ENSG00000062038	Protein Coding	1.566517354	AGO2_Dicer_Intersection
MPP7	ENSG00000150054	Protein Coding	1.563848078	AGO2_Dicer_Intersection
ENSG00000249967	ENSG00000249967	Protein Coding	1.551976338	AGO2_Dicer_Intersection
KLHL3	ENSG00000146021	Protein Coding	1.547873249	AGO2_Dicer_Intersection
TRIM7	ENSG00000146054	Protein Coding	1.547364897	AGO2_Dicer_Intersection
LIFR	ENSG00000113594	Protein Coding	1.540193675	AGO2_Dicer_Intersection
S100Z	ENSG00000171643	Protein Coding	1.532771911	AGO2_Dicer_Intersection
ZNF814	ENSG00000204514	Protein Coding	1.514572713	AGO2_Dicer_Intersection
PRUNE2	ENSG00000106772	Protein Coding	1.504216935	AGO2_Dicer_Intersection
PARP8	ENSG00000151883	Protein Coding	1.503957163	AGO2_Dicer_Intersection
IFNLR1	ENSG00000185436	Protein Coding	1.503028684	AGO2_Dicer_Intersection
MTUS1	ENSG00000129422	Protein Coding	1.496369469	AGO2_Dicer_Intersection

Table 3 continued from previous page

Gene name	Gene ID	Gene Type	L2FC	Description
H6PD	ENSG00000049239	Protein Coding	1.492976501	AGO2_Dicer_Intersection
SLC24A3	ENSG00000185052	Protein Coding	1.485364258	AGO2_Dicer_Intersection
CATSPERG	ENSG00000099338	Protein Coding	1.451978142	AGO2_Dicer_Intersection
PDE10A	ENSG00000112541	Protein Coding	1.428172126	AGO2_Dicer_Intersection
FRMD5	ENSG00000171877	Protein Coding	1.426385272	AGO2_Dicer_Intersection
ZNF566	ENSG00000186017	Protein Coding	1.413167703	AGO2_Dicer_Intersection
SLCO3A1	ENSG00000176463	Protein Coding	1.411493598	AGO2_Dicer_Intersection
ARHGEF16	ENSG00000130762	Protein Coding	1.4111223	AGO2_Dicer_Intersection
ENSG00000283782	ENSG00000283782	Protein Coding	1.409414061	AGO2_Dicer_Intersection
CRYBG2	ENSG00000176092	Protein Coding	1.389840558	AGO2_Dicer_Intersection
ANKRD6	ENSG00000135299	Protein Coding	1.389327544	AGO2_Dicer_Intersection
RCOR2	ENSG00000167771	Protein Coding	1.387969572	AGO2_Dicer_Intersection
STK31	ENSG00000196335	Protein Coding	1.382320154	AGO2_Dicer_Intersection
ENPEP	ENSG00000138792	Protein Coding	1.370342535	AGO2_Dicer_Intersection
SLC4A4	ENSG00000080493	Protein Coding	1.366545412	AGO2_Dicer_Intersection
DNAH2	ENSG00000183914	Protein Coding	1.359822295	AGO2_Dicer_Intersection
CASP8	ENSG00000064012	Protein Coding	1.358779931	AGO2_Dicer_Intersection
SLC9A7	ENSG00000065923	Protein Coding	1.353003861	AGO2_Dicer_Intersection
SNCA	ENSG00000145335	Protein Coding	1.344296243	AGO2_Dicer_Intersection
RNF213	ENSG00000173821	Protein Coding	1.334692948	AGO2_Dicer_Intersection
AGBL1	ENSG00000273540	Protein Coding	1.330869532	AGO2_Dicer_Intersection
COL28A1	ENSG00000215018	Protein Coding	1.329450311	AGO2_Dicer_Intersection
PLAC1	ENSG00000170965	Protein Coding	1.30992532	AGO2_Dicer_Intersection
ADCY7	ENSG00000121281	Protein Coding	1.308749183	AGO2_Dicer_Intersection
ENSG00000269476	ENSG00000269476	Protein Coding	1.305080844	AGO2_Dicer_Intersection
MAP7	ENSG00000135525	Protein Coding	1.275734267	AGO2_Dicer_Intersection
ENSG00000288520	ENSG00000288520	Protein Coding	1.275268903	AGO2_Dicer_Intersection
PTPRK	ENSG00000152894	Protein Coding	1.266143648	AGO2_Dicer_Intersection
PPP1R9A	ENSG00000158528	Protein Coding	1.256670541	AGO2_Dicer_Intersection
QRFPR	ENSG00000186867	Protein Coding	1.255757351	AGO2_Dicer_Intersection
ZNF565	ENSG00000196357	Protein Coding	1.245931225	AGO2_Dicer_Intersection
TRIM46	ENSG00000163462	Protein Coding	1.239239113	AGO2_Dicer_Intersection
MARVELD2	ENSG00000152939	Protein Coding	1.220706176	AGO2_Dicer_Intersection

Table 3 continued from previous page

Gene name	Gene ID	Gene Type	L2FC	Description
AKAP6	ENSG00000151320	Protein Coding	1.215981489	AGO2_Dicer_Intersection
KCNMA1	ENSG00000156113	Protein Coding	1.21255527	AGO2_Dicer_Intersection
ANXA4	ENSG00000196975	Protein Coding	1.210846947	AGO2_Dicer_Intersection
FIGNL2	ENSG00000261308	Protein Coding	1.199942875	AGO2_Dicer_Intersection
CPQ	ENSG00000104324	Protein Coding	1.174916334	AGO2_Dicer_Intersection
NLGN1	ENSG00000169760	Protein Coding	1.173342944	AGO2_Dicer_Intersection
DAPK2	ENSG00000035664	Protein Coding	1.171530313	AGO2_Dicer_Intersection
PDGFD	ENSG00000170962	Protein Coding	1.168742863	AGO2_Dicer_Intersection
CLSTN2	ENSG00000158258	Protein Coding	1.163190193	AGO2_Dicer_Intersection
KCNQ4	ENSG00000117013	Protein Coding	1.160414818	AGO2_Dicer_Intersection
PLEKHG2	ENSG00000090924	Protein Coding	1.139752529	AGO2_Dicer_Intersection
NCOA7	ENSG00000111912	Protein Coding	1.132722732	AGO2_Dicer_Intersection
MYO10	ENSG00000145555	Protein Coding	1.128500686	AGO2_Dicer_Intersection
OLFM2	ENSG00000105088	Protein Coding	1.099747304	AGO2_Dicer_Intersection
ENSG00000272647	ENSG00000272647	Protein Coding	1.099211905	AGO2_Dicer_Intersection
FKBP9	ENSG00000122642	Protein Coding	1.095395477	AGO2_Dicer_Intersection
HTR7	ENSG00000148680	Protein Coding	1.086049568	AGO2_Dicer_Intersection
TEC	ENSG00000135605	Protein Coding	1.080404925	AGO2_Dicer_Intersection
LSR	ENSG00000105699	Protein Coding	1.077786724	AGO2_Dicer_Intersection
RGS7	ENSG00000182901	Protein Coding	1.075791903	AGO2_Dicer_Intersection
SIPA1L2	ENSG00000116991	Protein Coding	1.070528191	AGO2_Dicer_Intersection
CAST	ENSG00000153113	Protein Coding	1.066608869	AGO2_Dicer_Intersection
ZNF157	ENSG00000147117	Protein Coding	1.066283828	AGO2_Dicer_Intersection
TBC1D5	ENSG00000131374	Protein Coding	1.05214538	AGO2_Dicer_Intersection
SRGAP3	ENSG00000196220	Protein Coding	1.046955672	AGO2_Dicer_Intersection
COL25A1	ENSG00000188517	Protein Coding	1.044626339	AGO2_Dicer_Intersection
CACNA2D3	ENSG00000157445	Protein Coding	1.028979802	AGO2_Dicer_Intersection
SLIT3	ENSG00000184347	Protein Coding	1.020545913	AGO2_Dicer_Intersection
DTNA	ENSG00000134769	Protein Coding	1.00587463	AGO2_Dicer_Intersection
AHNAK2	ENSG00000185567	Protein Coding	1.002373804	AGO2_Dicer_Intersection

Table 4: Top 10 Gene Ontology terms for Up-regulated genes (vtRNA1-2 KO)

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	GO ID	Genes
6.51E-14	49	451	4.24824575	Reg. of membrane potential	GO:0042391	CFTR CACNA1G KCNQ1 TRPC5 CACNG4 CHRNA3 KCNK2 STX1B CACNA1I KCNK10 SLC8A3 CHRNA4 GABRE NALCN FGF14 CBLN1 SLC1A6 ASIC2 FGF12 GRIA3 IFI6 FGF13 KCNH3 HCN4 CACNA1A KCNH1 PID1 PIEZO2 KCNMA1 DGKI KCNB1 SHANK1 TAFA4 GRIK3 KCNK9 SLC26A5 KCNMB3 KCND3 KCNH6 CNIH2 SCN4B KCNA2 GRIN2A KCNQ3 KCNK12 KCND2 P2RX2 RELN CACNA1E
1.26E-09	45	533	3.30122707	Reg. of ion transmembrane transport	GO:0034765	CFTR CACNA1G KCNQ1 RASGRF1 TRPM5 CACNG5 CACNG4 CACNA1I KCNK10 NALCN FGF14 CACNG7 ASIC2 FGF12 OPRL1 GRIA3 HSPA2 FGF13 DPP6 CACNG6 KCNH3 HCN4 CACNA1A KCNH1 KCNMA1 KCNB1 SHANK1 GSG1L SLC26A5 KCNS3 KCNMB3 KCND3 KCNH6 CNIH2 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 RELN KEL CACNA1E
3.79E-12	61	763	3.12604619	Reg. of ion transport	GO:0043269	CFTR CACNA1G KCNQ1 RASGRF1 TRPM5 CACNG5 CACNG4 CHRNA3 GNAO1 PDGFB CACNA1I KCNK10 CHRNA4 GABRE NALCN FGF14 CACNG7 ASIC2 FGF12 GRIN3B OPRL1 GRIA3 HSPA2 BEST3 FGF13 DPP6 CACNG6 LGALS3 SYT11 KCNH3 HCN4 TRPC3 CACNA1A KCNH1 SYT2 KCNMA1 KCNB1 SHANK1 AVPR1A TRPV3 GSG1L SLC26A5 SYT9 KCNS3 KCNMB3 KCND3 IL16 KCNH6 CNIH2 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 P2RX2 RELN KEL CACNA1E
1.61E-12	72	986	2.85526009	Cation transmembrane transport	GO:0098655	CACNA1G KCNQ1 RASGRF1 TRPM5 TRPC5 CLCN4 VPS9D1 CACNG5 CACNG4 KCNK2 CACNA1I KCNK10 SLC8A3 NALCN FGF14 SLC6A2 CACNG7 KCNT1 ASIC2 SLC1A2 TRPV4 FGF12 PEX5L GRIN3B OPRL1 GRIA3 HSPA2 FGF13 DPP6 CACNG6 ANO1 CNGA4 KCNH3 HCN4 TRPC3 CACNA1A KCNH1 UNC80 SHROOM2 PIEZO2 SLC24A2 KCNMA1 ATP2B2 KCNB1 SHANK1 SLC6A5 TTYH1 TRPV3 SCARA5 GSG1L KCNK9 KCNS3 KCNMB3 KCND3 KCNH6 CNIH2 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 SLC24A3 P2RX2 RELN SLC6A17 KEL CACNA1E FAM155A NOX5
4.56E-10	63	909	2.70998421	Inorganic ion transmembrane transport	GO:0098660	CFTR CACNA1G KCNQ1 TRPM5 TRPC5 CLCN4 VPS9D1 CACNG5 CACNG4 KCNK2 CACNA1I KCNK10 SLC8A3 GABRE NALCN FGF14 SLC6A2 CACNG7 KCNT1 ASIC2 TRPV4 FGF12 GRIN3B OPRL1 HSPA2 FGF13 DPP6 CACNG6 ANO1 KCNH3 HCN4 TRPC3 CACNA1A KCNH1 SHROOM2 SLC24A2 KCNMA1 ATP2B2 KCNB1 SLC6A5 TTYH1 TRPV3 SCARA5 KCNK9 SLC26A5 KCNS3 KCNMB3 KCND3 KCNH6 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 SLC24A3 SLC6A17 KEL CACNA1E FAM155A NOX5
2.07E-10	67	984	2.6623785	Metal ion transport	GO:0030001	CACNA1G KCNQ1 TRPM5 TRPC5 CACNG5 CACNG4 KCNK2 GNAO1 PDGFB CACNA1I KCNK10 SLC8A3 CHRNA4 NALCN FGF14 SLC6A2 CACNG7 KCNT1 CDH23 ASIC2 TRPV4 FGF12 GRIN3B OPRL1 HSPA2 FGF13 DPP6 CACNG6 LGALS3 KCNH3 HCN4 TRPC3 SLC13A5 CACNA1A KCNH1 SHROOM2 SLC24A2 KCNMA1 ATP2B2 KCNB1 SLC6A5 TTYH1 TRPV3 SCARA5 KCNK9 KCNS3 KCNMB3 KCND3 NMUR1 IL16 KCNH6 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 SLC24A3 P2RX2 SLC6A17 KEL CACNA1E FAM155A NOX5 SLC5A8

Table 4 continued from previous page

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	GO ID	Genes
5.57E-13	86	1286	2.61485479	Cation transport	GO:0006812	CACNA1G KCNQ1 RASGRF1 TRPM5 TRPC5 CLCN4 VPS9D1 CACNG5 CACNG4 CHRNA3 KCNK2 GNAO1 PDGFB CACNA1I KCNK10 SLC8A3 CHRNA4 NALCN FGF14 SLC6A2 CACNG7 KCNT1 CDH23 ASIC2 SLC1A2 TRPV4 FGF12 PEX5L GRIN3B OPRL1 GRIA3 HSPA2 FGF13 DPP6 CACNG6 SLC7A10 ANO1 LGALS3 CNGA4 SYT11 KCNH3 HCN4 TRPC3 SLC13A5 CACNA1A KCNH1 SYT2 UNC80 SHROOM2 PIEZO2 SLC24A2 KCNMA1 ATP2B2 KCNB1 SHANK1 SLC6A5 TTYH1 TRPV3 SCARA5 GSG1L KCNK9 SYT9 KCNS3 KCNMB3 KCND3 NMUR1 IL16 KCNH6 CNIH2 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 SLC24A3 P2RX2 RELN SLC6A17 KEL CACNA1E FAM155A NOX5 SLC5A8
3.79E-12	82	1251	2.56298837	Ion transmembrane transport	GO:0034220	CFTR CACNA1G KCNQ1 RASGRF1 TRPM5 TRPC5 CLCN4 VPS9D1 CACNG5 CACNG4 CHRNA3 KCNK2 CACNA1I KCNK10 SLC8A3 CHRNA4 GABRE NALCN FGF14 SLC6A2 SLC1A6 CACNG7 KCNT1 ASIC2 SLC1A2 TRPV4 FGF12 PEX5L GRIN3B OPRL1 GRIA3 HSPA2 FGF13 DPP6 CACNG6 ANO1 CNGA4 KCNH3 HCN4 TRPC3 SLC13A5 CACNA1A KCNH1 UNC80 SHROOM2 SLC16A12 PIEZO2 SLC24A2 KCNMA1 ATP2B2 KCNB1 SHANK1 GRIK3 SLC6A5 TTYH1 TRPV3 SCARA5 GSG1L KCNK9 SLC26A5 KCNS3 KCNMB3 KCND3 KCNH6 CNIH2 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 SLC24A3 P2RX2 RELN SLC6A17 KEL CACNA1E FAM155A NOX5 SLC5A8
2.07E-10	96	1720	2.1823926	Ion transport	GO:0006811	CFTR CACNA1G KCNQ1 RASGRF1 TRPM5 TRPC5 CLCN4 VPS9D1 CACNG5 CACNG4 CHRNA3 KCNK2 GNAO1 PDGFB CACNA1I KCNK10 SLC8A3 CHRNA4 GABRE NALCN FGF14 SLC6A2 SLC1A6 CACNG7 KCNT1 CDH23 ASIC2 SLC1A2 TRPV4 FGF12 PEX5L GRIN3B OPRL1 GRIA3 HSPA2 BEST3 FGF13 DPP6 CACNG6 SLC7A10 ANO1 LGALS3 CNGA4 SYT11 ANXA1 KCNH3 HCN4 TRPC3 SLC13A5 CACNA1A KCNH1 SYT2 UNC80 SHROOM2 SLC16A12 PIEZO2 SLC24A2 KCNMA1 ATP2B2 KCNB1 SHANK1 GRIK3 SLC6A5 AVPR1A TTYH1 TRPV3 SCARA5 GSG1L KCNK9 SLC26A5 SYT9 KCNS3 KCNMB3 KCND3 NMUR1 IL16 KCNH6 CNIH2 SLCO3A1 SCN4B KCNA3 KCNA2 KCNG2 GRIN2A KCNQ3 KCNK12 KCND2 SLC24A3 P2RX2 RELN SLC6A17 KEL CACNA1E FAM155A NOX5 SLC5A8
5.16E-10	96	1757	2.13643441	Neurogenesis	GO:0022008	ARHGAP33 ARX ANOS1 TNFRSF1B TNC USH2A LAMC2 RASGRF1 DGKG NGFR TRPC5 SEMA3C PLXNA2 MAP2 CHRNA3 GNAO1 STX1B EFNA2 SLC8A3 CBLN1 CACNG7 LHX2 CDH23 NEURL1 RND2 TRPV4 FZD10 MDGA1 PROX1 TRIM67 INHBA BHLHE41 MYRF BHLHE23 FOXA2 BMP2 MMP24 FGF13 OLFM1 ATP8A2 NGF MYCN ANXA1 DTX1 DOCK10 PLXNC1 MMD2 NTRK3 CDH11 DUSP10 SYT2 CNTN4 TENM2 ANKRD1 NKX6-2 ASTN1 SPOCK1 UNC5D ATP2B2 NRG1 SHANK1 ELAVL4 NFASC FAT3 SPINT1 DCHS1 BMERB1 SEMA6B GFRA2 CTNND2 ALK MAP6 KNDC1 FOXB1 NEGR1 DSCAML1 ADAMTSL1 ERBB4 CDH4 TMEM132E GRIN2A SLIT3 FLRT2 NTF3 RTN4RL1 AGBL4 FSCN2 DNER RELN NTNG2 KEL BRINP2 L1CAM PLXNA4 STMND1 ETV5

Table 5: Top 10 Gene Ontology terms for Down-regulated genes (vtRNA1-2 KO)

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	GO ID	Genes
4.01E-05	15	225	5.54647202	Reg. of synapse organization	GO:0050807	MEF2C SPARC FLRT3 ADGRB3 PPFIA2 GPM6A SLC7A11 PTPRO GRID2 GHRL CAMKV SLITRK5 LRFN5 IL1RAPL1 FRMPD4
1.83E-05	22	451	4.05839416	Synapse organization	GO:0050808	MEF2C NDRG1 DKK1 SPARC PCDH17 FLRT3 ADGRB3 PPFIA2 ADGRL3 GPM6A SLC7A11 PTPRO GRID2 GHRL CAMKV SLITRK5 LRFN5 IL1RAPL1 FRMPD4 SEMA3E GAP43 LRRTM4
3.26E-05	27	686	3.27452065	Neuron projection morphogenesis	GO:0048812	EPHA3 SEMA3A EPHA6 PTPRZ1 RASAL1 ITGA4 PLPPR4 FLRT3 ADGRB3 PPFIA2 EPHA5 DCDC2 GPM6A PTPRO GJA1 SEMA3D UCHL1 SLITRK5 IL1RAPL1 ALCAM SEMA3E GAP43 ATOH1 UGT8 UNC5C OSTN LAMA2
2.10E-05	29	757	3.18720651	Cell junction organization	GO:0034330	CDH10 EPHA3 MEF2C NDRG1 DKK1 SPARC PCDH17 FLRT3 ADGRB3 PPFIA2 CDH18 ADGRL3 GPM6A SLC7A11 PTPRO GRID2 GJA1 CDH12 GHRL CAMKV SLITRK5 LRFN5 IL1RAPL1 FRMPD4 SEMA3E GAP43 UGT8 CADM2 LRRTM4
7.09E-08	40	1045	3.18457724	Neuron projection development	GO:0031175	CNTN1 EPHA3 SEMA3A EPHA6 MEF2C PTPRZ1 RASAL1 DPYSL3 ITGA4 PLPPR5 PLPPR4 FLRT3 SERPINF1 NREP ADGRB3 MYO7A PPFIA2 EPHA5 DCDC2 GPM6A PTPRO GRID2 GJA1 SEMA3D UCHL1 NCAM2 GHRL SLITRK5 IL1RAPL1 ALCAM SEMA3E GAP43 ATOH1 FUT9 UGT8 UNC5C OSTN LAMA2 ALKAL1 TOX
5.22E-08	44	1203	3.04295223	Neuron development	GO:0048666	CNTN1 EPHA3 SEMA3A SLC1A3 EPHA6 MEF2C PTPRZ1 RASAL1 DPYSL3 ITGA4 PLPPR5 PLPPR4 FLRT3 SERPINF1 NREP ADGRB3 MYO7A PPFIA2 EPHA5 DCDC2 GPM6A PTPRO GRID2 GJA1 SEMA3D UCHL1 NCAM2 GHRL LHX8 UQCRQ SLITRK5 IL1RAPL1 ALCAM SEMA3E GAP43 ATOH1 FUT9 UGT8 UNC5C OSTN LAMA2 ALKAL1 TOX RORB

Table 5 continued from previous page

Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathway	GO ID	Genes
6.36E-07	39	1085	2.99049413	Cell morphogenesis	GO:0000902	CDH10 EPHA3 SEMA3A SLC1A3 EPHA6 MEF2C LAMB4 PTPRZ1 RASAL1 SPARC ITGA4 RND3 PLPPR4 MYL12B FLRT3 ADGRB3 MYO7A PARVG PPFIA2 EPHA5 CDH18 DCDC2 GPM6A PTPRO GJA1 SEMA3D CDH12 UCHL1 NBEAL2 SLITRK5 IL1RAPL1 ALCAM SEMA3E GAP43 ATOH1 UGT8 UNC5C OSTN LAMA2
3.83E-09	55	1624	2.81763511	Generation of neurons	GO:0048699	CNTN1 EPHA3 SEMA3A SLC1A3 EPHA6 MEF2C PTPRZ1 DKK1 RASAL1 DPYSL3 BCHE ITGA4 PLPPR5 PLPPR4 FLRT3 SERPINF1 SPINK5 NREP ADGRB3 MYC MYO7A PPFIA2 MDGA2 EPHA5 DCDC2 ADGRL3 GPM6A VEGFC PTPRO GRID2 IGSF10 GJA1 SEMA3D UCHL1 NCAM2 GHRL LHX8 UQCRQ SLITRK5 DDIT4 IL1RAPL1 ALCAM SEMA3E GAP43 ATOH1 FUT9 UGT8 HES7 BHLHA15 UNC5C OSTN LAMA2 ALKAL1 TOX RORB
7.09E-08	49	1473	2.76758787	Neuron differentiation	GO:0030182	CNTN1 EPHA3 SEMA3A SLC1A3 EPHA6 MEF2C PTPRZ1 DKK1 RASAL1 DPYSL3 ITGA4 PLPPR5 PLPPR4 FLRT3 SERPINF1 SPINK5 NREP ADGRB3 MYO7A PPFIA2 MDGA2 EPHA5 DCDC2 GPM6A PTPRO GRID2 GJA1 SEMA3D UCHL1 NCAM2 GHRL LHX8 UQCRQ SLITRK5 DDIT4 IL1RAPL1 ALCAM SEMA3E GAP43 ATOH1 FUT9 UGT8 BHLHA15 UNC5C OSTN LAMA2 ALKAL1 TOX RORB
3.83E-09	58	1757	2.74640333	Neurogenesis	GO:0022008	CNTN1 EPHA3 SEMA3A SLC1A3 EPHA6 MEF2C NDRG1 PTPRZ1 DKK1 RASAL1 DPYSL3 BCHE ITGA4 PLPPR5 PLPPR4 FLRT3 SERPINF1 TRPC4 SPINK5 NREP ADGRB3 MYC MYO7A PPFIA2 MDGA2 EPHA5 DCDC2 ADGRL3 GPM6A VEGFC PTPRO GRID2 IGSF10 GJA1 SEMA3D UCHL1 NCAM2 GHRL LHX8 KLF15 UQCRQ SLITRK5 DDIT4 IL1RAPL1 ALCAM SEMA3E GAP43 ATOH1 FUT9 UGT8 HES7 BHLHA15 UNC5C OSTN LAMA2 ALKAL1 TOX RORB

Gene ID	Gene Symbol	Gene Name	Gene % GC content	log2FC	Gene Type
ENSG00000116183	PAPPA2	pappalysin 2	39.35	1.03209129	Protein Coding
ENSG00000119681	LTBP2	latent transforming growth factor beta binding protein 2	51.68	1.415322159	Protein Coding
ENSG00000120278	PLEKHG1	pleckstrin homology and RhoGEF domain containing G1	42.32	1.254535069	Protein Coding
ENSG00000125675	GRIA3	glutamate ionotropic receptor AMPA type subunit 3	39.27	2.921167577	Protein Coding
ENSG00000132874	SLC14A2	solute carrier family 14 member 2	42.88	3.117133393	Protein Coding
ENSG00000137142	IGFBPL1	insulin like growth factor binding protein like 1	49.21	2.584936553	Protein Coding
ENSG00000140873	ADAMTS18	ADAM metallopeptidase with thrombospondin type 1 motif 18	39.74	2.714650548	Protein Coding
ENSG00000141293	SKAP1	src kinase associated phosphoprotein 1	39.13	1.480551075	Protein Coding
ENSG00000143473	KCNH1	potassium voltage-gated channel subfamily H member 1	41.06	3.627190405	Protein Coding
ENSG00000148848	ADAM12	ADAM metallopeptidase domain 12	43.11	4.807687037	Protein Coding
ENSG00000156113	KCNMA1	potassium calcium-activated channel subfamily M alpha 1	43.94	1.011746296	Protein Coding
ENSG00000162444	RBP7	retinol binding protein 7	48.33	1.155617161	Protein Coding
ENSG00000166145	SPINT1	serine peptidase inhibitor, Kunitz type 1	53.63	1.310231226	Protein Coding
ENSG00000170962	PDGFD	platelet derived growth factor D	37.11	1.441645195	Protein Coding
ENSG00000171724	VAT1L	vesicle amine transport 1 like	44.32	2.770894449	Protein Coding
ENSG00000173918	C1QTNF1	C1q and TNF related 1	54.12	1.02378412	Protein Coding
ENSG00000176463	SLCO3A1	solute carrier organic anion transporter family member 3A1	45.12	1.514743137	Protein Coding
ENSG00000178568	ERBB4	erb-b2 receptor tyrosine kinase 4	34.69	1.432193313	Protein Coding
ENSG00000181381	DDX60L	DExD/H-box 60 like	38.14	2.060993407	Protein Coding
ENSG00000184347	SLIT3	slit guidance ligand 3	45.01	1.730049183	Protein Coding
ENSG00000185052	SLC24A3	solute carrier family 24 member 3	43.88	1.458172478	Protein Coding
ENSG00000185070	FLRT2	fibronectin leucine rich transmembrane protein 2	39.33	3.093101076	Protein Coding
ENSG00000187775	DNAH17	dynein axonemal heavy chain 17	52.35	1.87299172	Protein Coding
ENSG00000197721	CR1L	complement C3b/C4b receptor 1 like	40.5	1.776321395	Protein Coding
ENSG00000221818	EBF2	EBF transcription factor 2	42.02	1.316008864	Protein Coding
ENSG00000221866	PLXNA4	plexin A4	46.94	2.952397277	Protein Coding
ENSG00000261308	FIGNL2	fidgetin like 2	55.63	1.110608308	Protein Coding
ENSG00000267022	K7ENM7	novel protein	41.65	1.229459564	Protein Coding
ENSG00000273540	AGBL1	AGBL carboxypeptidase 1	39.11	2.924426853	Protein Coding

Table 6: 29 predicted svtRNA1-2 gene targets

Box : svtRNA1-2 WT miRanda predicted bindings

```
Read Sequence:ENSG00000181143:ENST00000397910:5.73969207442768:protein_coding::chr19:8848843-8981342(-) (132499 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000181143:ENST00000397910:5.73969207442768:protein_coding::chr19:8848843-8981342(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 160.000000  Q:2 to 22  R:55380 to 55407 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ggAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.799999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000181143:ENST00000397910:5.73969207442768:protein_coding::chr19:8848843-8981342(-) 160.00 -31.80 2 22 55380 55407 25
64.00% 76.00%
--
Read Sequence:ENSG00000148848:ENST00000448723:4.80768703711831:protein_coding::chr10:126012390-126388477(-) (376087 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000148848:ENST00000448723:4.80768703711831:protein_coding::chr10:126012390-126388477(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 160.000000  Q:2 to 22  R:30407 to 30434 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000148848:ENST00000448723:4.80768703711831:protein_coding::chr10:126012390-126388477(-) 160.00 -31.53 2 22
30407 30434 25 64.00% 76.00%
--
Read Sequence:ENSG00000143473:ENST00000271751:3.62719040509158:protein_coding::chr1:210678313-211134148(-) (455835 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000143473:ENST00000271751:3.62719040509158:protein_coding::chr1:210678313-211134148(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 160.000000  Q:2 to 22  R:9612 to 9639 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGCAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000143473:ENST00000271751:3.62719040509158:protein_coding::chr1:210678313-211134148(-) 160.00 -31.53 2 22 9612 9639
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:258453 to 258480 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'
```

```
   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000143473:ENST00000271751:3.62719040509158:protein_coding::chr1:210678313-211134148(-) 160.00 -31.40 2 22
258453 258480 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:306920 to 306947 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000143473:ENST00000271751:3.62719040509158:protein_coding::chr1:210678313-211134148(-) 160.00 -31.53 2 22
306920 306947 25 64.00% 76.00%
--
Read Sequence:ENSG00000144406:ENST00000673920:3.43142738293305:protein_coding::chr2:209771831-209999296(+) (227465 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000144406:ENST00000673920:3.43142738293305:protein_coding::chr2:209771831-209999296(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:85082 to 85110 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000144406:ENST00000673920:3.43142738293305:protein_coding::chr2:209771831-209999296(+) 163.00 -31.81 2 22 85082 85110
26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:177771 to 177798 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000144406:ENST00000673920:3.43142738293305:protein_coding::chr2:209771831-209999296(+) 160.00 -31.40 2 22
177771 177798 25 64.00% 76.00%
--
Read Sequence:ENSG00000188833:ENST00000371506:3.26959109553511:protein_coding::chr9:137434363-137441357(-) (6994 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000188833:ENST00000371506:3.26959109553511:protein_coding::chr9:137434363-137441357(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 173.000000  Q:2 to 22  R:1856 to 1878 Align Len (20) (75.00%) (85.00%)

   Query:    3' uuUUCUCGACCUUUCGUGGGCGc 5'
                 |:|| | ||: |||||||||
   Ref:      5' ucAGGACCAGGGCAGCACCCGCg 3'
```

```
   Energy:  -30.820000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000188833:ENST00000371506:3.26959109553511:protein_coding::chr9:137434363-137441357(-) 173.00 -30.82 2 22 1856 1878
20 75.00% 85.00%
--
Read Sequence:ENSG00000132874:ENST00000255226:3.11713339314946:protein_coding::chr18:45615475-45683688(+) (68213 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000132874:ENST00000255226:3.11713339314946:protein_coding::chr18:45615475-45683688(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:17326 to 17354 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' acAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000132874:ENST00000255226:3.11713339314946:protein_coding::chr18:45615475-45683688(+) 163.00 -31.81 2 22 17326 17354
26 69.23% 76.92%
--
Read Sequence:ENSG00000185070:ENST00000330753:3.09310107619429:protein_coding::chr14:85530143-85654428(+) (124285 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000185070:ENST00000330753:3.09310107619429:protein_coding::chr14:85530143-85654428(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:40563 to 40590 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUCCAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000185070:ENST00000330753:3.09310107619429:protein_coding::chr14:85530143-85654428(+) 160.00 -31.40 2 22 40563 40590
25 64.00% 76.00%
--
Read Sequence:ENSG00000152953:ENST00000282908:3.04354513152249:protein_coding::chr4:5051479-5500989(+) (449510 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000152953:ENST00000282908:3.04354513152249:protein_coding::chr4:5051479-5500989(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:369734 to 369762 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000152953:ENST00000282908:3.04354513152249:protein_coding::chr4:5051479-5500989(+) 163.00 -31.81 2 22 369734 369762
26 69.23% 76.92%


   Forward: Score: 163.000000  Q:2 to 22  R:424539 to 424567 Align Len (26) (69.23%) (76.92%)
```

```
   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000152953:ENST00000282908:3.04354513152249:protein_coding::chr4:5051479-5500989(+) 163.00 -31.81 2 22 424539 424567
26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:339604 to 339631 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' agAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -32.480000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000152953:ENST00000282908:3.04354513152249:protein_coding::chr4:5051479-5500989(+) 160.00 -32.48 2 22 339604 339631
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:405861 to 405888 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000152953:ENST00000282908:3.04354513152249:protein_coding::chr4:5051479-5500989(+) 160.00 -31.53 2 22 405861 405888
25 64.00% 76.00%
--
Read Sequence:ENSG00000169436:ENST00000303045:3.03826746778517:protein_coding::chr8:138588234-138914041(-) (325807 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000169436:ENST00000303045:3.03826746778517:protein_coding::chr8:138588234-138914041(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 173.000000  Q:2 to 22  R:136830 to 136852 Align Len (20) (70.00%) (90.00%)

   Query:    3' uuUUCUCGACCUUUCGUGGGCGc 5'
                  :||: ||||: :||||||||
   Ref:      5' agGAGGCCUGGGUGGCACCCGCu 3'

   Energy:  -32.060001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000169436:ENST00000303045:3.03826746778517:protein_coding::chr8:138588234-138914041(-) 173.00 -32.06 2 22
136830 136852 20 70.00% 90.00%
--
Read Sequence:ENSG00000221866:ENST00000321063:2.95239727689986:protein_coding::chr7:132123339-132576470(-) (453131 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000221866:ENST00000321063:2.95239727689986:protein_coding::chr7:132123339-132576470(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:210054 to 210081 Align Len (25) (64.00%) (76.00%)
```

```
   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGGUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000221866:ENST00000321063:2.95239727689986:protein_coding::chr7:132123339-132576470(-) 160.00 -31.53 2 22
210054 210081 25 64.00% 76.00%
--
Read Sequence:ENSG00000273540:ENST00000614907:2.92442685303074:protein_coding::chr15:86079870-86916156(+) (836286 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000273540:ENST00000614907:2.92442685303074:protein_coding::chr15:86079870-86916156(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:240011 to 240038 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000273540:ENST00000614907:2.92442685303074:protein_coding::chr15:86079870-86916156(+) 160.00 -31.40 2 22
240011 240038 25 64.00% 76.00%
--
Read Sequence:ENSG00000125675:ENST00000620443:2.9211675766781:protein_coding::chrX:123184277-123490915(+) (306638 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000125675:ENST00000620443:2.9211675766781:protein_coding::chrX:123184277-123490915(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 171.000000  Q:2 to 22  R:88693 to 88718 Align Len (23) (78.26%) (82.61%)

   Query:    3' uuUUCUCGAC--CUUUC-GUGGGCGc 5'
                :|||||||  ||| | |||||||
   Ref:      5' uuGAGAGCUGGUGAAUGCCACCCGCa 3'

   Energy:  -30.590000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000125675:ENST00000620443:2.9211675766781:protein_coding::chrX:123184277-123490915(+) 171.00 -30.59 2 22 88693 88718
23 78.26% 82.61%
--
Read Sequence:ENSG00000171724:ENST00000302536:2.77089444930859:protein_coding::chr16:77788563-77980107(+) (191544 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000171724:ENST00000302536:2.77089444930859:protein_coding::chr16:77788563-77980107(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:53449 to 53476 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: |||||||     |:||||||||
   Ref:      5' cgAGUAGCUGGAACUACAGGCACCCGCc 3'

   Energy:  -31.360001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000171724:ENST00000302536:2.77089444930859:protein_coding::chr16:77788563-77980107(+) 164.00 -31.36 2 22 53449 53476
25 68.00% 76.00%
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:157845 to 157872 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUAUAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000171724:ENST00000302536:2.77089444930859:protein_coding::chr16:77788563-77980107(+) 160.00 -31.40 2 22
157845 157872 25 64.00% 76.00%
--
Read Sequence:ENSG00000140873:ENST00000282849:2.71465054795494:protein_coding::chr16:77282127-77435034(-) (152907 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000140873:ENST00000282849:2.71465054795494:protein_coding::chr16:77282127-77435034(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 159.000000  Q:2 to 22  R:33887 to 33914 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACC----U-UUCGUGGGCGc 5'
                |: ||||||    | |:||||||||
   Ref:      5' ugAGUAGCUGGUAUUAUAGGCACCCGCc 3'

   Energy:  -30.500000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000140873:ENST00000282849:2.71465054795494:protein_coding::chr16:77282127-77435034(-) 159.00 -30.50 2 22 33887 33914
25 68.00% 76.00%
--
Read Sequence:ENSG00000108018:ENST00000263054:2.70815166333498:protein_coding::chr10:106573662-107164706(-) (591044 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000108018:ENST00000263054:2.70815166333498:protein_coding::chr10:106573662-107164706(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:395328 to 395356 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCUU-----UCGUGGGCGc 5'
                ||| ||||||:|     :||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACUGGCACCCGCc 3'

   Energy:  -30.910000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000108018:ENST00000263054:2.70815166333498:protein_coding::chr10:106573662-107164706(-) 163.00 -30.91 2 22
395328 395356 26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:143300 to 143327 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000108018:ENST00000263054:2.70815166333498:protein_coding::chr10:106573662-107164706(-) 160.00 -31.40 2 22
143300 143327 25 64.00% 76.00%
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:437728 to 437755 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCu 3'

   Energy:  -31.350000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000108018:ENST00000263054:2.70815166333498:protein_coding::chr10:106573662-107164706(-) 160.00 -31.35 2 22
437728 437755 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:500172 to 500199 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000108018:ENST00000263054:2.70815166333498:protein_coding::chr10:106573662-107164706(-) 160.00 -31.40 2 22
500172 500199 25 64.00% 76.00%
--
Read Sequence:ENSG00000189056:ENST00000428762:2.68924223602589:protein_coding::chr7:103471788-103989658(-) (517870 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000189056:ENST00000428762:2.68924223602589:protein_coding::chr7:103471788-103989658(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:104426 to 104454 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCu 3'

   Energy:  -31.629999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000189056:ENST00000428762:2.68924223602589:protein_coding::chr7:103471788-103989658(-) 163.00 -31.63 2 22
104426 104454 26 69.23% 76.92%


   Forward: Score: 162.000000  Q:2 to 22  R:45917 to 45945 Align Len (26) (73.08%) (76.92%)

   Query:    3' uuUUC-UCGACC----U-UUCGUGGGCGc 5'
                 ||| ||||||    | |:||||||||
   Ref:      5' ccAAGUAGCUGGUAUUACAGGCACCCGCc 3'

   Energy:  -30.080000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000189056:ENST00000428762:2.68924223602589:protein_coding::chr7:103471788-103989658(-) 162.00 -30.08 2 22 45917 45945
26 73.08% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:233950 to 233977 Align Len (25) (64.00%) (76.00%)
```

```
   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000189056:ENST00000428762:2.68924223602589:protein_coding::chr7:103471788-103989658(-) 160.00 -31.53 2 22
233950 233977 25 64.00% 76.00%
--
Read Sequence:ENSG00000153823:ENST00000392055:2.67843827676608:protein_coding::chr2:229023972-229271287(-) (247315 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000153823:ENST00000392055:2.67843827676608:protein_coding::chr2:229023972-229271287(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:82655 to 82683 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000153823:ENST00000392055:2.67843827676608:protein_coding::chr2:229023972-229271287(-) 163.00 -31.81 2 22 82655 82683
26 69.23% 76.92%
--
Read Sequence:ENSG00000137142:ENST00000377694:2.58493655336517:protein_coding::chr9:38406527-38424454(-) (17927 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000137142:ENST00000377694:2.58493655336517:protein_coding::chr9:38406527-38424454(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:14079 to 14106 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000137142:ENST00000377694:2.58493655336517:protein_coding::chr9:38406527-38424454(-) 160.00 -31.40 2 22 14079 14106
25 64.00% 76.00%
--
Read Sequence:ENSG00000198216:ENST00000367573:2.52810774676787:protein_coding::chr1:181483516-181808084(+) (324568 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000198216:ENST00000367573:2.52810774676787:protein_coding::chr1:181483516-181808084(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:303728 to 303755 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000198216:ENST00000367573:2.52810774676787:protein_coding::chr1:181483516-181808084(+) 160.00 -31.40 2 22
303728 303755 25 64.00% 76.00%
```

```
--
Read Sequence:ENSG00000140538:ENST00000629765:2.46839541094856:protein_coding::chr15:87859750-88256739(-) (396989 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000140538:ENST00000629765:2.46839541094856:protein_coding::chr15:87859750-88256739(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:230576 to 230603 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000140538:ENST00000629765:2.46839541094856:protein_coding::chr15:87859750-88256739(-) 160.00 -31.53 2 22
230576 230603 25 64.00% 76.00%
--
Read Sequence:ENSG00000104967:ENST00000263257:2.46136819734812:protein_coding::chr19:45933733-45973865(-) (40132 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000104967:ENST00000263257:2.46136819734812:protein_coding::chr19:45933733-45973865(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:25373 to 25400 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000104967:ENST00000263257:2.46136819734812:protein_coding::chr19:45933733-45973865(-) 160.00 -31.40 2 22 25373 25400
25 64.00% 76.00%
--
Read Sequence:ENSG00000168546:ENST00000524240:2.41610333912622:protein_coding::chr8:21690397-21788875(-) (98478 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000168546:ENST00000524240:2.41610333912622:protein_coding::chr8:21690397-21788875(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:87589 to 87616 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000168546:ENST00000524240:2.41610333912622:protein_coding::chr8:21690397-21788875(-) 160.00 -31.40 2 22 87589 87616
25 64.00% 76.00%
--
Read Sequence:ENSG00000184261:ENST00000327876:2.39703696327759:protein_coding::chr2:47509289-47570985(-) (61696 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000184261:ENST00000327876:2.39703696327759:protein_coding::chr2:47509289-47570985(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:44687 to 44714 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
```

```
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184261:ENST00000327876:2.39703696327759:protein_coding::chr2:47509289-47570985(-) 160.00 -31.40 2 22 44687 44714
25 64.00% 76.00%
--
Read Sequence:ENSG00000153930:ENST00000653862:2.36819803460972:protein_coding::chr17:55962061-56511659(+) (549598 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000153930:ENST00000653862:2.36819803460972:protein_coding::chr17:55962061-56511659(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 156.000000  Q:2 to 22  R:119416 to 119443 Align Len (25) (60.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: :|||||:     |:||||||||
   Ref:      5' agAGUGGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.549999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000153930:ENST00000653862:2.36819803460972:protein_coding::chr17:55962061-56511659(+) 156.00 -31.55 2 22
119416 119443 25 60.00% 76.00%
--
Read Sequence:ENSG00000186094:ENST00000371839:2.24816316029702:protein_coding::chr1:48532853-50023954(-) (1491101 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000186094:ENST00000371839:2.24816316029702:protein_coding::chr1:48532853-50023954(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:497970 to 497997 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000186094:ENST00000371839:2.24816316029702:protein_coding::chr1:48532853-50023954(-) 160.00 -31.40 2 22 497970 497997
25 64.00% 76.00%


   Forward: Score: 159.000000  Q:2 to 22  R:17722 to 17750 Align Len (26) (65.38%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                ||| :|||||:     |:||||||||
   Ref:      5' ccAAGUGGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -30.879999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000186094:ENST00000371839:2.24816316029702:protein_coding::chr1:48532853-50023954(-) 159.00 -30.88 2 22 17722 17750
26 65.38% 76.92%
--
Read Sequence:ENSG00000183580:ENST00000504595:2.19514851358254:protein_coding::chr5:15500179-15939793(+) (439614 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000183580:ENST00000504595:2.19514851358254:protein_coding::chr5:15500179-15939793(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:151042 to 151069 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000183580:ENST00000504595:2.19514851358254:protein_coding::chr5:15500179-15939793(+) 160.00 -31.53 2 22 151042 151069
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:233038 to 233065 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000183580:ENST00000504595:2.19514851358254:protein_coding::chr5:15500179-15939793(+) 160.00 -31.40 2 22 233038 233065
25 64.00% 76.00%
--
Read Sequence:ENSG00000100095:ENST00000248933:2.13762757895576:protein_coding::chr22:26169461-26383596(+) (214135 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000100095:ENST00000248933:2.13762757895576:protein_coding::chr22:26169461-26383596(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 172.000000  Q:2 to 22  R:110652 to 110676 Align Len (22) (72.73%) (86.36%)

   Query:    3' uuUUCUCGACCUUUC--GUGGGCGc 5'
                  :||:|||||:| |  |||||||
   Ref:      5' guGAGGGCUGGGAUGCACACCCGCc 3'

   Energy:  -31.410000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000100095:ENST00000248933:2.13762757895576:protein_coding::chr22:26169461-26383596(+) 172.00 -31.41 2 22
110652 110676 22 72.73% 86.36%
--
Read Sequence:ENSG00000187957:ENST00000341772:2.12438962801294:protein_coding::chr2:229357628-229714555(-) (356927 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000187957:ENST00000341772:2.12438962801294:protein_coding::chr2:229357628-229714555(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:288222 to 288249 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ggAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.799999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000187957:ENST00000341772:2.12438962801294:protein_coding::chr2:229357628-229714555(-) 160.00 -31.80 2 22
288222 288249 25 64.00% 76.00%
--
```

```
Read Sequence:ENSG00000181381:ENST00000682922:2.06099340718334:protein_coding::chr4:168356734-168480492(-) (123758 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000181381:ENST00000682922:2.06099340718334:protein_coding::chr4:168356734-168480492(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:616 to 643 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000181381:ENST00000682922:2.06099340718334:protein_coding::chr4:168356734-168480492(-) 160.00 -31.40 2 22 616 643 25
64.00% 76.00%
--
Read Sequence:ENSG00000082482:ENST00000444842:2.0484139174672:protein_coding::chr1:215082736-215237090(+) (154354 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000082482:ENST00000444842:2.0484139174672:protein_coding::chr1:215082736-215237090(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:87034 to 87061 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: |||||||     |:||||||||
   Ref:      5' ugAGUAGCUGGAAUUAUAGGCACCCGCc 3'

   Energy:  -31.490000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000082482:ENST00000444842:2.0484139174672:protein_coding::chr1:215082736-215237090(+) 164.00 -31.49 2 22 87034 87061
25 68.00% 76.00%
--
Read Sequence:ENSG00000170743:ENST00000318881:2.02845921913966:protein_coding::chr11:7251903-7469043(+) (217140 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000170743:ENST00000318881:2.02845921913966:protein_coding::chr11:7251903-7469043(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:203554 to 203582 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000170743:ENST00000318881:2.02845921913966:protein_coding::chr11:7251903-7469043(+) 163.00 -31.81 2 22 203554 203582
26 69.23% 76.92%
--
Read Sequence:ENSG00000164619:ENST00000649409:2.02396626479879:protein_coding::chr7:33905538-34156427(+) (250889 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000164619:ENST00000649409:2.02396626479879:protein_coding::chr7:33905538-34156427(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:66489 to 66516 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
```

```
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000164619:ENST00000649409:2.02396626479879:protein_coding::chr7:33905538-34156427(+) 160.00 -31.40 2 22 66489 66516
25 64.00% 76.00%
--
Read Sequence:ENSG00000131620:ENST00000355303:1.97335379190681:protein_coding::chr11:70078316-70189530(+) (111214 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000131620:ENST00000355303:1.97335379190681:protein_coding::chr11:70078316-70189530(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:103991 to 104015 Align Len (22) (63.64%) (86.36%)

   Query:    3' uuUUCUCGACCUUUC--GUGGGCGc 5'
                  ::|:|||| :|:|  |||||||
   Ref:      5' caGGGGGCUGAGAGGAACACCCGCg 3'

   Energy:  -35.099998 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000131620:ENST00000355303:1.97335379190681:protein_coding::chr11:70078316-70189530(+) 164.00 -35.10 2 22
103991 104015 22 63.64% 86.36%
--
Read Sequence:ENSG00000105251:ENST00000543264:1.96833030874395:protein_coding::chr19:4279265-4290722(+) (11457 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000105251:ENST00000543264:1.96833030874395:protein_coding::chr19:4279265-4290722(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:8761 to 8788 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000105251:ENST00000543264:1.96833030874395:protein_coding::chr19:4279265-4290722(+) 160.00 -31.53 2 22 8761 8788 25
64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:10735 to 10762 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACCACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000105251:ENST00000543264:1.96833030874395:protein_coding::chr19:4279265-4290722(+) 160.00 -31.53 2 22 10735 10762 25
64.00% 76.00%
--
Read Sequence:ENSG00000157680:ENST00000614521:1.96646894781994:protein_coding::chr7:137381036-137846974(-) (465938 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000157680:ENST00000614521:1.96646894781994:protein_coding::chr7:137381036-137846974(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
```

```
   Forward: Score: 163.000000  Q:2 to 22  R:279817 to 279845 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000157680:ENST00000614521:1.96646894781994:protein_coding::chr7:137381036-137846974(-) 163.00 -31.81 2 22
279817 279845 26 69.23% 76.92%
--
Read Sequence:ENSG00000136297:ENST00000401401:1.96053041043586:protein_coding::chr7:4905997-4959187(-) (53190 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000136297:ENST00000401401:1.96053041043586:protein_coding::chr7:4905997-4959187(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:17665 to 17692 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000136297:ENST00000401401:1.96053041043586:protein_coding::chr7:4905997-4959187(-) 160.00 -31.53 2 22 17665 17692 25
64.00% 76.00%
--
Read Sequence:ENSG00000163531:ENST00000339876:1.93324869312256:protein_coding::chr1:204828651-205022822(+) (194171 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000163531:ENST00000339876:1.93324869312256:protein_coding::chr1:204828651-205022822(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:51828 to 51856 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000163531:ENST00000339876:1.93324869312256:protein_coding::chr1:204828651-205022822(+) 163.00 -31.81 2 22 51828 51856
26 69.23% 76.92%


   Forward: Score: 163.000000  Q:2 to 22  R:124770 to 124798 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000163531:ENST00000339876:1.93324869312256:protein_coding::chr1:204828651-205022822(+) 163.00 -31.81 2 22
124770 124798 26 69.23% 76.92%
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:78157 to 78184 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000163531:ENST00000339876:1.93324869312256:protein_coding::chr1:204828651-205022822(+) 160.00 -31.40 2 22 78157 78184
25 64.00% 76.00%
--
Read Sequence:ENSG00000171533:ENST00000304771:1.92508020749681:protein_coding::chr11:75586917-75669038(-) (82121 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000171533:ENST00000304771:1.92508020749681:protein_coding::chr11:75586917-75669038(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:51630 to 51657 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     ||||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAAGCACCCGCc 3'

   Energy:  -31.980000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000171533:ENST00000304771:1.92508020749681:protein_coding::chr11:75586917-75669038(-) 164.00 -31.98 2 22 51630 51657
25 68.00% 76.00%
--
Read Sequence:ENSG00000179242:ENST00000614565:1.92076954502748:protein_coding::chr20:61252260-61940617(+) (688357 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000179242:ENST00000614565:1.92076954502748:protein_coding::chr20:61252260-61940617(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 174.000000  Q:2 to 20  R:396031 to 396054 Align Len (19) (78.95%) (94.74%)

   Query:    3' uuuuCUCGACCU-UUCGUGGGCGc 5'
                    |:|:|||| |:||||||||
   Ref:      5' gcccGGGUUGGAGAGGCACCCGCa 3'

   Energy:  -32.400002 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000179242:ENST00000614565:1.92076954502748:protein_coding::chr20:61252260-61940617(+) 174.00 -32.40 2 20
396031 396054 19 78.95% 94.74%


   Forward: Score: 163.000000  Q:2 to 22  R:596389 to 596417 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  |:| |||||||     |:||||||||
   Ref:      5' ccAGGUAGCUGGAACUACAGGCACCCGCc 3'

   Energy:  -31.320000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000179242:ENST00000614565:1.92076954502748:protein_coding::chr20:61252260-61940617(+) 163.00 -31.32 2 22
596389 596417 26 69.23% 76.92%
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:489356 to 489383 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000179242:ENST00000614565:1.92076954502748:protein_coding::chr20:61252260-61940617(+) 160.00 -31.40 2 22
489356 489383 25 64.00% 76.00%


   Forward: Score: 156.000000  Q:2 to 19  R:683757 to 683782 Align Len (20) (70.00%) (85.00%)

   Query:    3' uuuucUCG-ACCUU--UCGUGGGCGc 5'
                   :|| ||||:  :||||||||
   Ref:      5' gugcuGGCAUGGAGGUGGCACCCGCg 3'

   Energy:  -31.129999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000179242:ENST00000614565:1.92076954502748:protein_coding::chr20:61252260-61940617(+) 156.00 -31.13 2 19
683757 683782 20 70.00% 85.00%
--
Read Sequence:ENSG00000178031:ENST00000380548:1.89287663944332:protein_coding::chr9:18474152-18910950(+) (436798 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000178031:ENST00000380548:1.89287663944332:protein_coding::chr9:18474152-18910950(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:40302 to 40329 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000178031:ENST00000380548:1.89287663944332:protein_coding::chr9:18474152-18910950(+) 160.00 -31.40 2 22 40302 40329
25 64.00% 76.00%
--
Read Sequence:ENSG00000102359:ENST00000373004:1.88960894435827:protein_coding::chrX:100644198-100675788(+) (31590 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000102359:ENST00000373004:1.88960894435827:protein_coding::chrX:100644198-100675788(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:12854 to 12881 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' ccAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -30.780001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000102359:ENST00000373004:1.88960894435827:protein_coding::chrX:100644198-100675788(+) 160.00 -30.78 2 22 12854 12881
25 64.00% 76.00%
--
Read Sequence:ENSG00000187775:ENST00000389840:1.87299171950429:protein_coding::chr17:78423696-78577396(-) (153700 nt)
```

```
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000187775:ENST00000389840:1.87299171950429:protein_coding::chr17:78423696-78577396(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:40326 to 40353 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000187775:ENST00000389840:1.87299171950429:protein_coding::chr17:78423696-78577396(-) 160.00 -31.40 2 22 40326 40353
25 64.00% 76.00%
--
Read Sequence:ENSG00000171811:ENST00000368586:1.81983816568735:protein_coding::chr10:132808391-132942570(-) (134179 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000171811:ENST00000368586:1.81983816568735:protein_coding::chr10:132808391-132942570(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 183.000000  Q:2 to 22  R:9748 to 9769 Align Len (20) (85.00%) (95.00%)

   Query:    3' uuUUCUCGACCUUUCGUGGGCGc 5'
                :|||||| ||:|||||||||
   Ref:      5' ugGAGAGCU-GAGAGCACCCGCu 3'

   Energy:  -35.099998 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000171811:ENST00000368586:1.81983816568735:protein_coding::chr10:132808391-132942570(-) 183.00 -35.10 2 22 9748 9769
20 85.00% 95.00%
--
Read Sequence:ENSG00000187398:ENST00000336930:1.81649050972898:protein_coding::chr11:24497052-25082638(+) (585586 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000187398:ENST00000336930:1.81649050972898:protein_coding::chr11:24497052-25082638(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:425921 to 425948 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000187398:ENST00000336930:1.81649050972898:protein_coding::chr11:24497052-25082638(+) 160.00 -31.53 2 22
425921 425948 25 64.00% 76.00%
--
Read Sequence:ENSG00000175262:ENST00000377004:1.81242764615932:protein_coding::chr1:10946472-10982037(-) (35565 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000175262:ENST00000377004:1.81242764615932:protein_coding::chr1:10946472-10982037(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:26300 to 26328 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'
```

```
   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000175262:ENST00000377004:1.81242764615932:protein_coding::chr1:10946472-10982037(-) 163.00 -31.81 2 22 26300 26328
26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:8772 to 8799 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCa 3'

   Energy:  -31.879999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000175262:ENST00000377004:1.81242764615932:protein_coding::chr1:10946472-10982037(-) 160.00 -31.88 2 22 8772 8799 25
64.00% 76.00%
--
Read Sequence:ENSG00000103489:ENST00000261381:1.80410703119625:protein_coding::chr16:17101768-17470960(-) (369192 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000103489:ENST00000261381:1.80410703119625:protein_coding::chr16:17101768-17470960(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:135843 to 135870 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACC-----UUUCGUGGGCGc 5'
                  |: ||||||     :|:||||||||
   Ref:      5' ugAGUAGCUGGGAUUAGAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000103489:ENST00000261381:1.80410703119625:protein_coding::chr16:17101768-17470960(-) 160.00 -31.53 2 22
135843 135870 25 64.00% 76.00%
--
Read Sequence:ENSG00000266714:ENST00000645453:1.79486000582748:protein_coding::chr17:75587799-75626849(+) (39050 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000266714:ENST00000645453:1.79486000582748:protein_coding::chr17:75587799-75626849(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:11599 to 11627 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000266714:ENST00000645453:1.79486000582748:protein_coding::chr17:75587799-75626849(+) 163.00 -31.81 2 22 11599 11627
26 69.23% 76.92%
--
Read Sequence:ENSG00000172260:ENST00000357731:1.78731154438963:protein_coding::chr1:71395942-72282539(-) (886597 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000172260:ENST00000357731:1.78731154438963:protein_coding::chr1:71395942-72282539(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:364430 to 364457 Align Len (25) (64.00%) (76.00%)
```

```
   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000172260:ENST00000357731:1.78731154438963:protein_coding::chr1:71395942-72282539(-) 160.00 -31.40 2 22 364430 364457
25 64.00% 76.00%
--
Read Sequence:ENSG00000197721:ENST00000508064:1.7763213954585:protein_coding::chr1:207645132-207723703(+) (78571 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000197721:ENST00000508064:1.7763213954585:protein_coding::chr1:207645132-207723703(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:20046 to 20073 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGGCUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000197721:ENST00000508064:1.7763213954585:protein_coding::chr1:207645132-207723703(+) 160.00 -31.40 2 22 20046 20073
25 64.00% 76.00%
--
Read Sequence:ENSG00000075461:ENST00000262138:1.73930554764103:protein_coding::chr17:66964706-67033398(+) (68692 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000075461:ENST00000262138:1.73930554764103:protein_coding::chr17:66964706-67033398(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:29023 to 29051 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000075461:ENST00000262138:1.73930554764103:protein_coding::chr17:66964706-67033398(+) 163.00 -31.81 2 22 29023 29051
26 69.23% 76.92%
--
Read Sequence:ENSG00000151702:ENST00000527786:1.73900178943758:protein_coding::chr11:128694071-128813267(+) (119196 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000151702:ENST00000527786:1.73900178943758:protein_coding::chr11:128694071-128813267(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:57579 to 57607 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000151702:ENST00000527786:1.73900178943758:protein_coding::chr11:128694071-128813267(+) 163.00 -31.81 2 22
```

```
57579 57607 26 69.23% 76.92%
--
Read Sequence:ENSG00000058335:ENST00000558480:1.73289122955772:protein_coding::chr15:78959905-79090780(-) (130875 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000058335:ENST00000558480:1.73289122955772:protein_coding::chr15:78959905-79090780(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:68448 to 68475 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: |||||||     |:||||||||
   Ref:      5' cgAGUAGCUGGAAUUAUAGGCACCCGCu 3'

   Energy:  -31.180000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000058335:ENST00000558480:1.73289122955772:protein_coding::chr15:78959905-79090780(-) 164.00 -31.18 2 22 68448 68475
25 68.00% 76.00%
--
Read Sequence:ENSG00000206384:ENST00000358511:1.73046153380702:protein_coding::chr3:130517176-130677042(+) (159866 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000206384:ENST00000358511:1.73046153380702:protein_coding::chr3:130517176-130677042(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:73516 to 73543 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACCACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000206384:ENST00000358511:1.73046153380702:protein_coding::chr3:130517176-130677042(+) 160.00 -31.53 2 22 73516 73543
25 64.00% 76.00%
--
Read Sequence:ENSG00000184347:ENST00000519560:1.73004918346393:protein_coding::chr5:168661739-169301139(-) (639400 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000184347:ENST00000519560:1.73004918346393:protein_coding::chr5:168661739-169301139(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:45364 to 45391 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACAACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184347:ENST00000519560:1.73004918346393:protein_coding::chr5:168661739-169301139(-) 160.00 -31.53 2 22 45364 45391
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:159508 to 159535 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' agAGUAGCUGGGACUACAGGCACCCGCc 3'
```

```
   Energy:  -32.480000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184347:ENST00000519560:1.73004918346393:protein_coding::chr5:168661739-169301139(-) 160.00 -32.48 2 22
159508 159535 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:344018 to 344045 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' agAGUAGCUGGGAUUAUAGGCACCCGCc 3'

   Energy:  -32.480000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184347:ENST00000519560:1.73004918346393:protein_coding::chr5:168661739-169301139(-) 160.00 -32.48 2 22
344018 344045 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:378790 to 378817 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184347:ENST00000519560:1.73004918346393:protein_coding::chr5:168661739-169301139(-) 160.00 -31.40 2 22
378790 378817 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:381987 to 382014 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184347:ENST00000519560:1.73004918346393:protein_coding::chr5:168661739-169301139(-) 160.00 -31.40 2 22
381987 382014 25 64.00% 76.00%
--
Read Sequence:ENSG00000154493:ENST00000488181:1.71732798373038:protein_coding::chr10:126424996-126670693(-) (245697 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000154493:ENST00000488181:1.71732798373038:protein_coding::chr10:126424996-126670693(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:3821 to 3848 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     ||||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAAGCACCCGCc 3'

   Energy:  -31.980000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000154493:ENST00000488181:1.71732798373038:protein_coding::chr10:126424996-126670693(-) 164.00 -31.98 2 22 3821 3848
```

```
25 68.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:137964 to 137991 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' ggAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.799999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000154493:ENST00000488181:1.71732798373038:protein_coding::chr10:126424996-126670693(-) 160.00 -31.80 2 22
137964 137991 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:139627 to 139654 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000154493:ENST00000488181:1.71732798373038:protein_coding::chr10:126424996-126670693(-) 160.00 -31.53 2 22
139627 139654 25 64.00% 76.00%
--
Read Sequence:ENSG00000105143:ENST00000594383:1.7126198948476:protein_coding::chr19:14950032-14979864(-) (29832 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000105143:ENST00000594383:1.7126198948476:protein_coding::chr19:14950032-14979864(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 164.000000  Q:2 to 22  R:9261 to 9288 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: |||||||     |:||||||||
   Ref:      5' ugAGUAGCUGGAACUACAGGCACCCGCc 3'

   Energy:  -31.490000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000105143:ENST00000594383:1.7126198948476:protein_coding::chr19:14950032-14979864(-) 164.00 -31.49 2 22 9261 9288 25
68.00% 76.00%
--
Read Sequence:ENSG00000159307:ENST00000360835:1.69706868227251:protein_coding::chr22:43197279-43343372(-) (146093 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000159307:ENST00000360835:1.69706868227251:protein_coding::chr22:43197279-43343372(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 163.000000  Q:2 to 22  R:44427 to 44455 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGCAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000159307:ENST00000360835:1.69706868227251:protein_coding::chr22:43197279-43343372(-) 163.00 -31.81 2 22 44427 44455
```

```
26 69.23% 76.92%
--
Read Sequence:ENSG00000132854:ENST00000371153:1.69402109407203:protein_coding::chr1:62236164-62319434(-) (83270 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000132854:ENST00000371153:1.69402109407203:protein_coding::chr1:62236164-62319434(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:67549 to 67577 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000132854:ENST00000371153:1.69402109407203:protein_coding::chr1:62236164-62319434(-) 163.00 -31.81 2 22 67549 67577
26 69.23% 76.92%
--
Read Sequence:ENSG00000169181:ENST00000447459:1.68716121909993:protein_coding::chr16:27787527-28063714(-) (276187 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000169181:ENST00000447459:1.68716121909993:protein_coding::chr16:27787527-28063714(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:140079 to 140107 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACCACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000169181:ENST00000447459:1.68716121909993:protein_coding::chr16:27787527-28063714(-) 163.00 -31.81 2 22
140079 140107 26 69.23% 76.92%
--
Read Sequence:ENSG00000126785:ENST00000316754:1.68649500042142:protein_coding::chr14:63204442-63293508(+) (89066 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000126785:ENST00000316754:1.68649500042142:protein_coding::chr14:63204442-63293508(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:39022 to 39049 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000126785:ENST00000316754:1.68649500042142:protein_coding::chr14:63204442-63293508(+) 160.00 -31.40 2 22 39022 39049
25 64.00% 76.00%
--
Read Sequence:ENSG00000058866:ENST00000265022:1.65988135861068:protein_coding::chr3:186147200-186362234(-) (215034 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000058866:ENST00000265022:1.65988135861068:protein_coding::chr3:186147200-186362234(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 157.000000  Q:2 to 19  R:69548 to 69575 Align Len (22) (68.18%) (77.27%)
```

```
   Query:    3' uuuucUCGACCU-----UUCGUGGGCGc 5'
                     ||||||:     |:||||||||
   Ref:      5' uggguAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -30.700001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000058866:ENST00000265022:1.65988135861068:protein_coding::chr3:186147200-186362234(-) 157.00 -30.70 2 19 69548 69575
22 68.18% 77.27%
--
Read Sequence:ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+) (1694317 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:1179106 to 1179134 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+) 163.00 -31.81 2 22
1179106 1179134 26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:161667 to 161694 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+) 160.00 -31.53 2 22 161667 161694
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:657777 to 657804 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cuAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -30.910000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+) 160.00 -30.91 2 22 657777 657804
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:1045270 to 1045297 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGCCUACAGGCACCCGCc 3'
```

```
   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+) 160.00 -31.53 2 22
1045270 1045297 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:1221068 to 1221095 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCUU-----UCGUGGGCGc 5'
                  |: ||||||:|     :||||||||
   Ref:      5' ugAGUAGCUGGGAUUACGGGCACCCGCc 3'

   Energy:  -30.770000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+) 160.00 -30.77 2 22
1221068 1221095 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:1500987 to 1501014 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078328:ENST00000550418:1.65241325044756:protein_coding::chr16:6019023-7713340(+) 160.00 -31.40 2 22
1500987 1501014 25 64.00% 76.00%
--
Read Sequence:ENSG00000110799:ENST00000261405:1.63518325259963:protein_coding::chr12:5948876-6124670(-) (175794 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000110799:ENST00000261405:1.63518325259963:protein_coding::chr12:5948876-6124670(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:15797 to 15824 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000110799:ENST00000261405:1.63518325259963:protein_coding::chr12:5948876-6124670(-) 160.00 -31.40 2 22 15797 15824 25
64.00% 76.00%
--
Read Sequence:ENSG00000183454:ENST00000330684:1.62560684600351:protein_coding::chr16:9753407-10182428(-) (429021 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000183454:ENST00000330684:1.62560684600351:protein_coding::chr16:9753407-10182428(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:195770 to 195798 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'
```

```
   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000183454:ENST00000330684:1.62560684600351:protein_coding::chr16:9753407-10182428(-) 163.00 -31.81 2 22 195770 195798
26 69.23% 76.92%
--
Read Sequence:ENSG00000147481:ENST00000642720:1.60617951174616:protein_coding::chr8:49911406-50796692(+) (885286 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000147481:ENST00000642720:1.60617951174616:protein_coding::chr8:49911406-50796692(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:55113 to 55140 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cuAGUAGCUGGGACUACAGGCACCCGCu 3'

   Energy:  -30.730000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000147481:ENST00000642720:1.60617951174616:protein_coding::chr8:49911406-50796692(+) 160.00 -30.73 2 22 55113 55140
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:82974 to 83001 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' auAGUAGCUGGGGUUACAGGCACCCGCc 3'

   Energy:  -30.910000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000147481:ENST00000642720:1.60617951174616:protein_coding::chr8:49911406-50796692(+) 160.00 -30.91 2 22 82974 83001
25 64.00% 76.00%
--
Read Sequence:ENSG00000078295:ENST00000338316:1.59397669379068:protein_coding::chr5:7396137-7830081(+) (433944 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000078295:ENST00000338316:1.59397669379068:protein_coding::chr5:7396137-7830081(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:110775 to 110802 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCUU-----UCGUGGGCGc 5'
                  |: ||||||:|     :||||||||
   Ref:      5' ugAGUAGCUGGGACUACGGGCACCCGCc 3'

   Energy:  -30.770000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078295:ENST00000338316:1.59397669379068:protein_coding::chr5:7396137-7830081(+) 160.00 -30.77 2 22 110775 110802
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:269826 to 269853 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'
```

```
   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078295:ENST00000338316:1.59397669379068:protein_coding::chr5:7396137-7830081(+) 160.00 -31.53 2 22 269826 269853
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:356886 to 356913 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' uuAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -30.910000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000078295:ENST00000338316:1.59397669379068:protein_coding::chr5:7396137-7830081(+) 160.00 -30.91 2 22 356886 356913
25 64.00% 76.00%
--
Read Sequence:ENSG00000178342:ENST00000316249:1.57990458805582:protein_coding::chr18:79797937-79900100(+) (102163 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000178342:ENST00000316249:1.57990458805582:protein_coding::chr18:79797937-79900100(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 172.000000  Q:2 to 22  R:90234 to 90257 Align Len (21) (71.43%) (90.48%)

   Query:    3' uuUUCUCG-ACCUUUCGUGGGCGc 5'
                  :||: | |||::|||||||||
   Ref:      5' agGAGGCCAUGGGGAGCACCCGCc 3'

   Energy:  -32.740002 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000178342:ENST00000316249:1.57990458805582:protein_coding::chr18:79797937-79900100(+) 172.00 -32.74 2 22 90234 90257
21 71.43% 90.48%
--
Read Sequence:ENSG00000075035:ENST00000547525:1.57864180322127:protein_coding::chr12:108129287-108250537(+) (121250 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000075035:ENST00000547525:1.57864180322127:protein_coding::chr12:108129287-108250537(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:46676 to 46704 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000075035:ENST00000547525:1.57864180322127:protein_coding::chr12:108129287-108250537(+) 163.00 -31.81 2 22
46676 46704 26 69.23% 76.92%
--
Read Sequence:ENSG00000225968:ENST00000424383:1.54678369840882:protein_coding::chr7:1670276-1747946(+) (77670 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000225968:ENST00000424383:1.54678369840882:protein_coding::chr7:1670276-1747946(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:41970 to 41997 Align Len (25) (64.00%) (76.00%)
```

```
   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000225968:ENST00000424383:1.54678369840882:protein_coding::chr7:1670276-1747946(+) 160.00 -31.53 2 22 41970 41997 25
64.00% 76.00%
--
Read Sequence:ENSG00000180336:ENST00000409122:1.5180563999524:protein_coding::chr17:44656467-44675801(+) (19334 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000180336:ENST00000409122:1.5180563999524:protein_coding::chr17:44656467-44675801(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:1812 to 1839 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000180336:ENST00000409122:1.5180563999524:protein_coding::chr17:44656467-44675801(+) 160.00 -31.40 2 22 1812 1839 25
64.00% 76.00%
--
Read Sequence:ENSG00000112562:ENST00000356284:1.51541882024909:protein_coding::chr6:168441183-168667992(+) (226809 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000112562:ENST00000356284:1.51541882024909:protein_coding::chr6:168441183-168667992(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:93912 to 93939 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ggAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.799999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000112562:ENST00000356284:1.51541882024909:protein_coding::chr6:168441183-168667992(+) 160.00 -31.80 2 22 93912 93939
25 64.00% 76.00%
--
Read Sequence:ENSG00000176463:ENST00000318445:1.51474313675348:protein_coding::chr15:91853707-92165903(+) (312196 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000176463:ENST00000318445:1.51474313675348:protein_coding::chr15:91853707-92165903(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:53623 to 53650 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000176463:ENST00000318445:1.51474313675348:protein_coding::chr15:91853707-92165903(+) 160.00 -31.40 2 22 53623 53650
25 64.00% 76.00%
```

```
--
Read Sequence:ENSG00000088002:ENST00000201586:1.50640931993015:protein_coding::chr19:48552171-48599423(+) (47252 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000088002:ENST00000201586:1.50640931993015:protein_coding::chr19:48552171-48599423(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 163.000000  Q:2 to 22  R:3044 to 3072 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000088002:ENST00000201586:1.50640931993015:protein_coding::chr19:48552171-48599423(+) 163.00 -31.81 2 22 3044 3072 26
69.23% 76.92%
--
Read Sequence:ENSG00000145934:ENST00000518659:1.49756217308624:protein_coding::chr5:166979028-168264157(+) (1285129 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000145934:ENST00000518659:1.49756217308624:protein_coding::chr5:166979028-168264157(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 160.000000  Q:2 to 22  R:172615 to 172642 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000145934:ENST00000518659:1.49756217308624:protein_coding::chr5:166979028-168264157(+) 160.00 -31.53 2 22
172615 172642 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:1217558 to 1217585 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000145934:ENST00000518659:1.49756217308624:protein_coding::chr5:166979028-168264157(+) 160.00 -31.40 2 22
1217558 1217585 25 64.00% 76.00%


   Forward: Score: 159.000000  Q:2 to 22  R:1278711 to 1278739 Align Len (26) (65.38%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 |:| ||||||:     |:||||||||
   Ref:      5' ccAGGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.360001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000145934:ENST00000518659:1.49756217308624:protein_coding::chr5:166979028-168264157(+) 159.00 -31.36 2 22
1278711 1278739 26 65.38% 76.92%
```

```
--
Read Sequence:ENSG00000185736:ENST00000381312:1.49390736653951:protein_coding::chr10:1177312-1737525(-) (560213 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000185736:ENST00000381312:1.49390736653951:protein_coding::chr10:1177312-1737525(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:544689 to 544717 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000185736:ENST00000381312:1.49390736653951:protein_coding::chr10:1177312-1737525(-) 163.00 -31.81 2 22 544689 544717
26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:255776 to 255803 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000185736:ENST00000381312:1.49390736653951:protein_coding::chr10:1177312-1737525(-) 160.00 -31.53 2 22 255776 255803
25 64.00% 76.00%


   Forward: Score: 158.000000  Q:2 to 22  R:46513 to 46533 Align Len (20) (65.00%) (85.00%)

   Query:    3' uuUUCUCGACCUUUCGUGGGCGc 5'
                  |:|:|:|||  : |||||||
   Ref:      5' agAGGGGUUGG--GUCACCCGCa 3'

   Energy:  -30.700001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000185736:ENST00000381312:1.49390736653951:protein_coding::chr10:1177312-1737525(-) 158.00 -30.70 2 22 46513 46533 20
65.00% 85.00%
--
Read Sequence:ENSG00000141293:ENST00000336915:1.48055107494198:protein_coding::chr17:48133441-48430214(-) (296773 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000141293:ENST00000336915:1.48055107494198:protein_coding::chr17:48133441-48430214(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:271754 to 271781 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000141293:ENST00000336915:1.48055107494198:protein_coding::chr17:48133441-48430214(-) 160.00 -31.40 2 22
271754 271781 25 64.00% 76.00%
```

```
--
Read Sequence:ENSG00000170615:ENST00000306312:1.46681047430855:protein_coding::chr7:103374206-103446207(-) (72001 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000170615:ENST00000306312:1.46681047430855:protein_coding::chr7:103374206-103446207(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:48774 to 48801 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGCAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000170615:ENST00000306312:1.46681047430855:protein_coding::chr7:103374206-103446207(-) 160.00 -31.40 2 22 48774 48801
25 64.00% 76.00%
--
Read Sequence:ENSG00000185052:ENST00000328041:1.45817247761174:protein_coding::chr20:19212641-19722926(+) (510285 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000185052:ENST00000328041:1.45817247761174:protein_coding::chr20:19212641-19722926(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:100520 to 100547 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUAUAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000185052:ENST00000328041:1.45817247761174:protein_coding::chr20:19212641-19722926(+) 160.00 -31.40 2 22
100520 100547 25 64.00% 76.00%
--
Read Sequence:ENSG00000158445:ENST00000371741:1.44997010117592:protein_coding::chr20:49363876-49482668(-) (118792 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000158445:ENST00000371741:1.44997010117592:protein_coding::chr20:49363876-49482668(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 159.000000  Q:2 to 22  R:13071 to 13099 Align Len (26) (65.38%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  |:| ||||||:     |:||||||||
   Ref:      5' ccAGGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.360001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000158445:ENST00000371741:1.44997010117592:protein_coding::chr20:49363876-49482668(-) 159.00 -31.36 2 22 13071 13099
26 65.38% 76.92%
--
Read Sequence:ENSG00000183018:ENST00000329078:1.44395394164594:protein_coding::chr17:4498880-4539035(+) (40155 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000183018:ENST00000329078:1.44395394164594:protein_coding::chr17:4498880-4539035(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:30180 to 30207 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
```

```
                |: ||||||:     |:||||||||
   Ref:      5' ccAGUAGCUGGGAUCACAGGCACCCGCc 3'

   Energy:  -30.780001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000183018:ENST00000329078:1.44395394164594:protein_coding::chr17:4498880-4539035(+) 160.00 -30.78 2 22 30180 30207 25
64.00% 76.00%
--
Read Sequence:ENSG00000144229:ENST00000409968:1.44249836284832:protein_coding::chr2:136765544-137677718(+) (912174 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000144229:ENST00000409968:1.44249836284832:protein_coding::chr2:136765544-137677718(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:498833 to 498860 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000144229:ENST00000409968:1.44249836284832:protein_coding::chr2:136765544-137677718(+) 160.00 -31.53 2 22
498833 498860 25 64.00% 76.00%
--
Read Sequence:ENSG00000170962:ENST00000393158:1.44164519536397:protein_coding::chr11:103907188-104164147(-) (256959 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000170962:ENST00000393158:1.44164519536397:protein_coding::chr11:103907188-104164147(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:173116 to 173143 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000170962:ENST00000393158:1.44164519536397:protein_coding::chr11:103907188-104164147(-) 160.00 -31.53 2 22
173116 173143 25 64.00% 76.00%
--
Read Sequence:ENSG00000206579:ENST00000327381:1.43994564621106:protein_coding::chr8:55102027-55542054(+) (440027 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000206579:ENST00000327381:1.43994564621106:protein_coding::chr8:55102027-55542054(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:95640 to 95668 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000206579:ENST00000327381:1.43994564621106:protein_coding::chr8:55102027-55542054(+) 163.00 -31.81 2 22 95640 95668
26 69.23% 76.92%
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:78629 to 78656 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGCAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000206579:ENST00000327381:1.43994564621106:protein_coding::chr8:55102027-55542054(+) 160.00 -31.40 2 22 78629 78656
25 64.00% 76.00%
--
Read Sequence:ENSG00000178568:ENST00000342788:1.43219331326049:protein_coding::chr2:211375716-212538802(-) (1163086 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000178568:ENST00000342788:1.43219331326049:protein_coding::chr2:211375716-212538802(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:411318 to 411345 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000178568:ENST00000342788:1.43219331326049:protein_coding::chr2:211375716-212538802(-) 160.00 -31.40 2 22
411318 411345 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:523644 to 523671 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000178568:ENST00000342788:1.43219331326049:protein_coding::chr2:211375716-212538802(-) 160.00 -31.40 2 22
523644 523671 25 64.00% 76.00%
--
Read Sequence:ENSG00000174059:ENST00000310833:1.42872519355218:protein_coding::chr1:207880971-207911125(-) (30154 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000174059:ENST00000310833:1.42872519355218:protein_coding::chr1:207880971-207911125(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:29572 to 29600 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000174059:ENST00000310833:1.42872519355218:protein_coding::chr1:207880971-207911125(-) 163.00 -31.81 2 22 29572 29600
26 69.23% 76.92%
--
```

```
Read Sequence:ENSG00000119681:ENST00000261978:1.41532215914513:protein_coding::chr14:74498182-74612237(-) (114055 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000119681:ENST00000261978:1.41532215914513:protein_coding::chr14:74498182-74612237(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 167.000000  Q:2 to 22  R:71579 to 71607 Align Len (26) (73.08%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| |||||||     |:||||||||
   Ref:      5' ccAAGUAGCUGGAACUACAGGCACCCGCc 3'

   Energy:  -31.770000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000119681:ENST00000261978:1.41532215914513:protein_coding::chr14:74498182-74612237(-) 167.00 -31.77 2 22 71579 71607
26 73.08% 76.92%
--
Read Sequence:ENSG00000168589:ENST00000305904:1.41005520777854:protein_coding::chr16:80540992-80550811(+) (9819 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000168589:ENST00000305904:1.41005520777854:protein_coding::chr16:80540992-80550811(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:4807 to 4835 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000168589:ENST00000305904:1.41005520777854:protein_coding::chr16:80540992-80550811(+) 163.00 -31.81 2 22 4807 4835 26
69.23% 76.92%
--
Read Sequence:ENSG00000156687:ENST00000404895:1.39301602528676:protein_coding::chr8:35235474-35796540(+) (561066 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000156687:ENST00000404895:1.39301602528676:protein_coding::chr8:35235474-35796540(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:499434 to 499461 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000156687:ENST00000404895:1.39301602528676:protein_coding::chr8:35235474-35796540(+) 160.00 -31.40 2 22 499434 499461
25 64.00% 76.00%
--
Read Sequence:ENSG00000148655:ENST00000611255:1.37031744891008:protein_coding::chr10:75431623-76560168(+) (1128545 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000148655:ENST00000611255:1.37031744891008:protein_coding::chr10:75431623-76560168(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:887411 to 887438 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  || ||||||:     |:||||||||
```

```
   Ref:      5' agAAUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -30.459999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000148655:ENST00000611255:1.37031744891008:protein_coding::chr10:75431623-76560168(+) 164.00 -30.46 2 22
887411 887438 25 68.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:435664 to 435691 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000148655:ENST00000611255:1.37031744891008:protein_coding::chr10:75431623-76560168(+) 160.00 -31.53 2 22
435664 435691 25 64.00% 76.00%
--
Read Sequence:ENSG00000196639:ENST00000431010:1.35666257301116:protein_coding::chr3:11154492-11263557(+) (109065 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000196639:ENST00000431010:1.35666257301116:protein_coding::chr3:11154492-11263557(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:19034 to 19061 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000196639:ENST00000431010:1.35666257301116:protein_coding::chr3:11154492-11263557(+) 160.00 -31.40 2 22 19034 19061
25 64.00% 76.00%
--
Read Sequence:ENSG00000181291:ENST00000631683:1.34170114217592:protein_coding::chr17:34579581-34639318(+) (59737 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000181291:ENST00000631683:1.34170114217592:protein_coding::chr17:34579581-34639318(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 156.000000  Q:2 to 22  R:2138 to 2164 Align Len (24) (58.33%) (79.17%)

   Query:    3' uuUUCUCGACCU----UUCGUGGGCGc 5'
                  :||:|| ||:    ::||||||||
   Ref:      5' cgGAGGGCGGGGGUCUGGGCACCCGCa 3'

   Energy:  -30.379999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000181291:ENST00000631683:1.34170114217592:protein_coding::chr17:34579581-34639318(+) 156.00 -30.38 2 22 2138 2164 24
58.33% 79.17%
--
Read Sequence:ENSG00000221818:ENST00000520164:1.31600886382693:protein_coding::chr8:25841724-26045413(-) (203689 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000221818:ENST00000520164:1.31600886382693:protein_coding::chr8:25841724-26045413(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:179496 to 179523 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000221818:ENST00000520164:1.31600886382693:protein_coding::chr8:25841724-26045413(-) 160.00 -31.40 2 22 179496 179523
25 64.00% 76.00%
--
Read Sequence:ENSG00000166145:ENST00000562057:1.31023122558137:protein_coding::chr15:40844047-40858207(+) (14160 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000166145:ENST00000562057:1.31023122558137:protein_coding::chr15:40844047-40858207(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:1229 to 1256 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: |||||||     |:||||||||
   Ref:      5' ggAGUAGCUGGAACUACAGGCACCCGCu 3'

   Energy:  -31.580000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000166145:ENST00000562057:1.31023122558137:protein_coding::chr15:40844047-40858207(+) 164.00 -31.58 2 22 1229 1256 25
68.00% 76.00%
--
Read Sequence:ENSG00000132718:ENST00000368324:1.27619159825752:protein_coding::chr1:155859566-155885199(+) (25633 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000132718:ENST00000368324:1.27619159825752:protein_coding::chr1:155859566-155885199(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:17525 to 17553 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' cuAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.940001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000132718:ENST00000368324:1.27619159825752:protein_coding::chr1:155859566-155885199(+) 163.00 -31.94 2 22 17525 17553
26 69.23% 76.92%
--
Read Sequence:ENSG00000184384:ENST00000524717:1.27161990173257:protein_coding::chr11:95976597-96343195(-) (366598 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000184384:ENST00000524717:1.27161990173257:protein_coding::chr11:95976597-96343195(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:241026 to 241054 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
```

```
>vtRNA1-2 ENSG00000184384:ENST00000524717:1.27161990173257:protein_coding::chr11:95976597-96343195(-) 163.00 -31.81 2 22
241026 241054 26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:303345 to 303372 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184384:ENST00000524717:1.27161990173257:protein_coding::chr11:95976597-96343195(-) 160.00 -31.53 2 22
303345 303372 25 64.00% 76.00%
--
Read Sequence:ENSG00000110076:ENST00000265459:1.27003453902695:protein_coding::chr11:64606173-64723197(-) (117024 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000110076:ENST00000265459:1.27003453902695:protein_coding::chr11:64606173-64723197(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:11583 to 11606 Align Len (21) (71.43%) (80.95%)

   Query:    3' uuUUCUC-GACCUUUCGUGGGCGc 5'
                |:|:| | ||  |||||||||
   Ref:      5' ggAGGGGUCCGGCCAGCACCCGCg 3'

   Energy:  -32.540001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000110076:ENST00000265459:1.27003453902695:protein_coding::chr11:64606173-64723197(-) 164.00 -32.54 2 22 11583 11606
21 71.43% 80.95%


   Forward: Score: 160.000000  Q:2 to 22  R:64280 to 64307 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000110076:ENST00000265459:1.27003453902695:protein_coding::chr11:64606173-64723197(-) 160.00 -31.40 2 22 64280 64307
25 64.00% 76.00%
--
Read Sequence:ENSG00000107736:ENST00000224721:1.26646037481865:protein_coding::chr10:71396919-71815947(+) (419028 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000107736:ENST00000224721:1.26646037481865:protein_coding::chr10:71396919-71815947(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:409781 to 409809 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
```

```
>vtRNA1-2 ENSG00000107736:ENST00000224721:1.26646037481865:protein_coding::chr10:71396919-71815947(+) 163.00 -31.81 2 22
409781 409809 26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:371376 to 371403 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000107736:ENST00000224721:1.26646037481865:protein_coding::chr10:71396919-71815947(+) 160.00 -31.40 2 22
371376 371403 25 64.00% 76.00%


   Forward: Score: 154.000000  Q:2 to 20  R:340841 to 340865 Align Len (20) (65.00%) (80.00%)

   Query:    3' uuuuCUCGACCUUU--CGUGGGCGc 5'
                    |: |||||::   |||||||
   Ref:      5' ccguGGACUGGAGGCCUCACCCGCg 3'

   Energy:  -30.540001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000107736:ENST00000224721:1.26646037481865:protein_coding::chr10:71396919-71815947(+) 154.00 -30.54 2 20
340841 340865 20 65.00% 80.00%
--
Read Sequence:ENSG00000186732:ENST00000443721:1.26406798343863:protein_coding::chr22:43412013-43507848(+) (95835 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000186732:ENST00000443721:1.26406798343863:protein_coding::chr22:43412013-43507848(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 160.000000  Q:2 to 22  R:41061 to 41088 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCu 3'

   Energy:  -31.219999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000186732:ENST00000443721:1.26406798343863:protein_coding::chr22:43412013-43507848(+) 160.00 -31.22 2 22 41061 41088
25 64.00% 76.00%
--
Read Sequence:ENSG00000112530:ENST00000366888:1.25489280221618:protein_coding::chr6:162727973-163315500(+) (587527 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000112530:ENST00000366888:1.25489280221618:protein_coding::chr6:162727973-163315500(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 160.000000  Q:2 to 22  R:27587 to 27614 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
```

```
>vtRNA1-2 ENSG00000112530:ENST00000366888:1.25489280221618:protein_coding::chr6:162727973-163315500(+) 160.00 -31.53 2 22 27587 27614
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:568446 to 568473 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000112530:ENST00000366888:1.25489280221618:protein_coding::chr6:162727973-163315500(+) 160.00 -31.40 2 22
568446 568473 25 64.00% 76.00%
--
Read Sequence:ENSG00000120278:ENST00000644968:1.25453506925117:protein_coding::chr6:150683611-150843350(+) (159739 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000120278:ENST00000644968:1.25453506925117:protein_coding::chr6:150683611-150843350(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:48475 to 48502 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000120278:ENST00000644968:1.25453506925117:protein_coding::chr6:150683611-150843350(+) 160.00 -31.40 2 22 48475 48502
25 64.00% 76.00%
--
Read Sequence:ENSG00000248905:ENST00000616417:1.24912252062789:protein_coding::chr15:32765543-33194714(-) (429171 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000248905:ENST00000616417:1.24912252062789:protein_coding::chr15:32765543-33194714(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:114070 to 114097 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  ||| |||||:     |:||||||||
   Ref:      5' caAAGUGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.940001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000248905:ENST00000616417:1.24912252062789:protein_coding::chr15:32765543-33194714(-) 164.00 -31.94 2 22
114070 114097 25 68.00% 76.00%


   Forward: Score: 163.000000  Q:2 to 22  R:317534 to 317562 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' caAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -32.840000 kCal/Mol

Scores for this hit:
```

```
>vtRNA1-2 ENSG00000248905:ENST00000616417:1.24912252062789:protein_coding::chr15:32765543-33194714(-) 163.00 -32.84 2 22
317534 317562 26 69.23% 76.92%
--
Read Sequence:ENSG00000184619:ENST00000396267:1.24705859275607:protein_coding::chr17:8358821-8376704(-) (17883 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000184619:ENST00000396267:1.24705859275607:protein_coding::chr17:8358821-8376704(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:14481 to 14509 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000184619:ENST00000396267:1.24705859275607:protein_coding::chr17:8358821-8376704(-) 163.00 -31.81 2 22 14481 14509 26
69.23% 76.92%
--
Read Sequence:ENSG00000157423:ENST00000393567:1.24046483485005:protein_coding::chr16:70802083-71230722(-) (428639 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000157423:ENST00000393567:1.24046483485005:protein_coding::chr16:70802083-71230722(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:61397 to 61424 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cuAGUAGCUGGGAUUAUAGGCACCCGCc 3'

   Energy:  -30.910000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000157423:ENST00000393567:1.24046483485005:protein_coding::chr16:70802083-71230722(-) 160.00 -30.91 2 22 61397 61424
25 64.00% 76.00%
--
Read Sequence:ENSG00000267022:ENST00000591793:1.22945956399843:protein_coding::chr19:44025353-44087318(+) (61965 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000267022:ENST00000591793:1.22945956399843:protein_coding::chr19:44025353-44087318(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:8005 to 8032 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000267022:ENST00000591793:1.22945956399843:protein_coding::chr19:44025353-44087318(+) 160.00 -31.53 2 22 8005 8032 25
64.00% 76.00%
--
Read Sequence:ENSG00000152377:ENST00000394945:1.21411956061556:protein_coding::chr5:136975297-137499326(-) (524029 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000152377:ENST00000394945:1.21411956061556:protein_coding::chr5:136975297-137499326(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 162.000000  Q:2 to 22  R:75820 to 75845 Align Len (24) (70.83%) (79.17%)
```

```
   Query:    3' uuUUCUCGACC----UUUCGUGGGCGc 5'
                  :|| |||||    :||||||||||
   Ref:      5' gcGAG-GCUGGGGGAGAAGCACCCGCc 3'

   Energy:  -30.860001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000152377:ENST00000394945:1.21411956061556:protein_coding::chr5:136975297-137499326(-) 162.00 -30.86 2 22 75820 75845
24 70.83% 79.17%
--
Read Sequence:ENSG00000169862:ENST00000304623:1.21191467590388:protein_coding::chr5:10971835-11904446(-) (932611 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000169862:ENST00000304623:1.21191467590388:protein_coding::chr5:10971835-11904446(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:864308 to 864335 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000169862:ENST00000304623:1.21191467590388:protein_coding::chr5:10971835-11904446(-) 160.00 -31.53 2 22 864308 864335
25 64.00% 76.00%
--
Read Sequence:ENSG00000114757:ENST00000467460:1.20364309042812:protein_coding::chr3:179794957-180036937(-) (241980 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000114757:ENST00000467460:1.20364309042812:protein_coding::chr3:179794957-180036937(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:205730 to 205757 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000114757:ENST00000467460:1.20364309042812:protein_coding::chr3:179794957-180036937(-) 160.00 -31.40 2 22
205730 205757 25 64.00% 76.00%
--
Read Sequence:ENSG00000146966:ENST00000496613:1.20316442026389:protein_coding::chr7:140518419-140640813(-) (122394 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000146966:ENST00000496613:1.20316442026389:protein_coding::chr7:140518419-140640813(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:65901 to 65928 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000146966:ENST00000496613:1.20316442026389:protein_coding::chr7:140518419-140640813(-) 160.00 -31.40 2 22 65901 65928
```

```
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:78259 to 78286 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000146966:ENST00000496613:1.20316442026389:protein_coding::chr7:140518419-140640813(-) 160.00 -31.53 2 22 78259 78286
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:108656 to 108683 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000146966:ENST00000496613:1.20316442026389:protein_coding::chr7:140518419-140640813(-) 160.00 -31.40 2 22
108656 108683 25 64.00% 76.00%
--
Read Sequence:ENSG00000142621:ENST00000688493:1.19739851997352:protein_coding::chr1:15247222-15398328(+) (151106 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000142621:ENST00000688493:1.19739851997352:protein_coding::chr1:15247222-15398328(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:128992 to 129019 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGCAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000142621:ENST00000688493:1.19739851997352:protein_coding::chr1:15247222-15398328(+) 160.00 -31.53 2 22 128992 129019
25 64.00% 76.00%
--
Read Sequence:ENSG00000132334:ENST00000254667:1.19402008324098:protein_coding::chr10:127907102-128085855(+) (178753 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000132334:ENST00000254667:1.19402008324098:protein_coding::chr10:127907102-128085855(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:108389 to 108416 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000132334:ENST00000254667:1.19402008324098:protein_coding::chr10:127907102-128085855(+) 160.00 -31.53 2 22
```

```
108389 108416 25 64.00% 76.00%
--
Read Sequence:ENSG00000130433:ENST00000252729:1.19338185923745:protein_coding::chr19:53991636-54012666(+) (21030 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000130433:ENST00000252729:1.19338185923745:protein_coding::chr19:53991636-54012666(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:15022 to 15049 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000130433:ENST00000252729:1.19338185923745:protein_coding::chr19:53991636-54012666(+) 160.00 -31.40 2 22 15022 15049
25 64.00% 76.00%
--
Read Sequence:ENSG00000132932:ENST00000381655:1.18544762032736:protein_coding::chr13:25371973-26025851(+) (653878 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000132932:ENST00000381655:1.18544762032736:protein_coding::chr13:25371973-26025851(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 172.000000  Q:2 to 19  R:97071 to 97092 Align Len (17) (88.24%) (94.12%)

   Query:    3' uuuucUCGACCUUUCGUGGGCGc 5'
                     |||||| |:||||||||
   Ref:      5' agaccAGCUGG-AGGCACCCGCc 3'

   Energy:  -30.540001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000132932:ENST00000381655:1.18544762032736:protein_coding::chr13:25371973-26025851(+) 172.00 -30.54 2 19 97071 97092
17 88.24% 94.12%


   Forward: Score: 160.000000  Q:2 to 22  R:604939 to 604966 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000132932:ENST00000381655:1.18544762032736:protein_coding::chr13:25371973-26025851(+) 160.00 -31.53 2 22
604939 604966 25 64.00% 76.00%
--
Read Sequence:ENSG00000112139:ENST00000434837:1.18111361682685:protein_coding::chr6:37630678-37697883(-) (67205 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000112139:ENST00000434837:1.18111361682685:protein_coding::chr6:37630678-37697883(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 181.000000  Q:2 to 22  R:64590 to 64612 Align Len (20) (80.00%) (90.00%)

   Query:    3' uuUUCUCGACCUUUCGUGGGCGc 5'
                  || :|| ||:||||||||||
   Ref:      5' ccAAAGGCCGGGAAGCACCCGCc 3'
```

```
   Energy:  -32.049999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000112139:ENST00000434837:1.18111361682685:protein_coding::chr6:37630678-37697883(-) 181.00 -32.05 2 22 64590 64612
20 80.00% 90.00%


   Forward: Score: 160.000000  Q:2 to 22  R:15507 to 15534 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000112139:ENST00000434837:1.18111361682685:protein_coding::chr6:37630678-37697883(-) 160.00 -31.40 2 22 15507 15534
25 64.00% 76.00%
--
Read Sequence:ENSG00000067066:ENST00000340126:1.18033525053466:protein_coding::chr2:230416200-230545606(+) (129406 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000067066:ENST00000340126:1.18033525053466:protein_coding::chr2:230416200-230545606(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:72337 to 72365 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000067066:ENST00000340126:1.18033525053466:protein_coding::chr2:230416200-230545606(+) 163.00 -31.81 2 22 72337 72365
26 69.23% 76.92%
--
Read Sequence:ENSG00000179603:ENST00000339582:1.16932253915751:protein_coding::chr7:126438597-127252941(-) (814344 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000179603:ENST00000339582:1.16932253915751:protein_coding::chr7:126438597-127252941(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 164.000000  Q:2 to 22  R:199261 to 199288 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: |||||||     |:||||||||
   Ref:      5' ugAGUAGCUGGAACUACAGGCACCCGCc 3'

   Energy:  -31.490000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000179603:ENST00000339582:1.16932253915751:protein_coding::chr7:126438597-127252941(-) 164.00 -31.49 2 22
199261 199288 25 68.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:119631 to 119658 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'
```

```
   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000179603:ENST00000339582:1.16932253915751:protein_coding::chr7:126438597-127252941(-) 160.00 -31.53 2 22
119631 119658 25 64.00% 76.00%
--
Read Sequence:ENSG00000162444:ENST00000294435:1.15561716112417:protein_coding::chr1:9997227-10016021(+) (18794 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000162444:ENST00000294435:1.15561716112417:protein_coding::chr1:9997227-10016021(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:13313 to 13341 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000162444:ENST00000294435:1.15561716112417:protein_coding::chr1:9997227-10016021(+) 163.00 -31.81 2 22 13313 13341 26
69.23% 76.92%
--
Read Sequence:ENSG00000130226:ENST00000377770:1.1523797453435:protein_coding::chr7:154052397-154894285(+) (841888 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000130226:ENST00000377770:1.1523797453435:protein_coding::chr7:154052397-154894285(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 172.000000  Q:2 to 21  R:9217 to 9239 Align Len (19) (73.68%) (89.47%)

   Query:    3' uuuUCUCGACCUUUCGUGGGCGc 5'
                  ||:|  ||:|:||||||||
   Ref:      5' cacAGGGGGGGGAGGCACCCGCu 3'

   Energy:  -32.580002 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000130226:ENST00000377770:1.1523797453435:protein_coding::chr7:154052397-154894285(+) 172.00 -32.58 2 21 9217 9239 19
73.68% 89.47%


   Forward: Score: 163.000000  Q:2 to 22  R:615947 to 615975 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000130226:ENST00000377770:1.1523797453435:protein_coding::chr7:154052397-154894285(+) 163.00 -31.81 2 22
615947 615975 26 69.23% 76.92%
--
Read Sequence:ENSG00000204442:ENST00000375915:1.14173152602822:protein_coding::chr13:107163509-107867496(-) (703987 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000204442:ENST00000375915:1.14173152602822:protein_coding::chr13:107163509-107867496(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:80872 to 80899 Align Len (25) (64.00%) (76.00%)
```

```
   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000204442:ENST00000375915:1.14173152602822:protein_coding::chr13:107163509-107867496(-) 160.00 -31.40 2 22
80872 80899 25 64.00% 76.00%
--
Read Sequence:ENSG00000171840:ENST00000305108:1.13543608087471:protein_coding::chr12:564295-663445(-) (99150 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000171840:ENST00000305108:1.13543608087471:protein_coding::chr12:564295-663445(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:61687 to 61714 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000171840:ENST00000305108:1.13543608087471:protein_coding::chr12:564295-663445(-) 160.00 -31.40 2 22 61687 61714 25
64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:78747 to 78774 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGCAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000171840:ENST00000305108:1.13543608087471:protein_coding::chr12:564295-663445(-) 160.00 -31.53 2 22 78747 78774 25
64.00% 76.00%
--
Read Sequence:ENSG00000180287:ENST00000536534:1.1320074674485:protein_coding::chr1:242082985-242524697(-) (441712 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000180287:ENST00000536534:1.1320074674485:protein_coding::chr1:242082985-242524697(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:12385 to 12412 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACC-----UUUCGUGGGCGc 5'
                  |: ||||||     :|:||||||||
   Ref:      5' ccAGUAGCUGGGACUAGAGGCACCCGCc 3'

   Energy:  -30.780001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000180287:ENST00000536534:1.1320074674485:protein_coding::chr1:242082985-242524697(-) 160.00 -30.78 2 22 12385 12412
25 64.00% 76.00%
--
Read Sequence:ENSG00000182601:ENST00000331351:1.13091595590314:protein_coding::chr16:25691958-26137685(+) (445727 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000182601:ENST00000331351:1.13091595590314:protein_coding::chr16:25691958-26137685(+)
```

```
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 160.000000  Q:2 to 22  R:129224 to 129251 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000182601:ENST00000331351:1.13091595590314:protein_coding::chr16:25691958-26137685(+) 160.00 -31.40 2 22
129224 129251 25 64.00% 76.00%


   Forward: Score: 159.000000  Q:2 to 22  R:136506 to 136534 Align Len (26) (65.38%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| :|||||:     |:||||||||
   Ref:      5' ccAAGUGGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -30.879999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000182601:ENST00000331351:1.13091595590314:protein_coding::chr16:25691958-26137685(+) 159.00 -30.88 2 22
136506 136534 26 65.38% 76.92%
--
Read Sequence:ENSG00000160712:ENST00000368485:1.11882670327008:protein_coding::chr1:154405342-154469450(+) (64108 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000160712:ENST00000368485:1.11882670327008:protein_coding::chr1:154405342-154469450(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:12516 to 12544 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000160712:ENST00000368485:1.11882670327008:protein_coding::chr1:154405342-154469450(+) 163.00 -31.81 2 22 12516 12544
26 69.23% 76.92%
--
Read Sequence:ENSG00000261308:ENST00000618634:1.11060830822487:protein_coding::chr12:51817898-51848718(-) (30820 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000261308:ENST00000618634:1.11060830822487:protein_coding::chr12:51817898-51848718(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:18536 to 18564 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                 ||| ||||||:     |:||||||||
   Ref:      5' ucAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000261308:ENST00000618634:1.11060830822487:protein_coding::chr12:51817898-51848718(-) 163.00 -31.81 2 22 18536 18564
26 69.23% 76.92%
```

```
--
Read Sequence:ENSG00000065717:ENST00000262953:1.11052856985362:protein_coding::chr19:2997643-3029280(-) (31637 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000065717:ENST00000262953:1.11052856985362:protein_coding::chr19:2997643-3029280(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 159.000000  Q:2 to 22  R:8299 to 8326 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACC----U-UUCGUGGGCGc 5'
                  |: ||||||    | |:||||||||
   Ref:      5' cgAGUAGCUGGUAUUAUAGGCACCCGCc 3'

   Energy:  -30.370001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000065717:ENST00000262953:1.11052856985362:protein_coding::chr19:2997643-3029280(-) 159.00 -30.37 2 22 8299 8326 25
68.00% 76.00%
--
Read Sequence:ENSG00000267179:ENST00000590798:1.08271751933249:protein_coding::chr19:11925074-11979290(+) (54216 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000267179:ENST00000590798:1.08271751933249:protein_coding::chr19:11925074-11979290(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 168.000000  Q:2 to 22  R:42455 to 42481 Align Len (24) (70.83%) (79.17%)

   Query:    3' uuUUCUCGACCU----UUCGUGGGCGc 5'
                  |: |||||||    |:||||||||
   Ref:      5' cgAGUAGCUGGACUACAGGCACCCGCc 3'

   Energy:  -31.750000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000267179:ENST00000590798:1.08271751933249:protein_coding::chr19:11925074-11979290(+) 168.00 -31.75 2 22 42455 42481
24 70.83% 79.17%


   Forward: Score: 163.000000  Q:2 to 22  R:7687 to 7715 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000267179:ENST00000590798:1.08271751933249:protein_coding::chr19:11925074-11979290(+) 163.00 -31.81 2 22 7687 7715 26
69.23% 76.92%
--
Read Sequence:ENSG00000188385:ENST00000668452:1.07586850149636:protein_coding::chr10:132036335-132184832(+) (148497 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000188385:ENST00000668452:1.07586850149636:protein_coding::chr10:132036335-132184832(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 171.000000  Q:2 to 22  R:130851 to 130877 Align Len (24) (75.00%) (83.33%)

   Query:    3' uuUUCUCGAC---CUUUC-GUGGGCGc 5'
                  |||:|:||   ||||| |||||||
   Ref:      5' aaAAGGGUUGUUAGAAAGCCACCCGCg 3'

   Energy:  -35.000000 kCal/Mol
```

```
Scores for this hit:
>vtRNA1-2 ENSG00000188385:ENST00000668452:1.07586850149636:protein_coding::chr10:132036335-132184832(+) 171.00 -35.00 2 22
130851 130877 24 75.00% 83.33%


   Forward: Score: 164.000000  Q:2 to 22  R:12426 to 12453 Align Len (25) (68.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                 |: |||||||     |:||||||||
   Ref:      5' ugAGUAGCUGGAACUACAGGCACCCGCc 3'

   Energy:  -31.490000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000188385:ENST00000668452:1.07586850149636:protein_coding::chr10:132036335-132184832(+) 164.00 -31.49 2 22
12426 12453 25 68.00% 76.00%
--
Read Sequence:ENSG00000183570:ENST00000681687:1.06058392223854:protein_coding::chr21:45643724-45942450(+) (298726 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000183570:ENST00000681687:1.06058392223854:protein_coding::chr21:45643724-45942450(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 175.000000  Q:2 to 21  R:283829 to 283853 Align Len (21) (80.95%) (85.71%)

   Query:    3' uuuUCUCGA--CCUUUCGUGGGCGc 5'
                  ||||||  || :|||||||||
   Ref:      5' ggcAGAGCUAGGGUGAGCACCCGCc 3'

   Energy:  -32.830002 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000183570:ENST00000681687:1.06058392223854:protein_coding::chr21:45643724-45942450(+) 175.00 -32.83 2 21
283829 283853 21 80.95% 85.71%


   Forward: Score: 170.000000  Q:2 to 20  R:201757 to 201780 Align Len (19) (84.21%) (84.21%)

   Query:    3' uuuuCUCGAC-CUUUCGUGGGCGc 5'
                   ||||||   ||||||||||
   Ref:      5' guguGAGCUGCUUAAGCACCCGCg 3'

   Energy:  -35.060001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000183570:ENST00000681687:1.06058392223854:protein_coding::chr21:45643724-45942450(+) 170.00 -35.06 2 20
201757 201780 19 84.21% 84.21%
--
Read Sequence:ENSG00000166780:ENST00000300006:1.05927917544278:protein_coding::chr16:15434587-15588259(+) (153672 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000166780:ENST00000300006:1.05927917544278:protein_coding::chr16:15434587-15588259(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 171.000000  Q:2 to 22  R:45649 to 45676 Align Len (25) (76.00%) (80.00%)

   Query:    3' uuUUC-UCGACCU----UUCGUGGGCGc 5'
                 ||| |||||||    |:||||||||
   Ref:      5' ccAAGUAGCUGGACUACAGGCACCCGCc 3'

   Energy:  -32.160000 kCal/Mol
```

```
Scores for this hit:
>vtRNA1-2 ENSG00000166780:ENST00000300006:1.05927917544278:protein_coding::chr16:15434587-15588259(+) 171.00 -32.16 2 22 45649 45676
25 76.00% 80.00%


   Forward: Score: 160.000000  Q:2 to 22  R:97769 to 97796 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cuAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -30.910000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000166780:ENST00000300006:1.05927917544278:protein_coding::chr16:15434587-15588259(+) 160.00 -30.91 2 22 97769 97796
25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:110510 to 110537 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGAUUAUAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000166780:ENST00000300006:1.05927917544278:protein_coding::chr16:15434587-15588259(+) 160.00 -31.40 2 22
110510 110537 25 64.00% 76.00%


   Forward: Score: 160.000000  Q:2 to 22  R:146085 to 146112 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000166780:ENST00000300006:1.05927917544278:protein_coding::chr16:15434587-15588259(+) 160.00 -31.40 2 22
146085 146112 25 64.00% 76.00%
--
Read Sequence:ENSG00000172349:ENST00000683961:1.05273239575006:protein_coding::chr15:81196878-81314058(+) (117180 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000172349:ENST00000683961:1.05273239575006:protein_coding::chr15:81196878-81314058(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 171.000000  Q:2 to 22  R:95966 to 95990 Align Len (22) (77.27%) (86.36%)

   Query:    3' uuUUCUCGACCUU-UC-GUGGGCGc 5'
                  :||||| ||:| || |||||||
   Ref:      5' caGAGAGCCGGGACAGCCACCCGCc 3'

   Energy:  -30.360001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000172349:ENST00000683961:1.05273239575006:protein_coding::chr15:81196878-81314058(+) 171.00 -30.36 2 22 95966 95990
22 77.27% 86.36%
```

```
--
Read Sequence:ENSG00000118922:ENST00000377669:1.04167555784089:protein_coding::chr13:73686088-74133929(-) (447841 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000118922:ENST00000377669:1.04167555784089:protein_coding::chr13:73686088-74133929(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:266015 to 266043 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ucAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000118922:ENST00000377669:1.04167555784089:protein_coding::chr13:73686088-74133929(-) 163.00 -31.81 2 22
266015 266043 26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:41740 to 41767 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  :| ||||||:     |:||||||||
   Ref:      5' caGAUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -30.129999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000118922:ENST00000377669:1.04167555784089:protein_coding::chr13:73686088-74133929(-) 160.00 -30.13 2 22 41740 41767
25 64.00% 76.00%
--
Read Sequence:ENSG00000172318:ENST00000392690:1.04076470116614:protein_coding::chr2:167293000-167874045(+) (581045 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000172318:ENST00000392690:1.04076470116614:protein_coding::chr2:167293000-167874045(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:215382 to 215410 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000172318:ENST00000392690:1.04076470116614:protein_coding::chr2:167293000-167874045(+) 163.00 -31.81 2 22
215382 215410 26 69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:574051 to 574078 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000172318:ENST00000392690:1.04076470116614:protein_coding::chr2:167293000-167874045(+) 160.00 -31.40 2 22
574051 574078 25 64.00% 76.00%
```

```
--
Read Sequence:ENSG00000185652:ENST00000423158:1.03592170287384:protein_coding::chr12:5432107-5495299(+) (63192 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000185652:ENST00000423158:1.03592170287384:protein_coding::chr12:5432107-5495299(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 161.000000  Q:2 to 22  R:1195 to 1214 Align Len (20) (75.00%) (80.00%)

   Query:    3' uuUUCUCGACCUUUCGUGGGCGc 5'
                ::||| ||||   ||||||||
   Ref:      5' ggGAGACCUGG---GCACCCGCg 3'

   Energy:  -31.450001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000185652:ENST00000423158:1.03592170287384:protein_coding::chr12:5432107-5495299(+) 161.00 -31.45 2 22 1195 1214 20
75.00% 80.00%
--
Read Sequence:ENSG00000116183:ENST00000367662:1.03209129019026:protein_coding::chr1:176463174-176845601(+) (382427 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000116183:ENST00000367662:1.03209129019026:protein_coding::chr1:176463174-176845601(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 160.000000  Q:2 to 22  R:253562 to 253589 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000116183:ENST00000367662:1.03209129019026:protein_coding::chr1:176463174-176845601(+) 160.00 -31.40 2 22
253562 253589 25 64.00% 76.00%
--
Read Sequence:ENSG00000136040:ENST00000258526:1.02595890569604:protein_coding::chr12:94148576-94307675(+) (159099 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000136040:ENST00000258526:1.02595890569604:protein_coding::chr12:94148576-94307675(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 163.000000  Q:2 to 22  R:83634 to 83662 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000136040:ENST00000258526:1.02595890569604:protein_coding::chr12:94148576-94307675(+) 163.00 -31.81 2 22 83634 83662
26 69.23% 76.92%
--
Read Sequence:ENSG00000173918:ENST00000579760:1.02378412000325:protein_coding::chr17:79024168-79049788(+) (25620 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-
Performing Scan: vtRNA1-2 vs ENSG00000173918:ENST00000579760:1.02378412000325:protein_coding::chr17:79024168-79049788(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-

   Forward: Score: 156.000000  Q:2 to 21  R:24116 to 24135 Align Len (19) (73.68%) (78.95%)

   Query:    3' uuuUCUCGACCUUUCGUGGGCGc 5'
```

```
                 ||:|| ||   ||||||||
   Ref:      5' accAGGGCGGG---GCACCCGCg 3'

   Energy:  -32.110001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000173918:ENST00000579760:1.02378412000325:protein_coding::chr17:79024168-79049788(+) 156.00 -32.11 2 21 24116 24135
19 73.68% 78.95%
--
Read Sequence:ENSG00000167216:ENST00000683218:1.01272301110128:protein_coding::chr18:46917601-47102243(+) (184642 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000167216:ENST00000683218:1.01272301110128:protein_coding::chr18:46917601-47102243(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 163.000000  Q:2 to 22  R:3650 to 3678 Align Len (26) (69.23%) (76.92%)

   Query:    3' uuUUC-UCGACCU-----UUCGUGGGCGc 5'
                  ||| ||||||:     |:||||||||
   Ref:      5' ccAAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.809999 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000167216:ENST00000683218:1.01272301110128:protein_coding::chr18:46917601-47102243(+) 163.00 -31.81 2 22 3650 3678 26
69.23% 76.92%


   Forward: Score: 160.000000  Q:2 to 22  R:131358 to 131385 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' cgAGUAGCUGGGACUACAGGCACCCGCc 3'

   Energy:  -31.400000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000167216:ENST00000683218:1.01272301110128:protein_coding::chr18:46917601-47102243(+) 160.00 -31.40 2 22
131358 131385 25 64.00% 76.00%
--
Read Sequence:ENSG00000156113:ENST00000639486:1.01174629638545:protein_coding::chr10:76872903-77637993(-) (765090 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000156113:ENST00000639486:1.01174629638545:protein_coding::chr10:76872903-77637993(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 159.000000  Q:2 to 22  R:398575 to 398603 Align Len (26) (69.23%) (73.08%)

   Query:    3' uuUUCUCGACCUU-----U-CGUGGGCGc 5'
                  :||||||||||     |  |||||||
   Ref:      5' uuGAGAGCUGGAAUUUCCAUACACCCGCu 3'

   Energy:  -30.090000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000156113:ENST00000639486:1.01174629638545:protein_coding::chr10:76872903-77637993(-) 159.00 -30.09 2 22
398575 398603 26 69.23% 73.08%
--
Read Sequence:ENSG00000108387:ENST00000672673:1.01108823992863:protein_coding::chr17:58520255-58544328(-) (24073 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000108387:ENST00000672673:1.01108823992863:protein_coding::chr17:58520255-58544328(-)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
```

```
   Forward: Score: 160.000000  Q:2 to 22  R:6606 to 6633 Align Len (25) (64.00%) (76.00%)

   Query:    3' uuUUCUCGACCU-----UUCGUGGGCGc 5'
                  |: ||||||:     |:||||||||
   Ref:      5' ugAGUAGCUGGGAUUACAGGCACCCGCc 3'

   Energy:  -31.530001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000108387:ENST00000672673:1.01108823992863:protein_coding::chr17:58520255-58544328(-) 160.00 -31.53 2 22 6606 6633 25
64.00% 76.00%
--
Read Sequence:ENSG00000053918:ENST00000155840:1.00084978623029:protein_coding::chr11:2445007-2849105(+) (404098 nt)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=
Performing Scan: vtRNA1-2 vs ENSG00000053918:ENST00000155840:1.00084978623029:protein_coding::chr11:2445007-2849105(+)
=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

   Forward: Score: 168.000000  Q:2 to 22  R:273074 to 273097 Align Len (21) (76.19%) (80.95%)

   Query:    3' uuUUCUCGACCUUUC-GUGGGCGc 5'
                  |||||| ||:  | |||||||
   Ref:      5' ggAAGAGCCGGGUUGACACCCGCg 3'

   Energy:  -34.200001 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000053918:ENST00000155840:1.00084978623029:protein_coding::chr11:2445007-2849105(+) 168.00 -34.20 2 22 273074 273097
21 76.19% 80.95%


   Forward: Score: 167.000000  Q:2 to 22  R:87976 to 87997 Align Len (20) (75.00%) (85.00%)

   Query:    3' uuUUCUCGACCUUUCGUGGGCGc 5'
                  |:||||| |  :||||||||
   Ref:      5' ucAGGAGCU-GCUGGCACCCGCc 3'

   Energy:  -30.520000 kCal/Mol

Scores for this hit:
>vtRNA1-2 ENSG00000053918:ENST00000155840:1.00084978623029:protein_coding::chr11:2445007-2849105(+) 167.00 -30.52 2 22 87976 87997 20
75.00% 85.00%
```