Computers in Biology and Medicine 106 (2019) 106–113

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

You are what you eat: Sequence analysis reveals how plant microRNAs may regulate the human genome

Bahareh Kashani[a,1], Mansoor Hasani Bidgoli[b,1], Seyed Abolfazl Motahari[b], Nafiseh Sedaghat[c], Mohammad Hossein Modarressi[a,*]

[a] Department of Medical Genetics, Faculty of Medicine, Tehran University of Medical Sciences (TUMS), Tehran, Iran
[b] Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
[c] Computer Engineering School, Iran University of Science and Technology, Tehran, Iran

ARTICLE INFO

Keywords:
Diet
Nutrigenomics
MicroRNA
Plant
Gene regulation
Target prediction

ABSTRACT

Background: Nutrigenomic has revolutionized our understanding of nutrition. As plants make up a noticeable part of our diet, in the present study we chose microRNAs of edible plants and investigated if they can perfectly match human genes, indicating potential regulatory functionalities.

Methods: miRNAs were obtained using the PNRD database. Edible plants were separated and microRNAs in common in at least four of them entered our analysis. Using vmatchPattern, these 64 miRNAs went through four steps of refinement to improve target prediction: Alignment with the whole genome (2581 results), filtered for those in gene regions (1371 results), filtered for exon regions (66 results) and finally alignment with the human CDS (41 results). The identified genes were further analyzed in-silico to find their functions and relations to human diseases.

Results: Four common plant miRNAs were identified to match perfectly with 22 human transcripts. The identified target genes were involved in a broad range of body functions, from muscle contraction to tumor suppression. We could also indicate some connections between these findings and folk herbology and botanical medicine.

Conclusions: The food that we regularly eat has a great potential in affecting our genome and altering body functions. Plant miRNAs can provide means of designing drugs for a vast range of health problems including obesity and cancer, since they target genes involved in cell cycle (CCNC), digestion (GIPR) and muscular contractions (MYLK). They can also target regions of CDS for which we still have no sufficient information, to help boost our knowledge of the human genome.

1. Introduction

Nutrigenomics and nutrigenetics are emerging scientific terms that are evolving rapidly; they represent how our diet has a great impact on different aspects of the genome functions [1–3]. The ultimate goal of nutrigenomics is to make dietary-intervention strategies possible for patients and even use nutrition to prevent diseases [4,5]. A very useful process by which nutrition and genome are connected is RNA interference [6].

RNA interference (RNAi) is a common post-transcription gene regulation mechanism which contains two main functional molecules: siRNAs (small interfering ribonucleic acids) and miRNAs (micro ribonucleic acids). MicroRNAs (miRNAs, miRs) are short, single-stranded ribonucleotides with an approximate length of 19–24 nucleotides. Genes coding for these molecules comprise 0.5 to 1% of the human genome, but it is estimated that at least half of the human genes are regulated by at least one miRNA [7]. Despite the known functions of endogenous miRNAs, the presence of active exogenous miRNAs (XenomiRs) seemed not possible, considering the sensitive nature of RNA which is caused by its 2′-OH active group and single-stranded structure. However, in 2012 Zhang et al. [8] discovered that not only plant miRs are present in the human serum, they can also be acquired orally and through food intake (Horizontal miRNA transfer) [9]. It is estimated that plant miRNAs can regulate 30% of protein-coding genes in

---

[*] Corresponding author. Department of Medical Genetics, School of Medicine, Tehran University of Medical Sciences, Poursina St., Tehran, P.O.Box: 14155-6447, Iran.

E-mail address: modaresi@sina.tums.ac.ir (M.H. Modarressi).

[1] These authors contributed equally to this paper as first authors.

https://doi.org/10.1016/j.compbiomed.2019.01.020
Received 24 October 2018; Received in revised form 1 January 2019; Accepted 21 January 2019

mammals [10]. This research started fierce debates on the possibility of the process and initiated other studies trying to reproduce the reported results. Some researches claimed that Zhang's findings might have been the result of a false positive report [11]. Corn feeding regimen was also performed on mice and no detectable plant miRNA increase was detected in the samples' blood [12]. Despite these controversies, the same phenomenon has been reported for the human breast milk miRNAs (Vertical miRNA transfer) [13].

After surviving the digestion, miRNAs use extracellular vesicles, e.g. exosomes [14], microvesicles, specific lipoproteins [15], some viruses and also argonate-containing bodies [16] to enter the blood stream. “Trancytosis” is another process to enter the body fluids, which means forming a vesicle from the intestine walls [17]. Immune system cells [18] and specific transporters [19,20] may also help this transportation.

When miRNAs enter the human body by food intake, they can regulate the gene expression mostly through affecting messenger RNAs [21]. Using Watson-Crick base pairing, miRNAs can attach their target mRNAs. According to the level of similarity, two pathways may be activated: If they match perfectly, Ago-dependent endonucleolysis degrades the mRNA. This process mostly begins by removing the cap and the polyA sites and then cleavage occurs between base 10 and 11 of miR [22]. Eventually, the target mRNA will be totally degraded and its translation can no longer occur. This mechanism happens in the P-bodies of the cells, where DCP1:DCP2 proteins (DeCapping Proteins) are present for decapping [23]. If the miRNA and the mRNA match partially, this heteroduplex competes with ribosome attachment and hence greatly reduces translation without actually decaying the mRNA. There are some researches suggesting that miRNA attachment to the promoter region of some genes may actually have a different function and increase the gene expression instead [24]. These different mechanisms of action suggest that each miRNA can have more than one mRNA as target.

As two main sources of exogenous miRNAs in our diet, plant and animal miRNA functions have some differences. The majority of reported plant miRNAs have a remarkable propensity to regulate genes involved in development, especially those functioning as transcription factors [25], but animal miRNAs seem to have a broader range of targets. This phenomenon may be due to the fact that plant miRNAs are mostly functional when they match perfectly with their target mRNA and this can greatly reduce the number of possible gene targets [21]. Whereas animal miRNAs are able to function whether the match is perfect or not. This dissimilarity makes computational target prediction easier in plants, due to the extensive level of complementarity. Also methods developed for animal lineages do not function properly in plants, since the hairpin part of plant miRNAs is more heterogeneous than those of animal miRNAs [26].

In this study, we used vmatchPattern function in Biostring package version 2.42.1 to detect human mRNAs that match perfectly with the desired set of plant miRNA. The selected miRNAs were those in common in at least four plant species, making it more likely to have a common function rather than a species-specific one and also making the miRNAs more versatile and probable in the human diet.

2. Materials and methods

2.1. The plant miRNAs database

Plant non-coding RNA database was used to obtain miRNAs of different species of plants (http://structuralbiology.cau.edu.cn/PNRD/index.php) [27].

2.2. Whole genome and CDS databases

The reference genome for finding targets on the whole genome was extracted from Bsgenome.Hsapiens.UCSC.hg38 package that encodes the human genome version hg38. First, Bsgenome.Hsapiens.UCSC.hg38 was filtered to only include the chromosomes and all the assembled contigs outside the chromosomes were removed. The filtered genome was used as the subject or reference and the selected microRNAs were used as patterns or queries for vmatchPattern function. The output of vmatchPattern presented the positions of alignments within the chromosomes with exact match for each selected microRNA.

The alignments of the selected miRNAs were also investigated in the exon regions and then with mRNAs’ sequences (the coding sequences, CDS). TxDb.Hsapiens.UCSC.hg38.knownGene package was used for acquiring the exon ranges of mRNAs in the genome and their orders, so that the sequences of all mRNAs could be retrieved. Next, similar to the alignment with the whole genome, mRNA sequences were used as the reference and each selected microRNA was used as a query for vmatchPattern function.

2.3. MiRNA alignment

The selected plant microRNAs were matched to the human whole genome, refined for the gene and then exon regions, and were finally matched to the whole CDS to find potential target genes. To this end, R software version 3.3.2 installed on Ubuntu operating system was used. In particular, vmatchPattern function in Biostring package version 2.42.1 was used for finding exact matches between the miRNAs and the targets.

2.4. In-silico gene identification

Reference websites such as Genecards (www.genecards.org), OMIM (www.omim.org), Genetic home reference (NIH) (https://ghr.nlm.nih.gov), UniProt (www.uniprot.org), Ensemble (www.ensemble.org), and NCBI (https://www.ncbi.nlm.nih.gov) were used to identify the roles of the predicted genes and how they could affect our health. Also miRBase (http://mirbase.org/) was used later to search for the human counterparts for the final selected plant miRs.

3. Results

3.1. Identification of the plant miRNA

As the first step, a list of the plant miRNAs was created. The list contained 5101 miRNAs from different plant species. The sequences were searched for identical sequences among different plants. The miRNAs which were in common in at least four different species were then selected as our study group. These plant miRNAs were considered to be involved in common processes of the organisms, rather than species-specific ones, and have a broader range of target mRNAs, since they are preserved and available in a larger group of species. A total of 64 miRNA sequences were selected to be studied and aligned with the human genome.

3.2. Identification of the target genes by aligning with the human whole genome

To start aligning the selected miRNAs with the genome, the human whole genome was filtered and only chromosomes with names chr 1, …, chr 22, chr X, chr Y, and chr M were retained, i.e. all the assembled contigs outside the chromosomes were removed. The number of alignments for every selected microRNA are shown in Fig. 1.

Among the 64 microRNAs that had been selected, only 10 had aligned with the whole genome. As it is evident from Fig. 1, microRNAs with ids 8, 9, 10, 30, 45, 46 and 58 had alignments that could be due to the repeat structure inside their sequences (Supplemental Table 1).

3.3. Refining the target prediction procedure by filtering for the gene regions

Since microRNAs were aligned with the whole genome at first, all

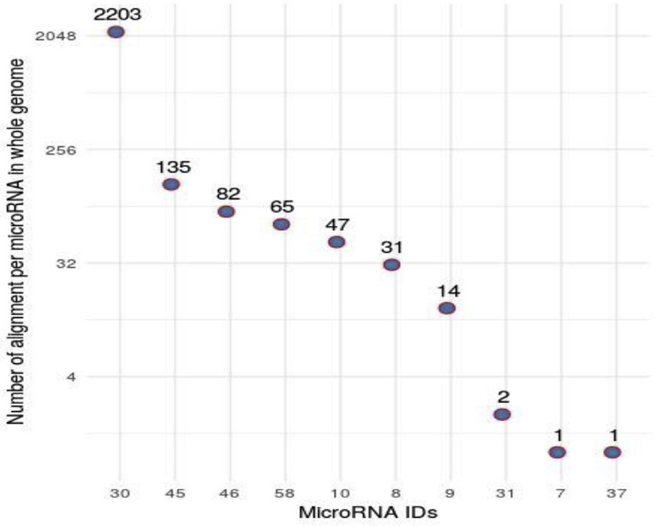

Fig. 1. Number of perfect match alignments with the whole human genome for the microRNAs with at least one match. Nine miRNAs were identified with 2580 alignments in total.

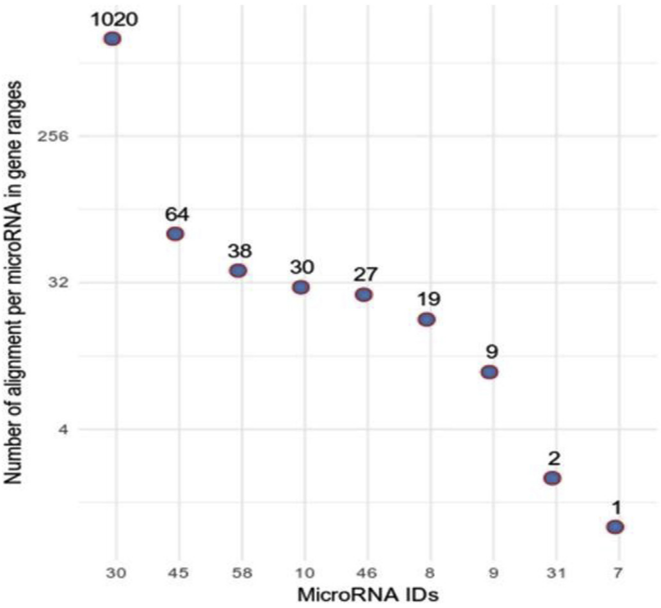

Fig. 2. Of the whole genome alignments, only those with alignment in the gene regions were retained. The number of alignments decreased from 2580 to 1210.

the alignments were searched for gene IDs and those alignments that occurred outside the gene ranges were later filtered. The result of this filtering is depicted in Fig. 2.

By comparing Figs. 1 and 2, it is evident that more than half of the alignments (1371/2581) occurred outside the gene regions (Fig. 3).

3.4. Refining the target prediction procedure by filtering for the exon regions

As microRNAs target mRNAs, the alignments that occur in the exon regions are biologically more plausible to be functional. Therefore, another filtering step was performed and only those alignments that occurred in the exon regions were kept, as shown in Fig. 4. Only 66/1371 alignments were retained. This is due to the facts that:

- It is possible that some microRNAs were aligned with the exon junctions that caused false positive targets. Also, some mRNA targets could be missed (false negative results) due to the intron removal, which brings distant sequences of exons together, or exon shuffling that happens in the process of mRNA transcription.
- Some of the mRNAs are transcribed from the reverse complement of the reference genome which is missing in the reference human genome that was used.

Supplemental Table 2 shows the genes that were the targets of plant microRNAs in the exon regions of the human genome.

3.5. Target prediction credibility

To have a more reliable analysis, aligning microRNAs with the CDS sequences was necessary as well. After retrieving all the human mRNA sequences, miRNAs were aligned with the CDS. Results are shown in the Supplemental Table 3. A notable point in this matter was that some microRNAs were aligned with mRNAs with no specified genes. Comparing the alignments with the human genome after filtering for the exon regions and then the CDS, there was only 1 plant miRNA that matched on both databases, whereas there were other sequences that matched only on one of them. So three groups of miRNAs were defined:

- The microRNAs that aligned with both exon regions and the whole CDS, targeted sequences within a single exon region, where its mRNA was transcribed from the forward strand of the genome (Supplemental Table 4).
- The set of microRNAs that aligned only with the CDS, for which no match could be found on the exon regions (Supplemental Table 5), could be indicators of two possibilities:
  - MicroRNAs were aligned with the splice junction regions, where two distinct exons were brought together after the intron removal. This case was not observed in our experiments.
  - MicroRNAs were aligned with the mRNAs that had been transcribed from the reverse strand complement of the genome.
- The set of microRNAs that only aligned with the exon regions (Supplemental Table 6) could also be indicators of two situations:
  - MicroRNAs aligned with the forward strand of exons, but the related mRNAs were transcribed from the reverse complement of the genome.
  - MicroRNAs aligned within a sequence that contained the border between an exon and an intron. Since these sequences would be detached during the mRNA maturation, these alignments were not valid for further analysis.

Considering the fact that miRNAs target mRNAs, and that transcription can occur from both strands of the DNA, our final analysis was performed on those miRNAs that could align with the CDS. Our final miRNAs' IDs were 8, 10, 30 and 45 which matched perfectly with 41 sequences of CDS and contained 22 different human transcripts (Fig. 5). The identified plant miRNAs had no identical matches among the human miRNAs as approved by miRBase database, but three similar human miRNAs were identified (Supplemental Table 7).

From a statistical point of view, the probability of finding a match between a microRNA and the whole human CDS is approximately 4×10^{-4}. To see why this is the case, we note that the total length of the human CDS is approximately 100 Mbp (123984959 bp) and the length of each microRNA is 21 bp. Therefore, the probability of observing a given sequence of the length of 21 within the CDS would be $10^8 \times 4^{-21}$. Given the fact that only 64 microRNAs were aligned with the CDS in this study, the probability of a random hit was relatively low; this suggested our results as statistically meaningful.

Table 1 presents the selected miRNAs, the plants that contain those, and their target genes. Among these genes are those associated with the cell cycle (*CCNC*), myosin (*MYLK*), actin (*ACTA2*), gastric polypeptide (*GIPR*) and angiogenesis (*FLT1*). Fig. 6 represents an overview of the target genes and their functions.

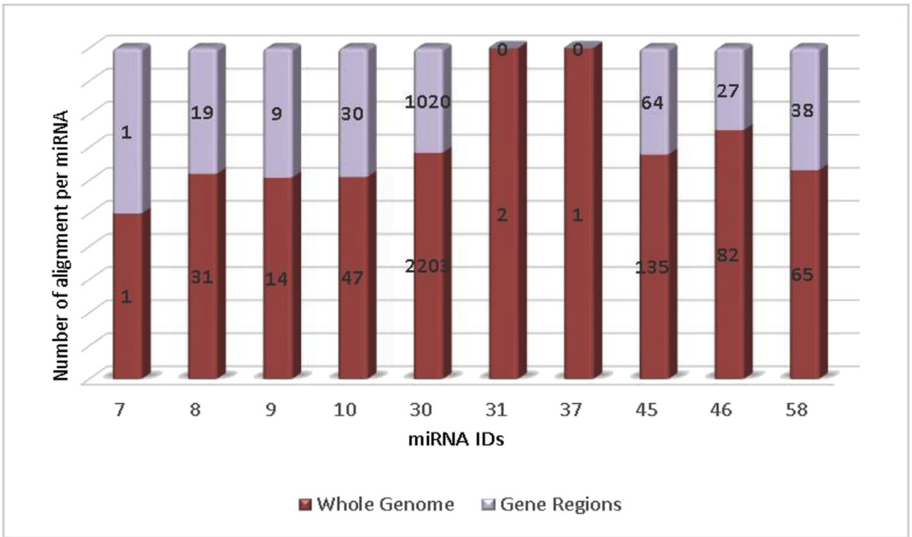

Fig. 3. The alignments with the whole genome were investigated for the gene ID and the results indicated that more than half of the alignments were placed outside the gene regions. As gene regions are more plausible to be targets for microRNAs, only those sequences were retained for further analysis.

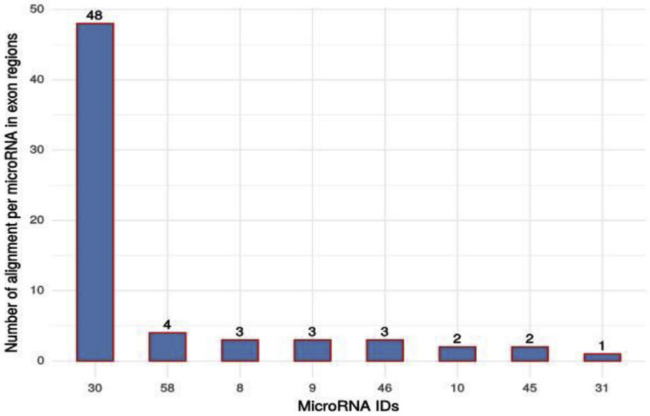

Fig. 4. Only alignments in the exon regions were retained for further analysis. Eight miRNAs had alignments in the exon regions.

3.6. Identifying the target genes

The selected miRNAs targeted diverse transcripts of the human genome and they had been detected in a broad range of edible plants. MiRNA number 8 targeted *CCNC*, a cell cycle regulator, and its best known containing plants were onion and soybean. The plants that contained miRNA number 10, e.g. black seed and wormwood, have long been known in herbal medicine. This miRNA aligned with the transcripts of *PTCHD4* (patched domain containing 4) and *FLT1* (fms related tyrosine kinase 1) that have mostly been studied for their role in cancer.

MiRNA number 30 had the greatest hit numbers and aligned with 13 transcripts. These transcripts contained well-known genes such as *MYLK* (myosin light-chain kinase), ACTA2 (Actin, Alpha 2, Smooth Muscle, Aorta) and *GIPR* (Gastric Inhibitory Polypeptide Receptor), as well as three transcripts of anti-sense sequences and two unknown transcripts. This miRNA was also found in the most diverse plant species, from apple and orange to tobacco and opium. MiRNA number 45 aligned with *ERICH3* (glutamate-rich 3, Chromosome 1 Open Reading Frame 173) as well as two unknown transcripts.

4. Discussion

Since Zhang et al. first published their paper and revealed that food

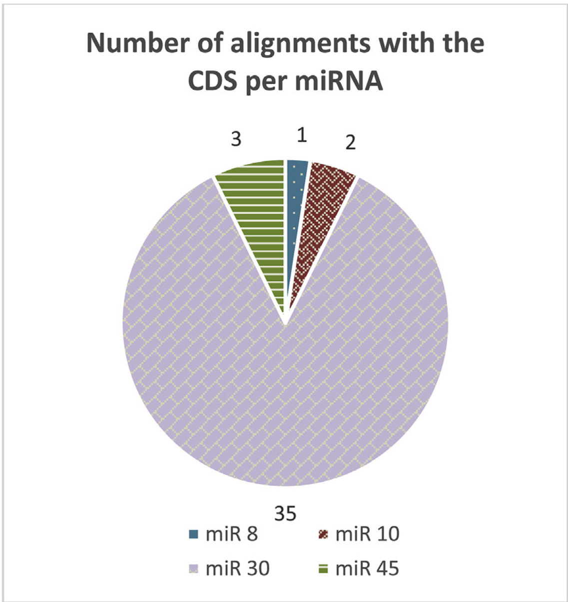

Fig. 5. As the final step, 41 alignments were identified in the human CDS that matched perfectly with 4 plant microRNAs. These alignments consisted of 22 individual genes.

microRNAs are able to enter the human body fluids [8], many scientists have been discussing the subject. Not every scientist has found this discovery as an “eureka” effect. Witwer and his colleagues questioned the possibility of the presence of plant miRs in the bloodstream by using RT-qPCR and described the results as “low quality and inconsistent with specific, reliable detection” [28], although they had previously approved the matter [20]. They also used digital PCR for copy number detection and stated that the levels of plasma xenomiRs did not appear to change after dietary intake. Among the six miRs that were tested (miR-156, 160, 166, 167 and 172), the miR-160 seemed to be the most promising one, yet its copy number was approximately one 1000th of the miR-16, a moderately abundant miRNA in the plasma. Kang et al. also questioned low abundance of miRNAs in the body fluids and tissues [29]. Hence there is still an important question to be answered: how much of a microRNA is needed to cause noticeable changes in the gene regulation?

Table 1
The final selected miRNAs, their associated plant species and their target genes (Unknown genes are identified by their transcript ID).

MicroRNA ID	Scientific Name of the plant	Common name of the plant	Gene/Transcript
8	*Carthamus_tinctorius*	Safflower	*CCNC*
	Trifolium_pratense	Red clover (Probable in honey)	
	Helianthus_paradoxus	Paradox Sunflower	
	Glycine_max	Soybean	
	Allium_cepa	Onion	
	Cucumis_melo_subsp._melo Euphorbia_esula	Kirkman Melon	
	Euphorbia_esula	Leafy Spurge	
10	*Helianthus_petiolaris*	Sunflower	*PTCHD4*
	Artemisia_annua	Wormwood	*FLT1*
	Guizotia_abyssinica	Black seed	
	Helianthus_tuberosus	Jerusalem Artichoke (Kind of sunflower grown for its edible tuberous roots)	
	Thellungiella_halophila	Saltwater Cress	
30	*Guizotia_abyssinica*	Black seed	*DEPDC1-AS1*
	Helianthus_tuberosus	Jerusalem artichoke (Kind of sunflower grown for its edible tuberous roots)	*ERICH3*
	Theobroma_cacao	Cacao tree	*GRM7-AS2*
	Phaseolus_acutifolius	Tepary bean	*MYLK*
	Citrus_clementina	Clementine	*COPB2*
	Carthamus_tinctorius	Safflower	*CDKN2B-AS1*
	Mimulus_guttatus_var._nasutus	Yellow monkeyflower	uc004brm.4
	Centaurea_solstitialis	Yellow cockspur	*ACTA2*
	Dioscorea_alata	Purple yam	*FLT1*
	Paullinia_cupana_var._sorbilis	Guarana	uc059mxz.1
	Artemisia_annua	Wormwood	*PKD1L2*
	Helianthus_ciliaris	Texas blueweed	*GIPR*
	Zinnia_violacea	Zinnia	*FAM83C*
	Mesembryanthemum_crystallinum	Ice plant	
	Raphanus_raphanistrum_subsp._maritimus	Wild radish	
	Papaver_somniferum	Opium poppy	
	Hordeum_vulgare	Barley	
	Fagus_sylvatica	Beech	
	Lactuca_perennis	Blue lettuce	
	Helianthus_paradoxus	Paradox sunflower	
	Volvox_carteri	Volvox	
	Linum_usitatissimum	Flax	
	Coffea_canephora	Robusta coffee	
	Cryptomeria_japonica	Japanese pine	
	Raphanus_raphanistrum	Wild radish	
	Helianthus_annuus	Sunflower	
	Glycine_max	Soybean	
	Arachis_duranensis	Wild Peanut	
	Nicotiana_tabacum	Tobacco	
	Citrus_sinensis	Sweet orange	
	Centaurea_maculosa	Knapweed	
	Beta_vulgaris	Beet	
	Trifolium_pratense	Red clover (Probable in honey)	
	Malus_domestica	Apple tree	
45	*Volvox_carteri*	Volvox	*ERICH3*
	Dioscorea_alata	Purple yam	uc004brm.4
	Mimulus_lewisii	Purple monkeyflower	uc059mxz.1
	Helianthus_annuus	Sunflower	
	Fagus_sylvatica	Beech	
	Arachis_ipaensis	Wild Peanut	
	Taraxacum_officinale	Dandelion	
	Aquilegia_formosa_x_Aquilegia_pubescens	Wolf's bane	
	Thellungiella_halophila	Saltwater cress	
	Glycyrrhiza_uralensis	Chinese liquorice	

Hirschi suggested that these conflicts might be due to the environmental factors, e.g. the gut microbiome which may facilitate the miR transfer into the bloodstream [30]. Whether or not this suggestion is correct needs repetitive experiments containing diverse ethnical populations. Despite the controversies, a growing body of literature is confirming the food miRNAs to be affecting the human genome. As Zhang targeted plant miRNAs in human, Yang et al. confirmed the same finding in animals [31] and Ashby et al. claimed that the plant miRNAs in the royal jelly determine the fate of honey bees as queen [32]. Later Chin et al. even used the plant miR159 to halt breast cancer cell growth [33]. Baier tested animal models and indicated the possibility of animal miRNA transfer, since chicken egg and cow milk could transfer 2 miRNAs through oral intake [34,35]. It has also been proved that the human breast milk can transfer 281 miRNAs from mothers to newborns, which mostly target their immune system. The crucial point about the breast milk miRNAs is that they are resistant to freeze and thaw cycles, very acidic environment (pH = 1, as in the stomach) and even known RNases [36]! It would confirm the possibility that these molecules could actually go through the digestion system unharmed. Yet if these xenomiRs lack endogenous human gene targets, the debate about their survival from the digestion would be irrelevant. Therefore bioinformatic studies such as ours may pave the way for the start of new experiments on food intake and nutrition, as they introduce new miRNAs and target transcripts for further detections.

The transcripts that we found can be categorized into 3 groups:

1 Transcripts with known genes and functions, which were the majority of the target transcripts of this study:

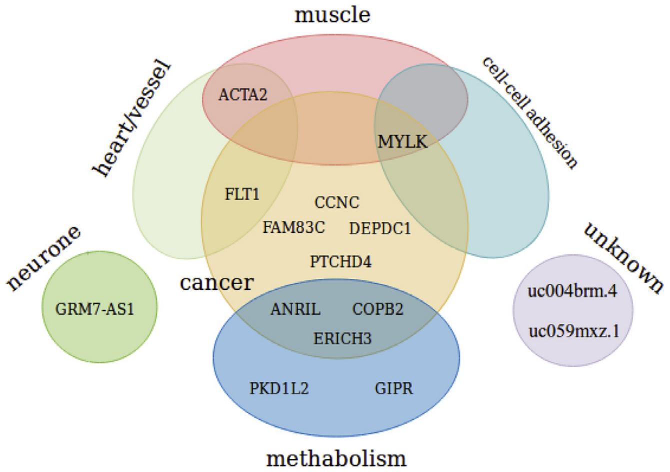

Fig. 6. An overview of the genes of the target transcripts and their function.

CCNC, which is translated to cyclin C protein, was a perfect-match target for miRNA number 8. Cyclin C interacts with CDK8 (cyclin-dependent kinase 8) to act as a major regulator of the G1 phase of cell cycle [37]. This complex is an important transcription regulation by phosphorylation induction in RNA polymerase II. Cyclin C expression has also been linked to the pathogenesis of Alzheimer's disease [38].

PTCHD4 were two targets for miRNA number 10. *PTCHD4* is a P53 target that can suppress the hedgehog pathway and acts as a tumor suppressor [39]. On the other hand, *FLT1* is a receptor for VEGF-A, VEGF-B (Vascular Endothelial Growth Factor A and B) and PGF (Placental Growth Factor) and plays an important role in inducing vasculogenesis and angiogenesis [40]. Having a miRNA that can affect both oncogenic and tumor suppressor genes makes it hard to predict the effects of these kinds of plant. There are many contributing factors to be considered: the transcription and translation rates of both genes, the threshold by which each protein can be functional in the body, how effectively the miRNA suppresses each transcript, etc. Having said that, folk Chinese manuscripts prescribed wormwood as an effective anticancer herb and there are also recent studies to prove the point [41,42]. *Artemisia annua* (aka wormwood) has an active ingredient known as artemisinin (ARS) which has long been known for the treatment of malaria, but there are also in-vivo and in-vitro studies that suggest its anti-cancer potential [43,44]. This finding could suggest other herbs in this group as candidate targets for exploring natural therapies against cancer. There is another study that has discussed the correlation between the reduced *FLT1* amounts and gestational diabetes mellitus [45], making it risky for pregnant women to have excessive amounts of these herbs.

One of the most interesting identified gene targets was *MYLK* which codes for myosin light-chain kinase. This calcium/calmodulin-dependent kinase phosphorylates myosin regulatory light chains to facilitate its interaction with actin, leading to the smooth muscle contractions. There is also a non-muscular isoform which participates in inflammatory responses, e.g. leukocyte diapedesis and apoptosis, cytoskeletal rearrangement, and airways hyperactivity which makes it a key element in causing asthma symptoms [46]. Purple yam, a less known herb that contains this miRNA, has been a folk remedy for respiratory diseases which can be justified by the anti-*MYLK* miRNA. Another aspect of *MYLK* function is how it can affect the cytoskeleton and the cell-cell interaction, which has made it a good target for cancer studies [47]. The fact that this protein is involved in such a broad range of body functions, makes it a curious target. The miRNA that targets this transcript is present in a great variety of plants as well: soybean, cocoa, apple, orange, sunflower, coffee and even tobacco and opium. Some of these plants have long been known as sedatives, which could be due to the *MYLK* inhibition. The role of *MYLK* in cancer progression seems to be controversial; there are studies suggesting that *MYLK* has a role in maintaining pro-survival signals in epithelial cells and its inhibition can lead to apoptosis of transformed cells. But there are more recent studies indicating that *MYLK* inactivation leads to decreased cell-cell adhesion and EGFR (Epidermal Growth Factor Receptor) signaling activation and hence facilitates breast epithelial cell migration. Wormwood is also present in this group and its anti-cancer properties have already been discussed.

Notably, a family member of myosin's most common accompanying protein was another target of the same miRNA. *ACTA2* codes for smooth muscle alpha 2 actin and it is mostly responsible for the vascular contractility and the blood pressure homeostasis. Ca-dependent apoptosis in T helper cells and GPCR (G Protein-Coupled Receptor) signaling pathway are the two most important pathways affected by *ACTA2*. There is a natural antisense RNA (*ACTA2-AS1*) on the same chromosome (Chromosome 10) to inhibit this transcript.

COPB2 (coatomer protein complex subunit beta2) was another transcript targeted by the same miRNA. This mRNA codes for a cytosolic protein associated with non-clathrin coated Golgi particles, so it is believed to be involved in the innate immune response and the endocytic recycling of LDL (Low-Density Lipoprotein) receptors. This can mark the *COPB2* influence on the lipid metabolism. The role of this protein in the cancer progression has recently been investigated as well. An siRNA study confirmed that the inhibition of *COPB2* had a positive effect on controlling prostate cancer. The results indicated that lack of this protein induced a decrease in the S-phase cell count and significantly increased the apoptotic cells number [48].

PKD1L2 (Polycystin 1 Like 2) is a member of the polycystin protein family which is believed to be a component of the cation channels, e.g. calcium. It may also be a G-protein receptor modulator since it regulates *GNAS* imprinted locus (Guanine Nucleotide Binding Protein Alpha Stimulating) and can affect MAPK (Mitogen-Activated Protein Kinase) pathway [49]. But the most important role of *PKD1L2* is interacting with the endogenous fatty acid synthase in the skeletal muscles [50]. Its role seems to be inhibitory since the up-regulation caused a neuromuscular disease in mice [51]. In this case, a dietary miRNA that inhibits excessive amounts of *PKD1L2* can provide a mean to study this disease. There is a curious fact: *PKD1L2* is a polymorphic gene in humans and some individuals have a non-functional (pseudogene) copy. These people can provide a great study group to assess the role of *PKD1L2* in neuromuscular disease susceptibility and how they react to a dietary regimen with foods containing this specific miRNA.

GIPR was another gene target of this group that is a G-protein coupled receptor for gastric inhibitory polypeptide, aka glucose-dependent insulinotropic Polypeptide. GIP is a hormone member of the incretin family which is secreted by the intestinal K cells. This hormone has two major roles: inhibiting the release of gastric acid and pepsin and also increasing insulin secretion from the beta cells of the islets of Langerhans in response to oral glucose consumption levels [52]. GIP also stimulates LipoProtein Lipase (LPL) activity and so it has been linked to obesity and diabetes. In a recent study, inhibition of GIPR signaling in the adipose tissue of mice proved to reduce insulin resistance and hepatic steatosis. Under the high-fat diet condition, *GIPR* knock-out mice showed lower body weight and plasma interleukin-6 levels, which led to less insulin resistance [53]. This study can provide a great hope for people dealing with diabetes; but since making knock-out human is not possible, using herbs and fruits that contain a miRNA for suppressing the transcript may be of a great use. According to our data in Table 1, these plants are amazingly versatile and easy to consume.

FAM83C (family with sequence similarity 83 member C) is a member of a family of 8 proteins (A to H) with different functions. The common feature of these members is the same domain of unknown function. But there has been a recent multi-omics study providing information on oncogenic properties of *FAM83* family and their association with increasing DNA copy numbers in different tumor types [54].

FAM83C may be involved in regulating MAPK signaling in cancer cells as well [55].

2 Target transcripts with unknown function, whether due to the lack of information on their protein or because they were antisense RNAs themselves:

ERICH3 is a transcript with no known function and no evidence in protein level. Based on the domain sequence, it has been suggested that ERICH3 has a physical association with BACH1, which is a transcription regulator protein [56]. BACH1 participates in cell cycle control in G1/M and G2/M transition [57], which would make *ERICH1* a probable cell cycle regulator. There are also genome-wide association studies (GWAS) indicating the role of *ERICH3* in serotonin reuptake and major depressive disorder (MDD). SNPs (Single Nucleotide Polymorphisms) across this gene affect the level of serotonin in the plasma and hence the therapeutic outcome of MDD patients [58].

DEPDC1-AS1 is an antisense RNA transcript for which no function has been suggested. *DEPDC1* (DEP Domain Containing 1) is not well known either, but it is said to be involved in a protein complex that acts as a transcriptional corepressor. In bladder carcinogenesis, DEPDC1A-ZNF224 complex leads to the suppression of the apoptosis in cancer cells, acting as an oncogenic agent [59]; the same mechanism has also been confirmed in prostate cancer through activation of E2F signaling, which regulates the cell cycle [60]. Considering these facts, *DEPDC1-AS1* may present a tumor suppression.

GRM7-AS2 has only been found in the placenta at low levels and at very higher levels in the testis. Unlike its antisense RNA, *GRM7* (Glutamate Metabotropic Receptor 7) is expressed in a broad range of tissues. Coding for a metabotropic glutamate receptor, the main function of *GMR7* is in neurotransmission in the brain and its down-regulation has been associated with Parkinson disease [61]. During the embryonic period, brain neurogenesis is regulated by *GMR7* via CREB (Cyclic AMP Response Element-Binding) and YAP (Yes-Associated Protein) signaling [62]. Yet more studies are needed to indicate why *GRM7-AS2* is not expressed as much in every other tissue.

CDKN2B-AS1 is also known as *ANRIL* (Antisense Noncoding RNA in the INK4 Locus) and its function is very well studied. This gene is encoded in the 9p21 chromosome region, within the *CDKN2B-CDKN2A* (Cyclin Dependent Kinase Inhibitor) gene cluster. This locus is a hotspot for polymorphisms in a variety of diseases such as type II diabetes and Alzheimer [63]. GWAS studies had also found a great association between this locus and different types of cancer, since ANRIL can regulate its neighbor tumor suppressor genes *CDKN2A/B* by epigenic mechanisms. The reason that *CDKN2B-AS1* have been categorized as a transcript with unknown function is that the exact role and effect of this lncRNA is not totally understood and does not show a clear pattern, as both oncogenic [64] and tumor suppressor effects [65] have been reported in different studies. Polymorphism of ANRIL gene have also been reported as cardiovascular risk alleles [64].

3 Transcripts with known sequences and positions for which no known gene was available:

The results of this study indicated that there were regions of chromosome 9 and 15 that are yet unknown but they seem to have active transcription. Our study suggests 2 dietary miRNA that can target these transcripts and help shed some light on these undiscovered regions of the human genome.

5. Conclusion

This study demonstrates that there are miRNAs in common in various species of plants which can be aligned with the human CDS sequences. The targeted mRNAs are involved in different body functions and metabolic pathways, therefore the nutritional plant miRNAs may be able to have an impact on the human body through altering gene regulation.

Conflicts of interest

None declared.

Financial support

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2019.01.020.

References

[1] D.M. Mutch, W. Wahli, G. Williamson, Nutrigenomics and nutrigenetics: the emerging faces of nutrition, FASEB J. 19 (12) (2005) 1602–1616.
[2] E.L. Beckett, et al., Nutrigenetics—personalized nutrition in the genetic age, Explor. Res. Hypothesis. Med. 2 (4) (2017) 1–8.
[3] M. Müller, S. Kersten, Nutrigenomics: goals and strategies, Nat. Rev. Genet. 4 (4) (2003) 315.
[4] L. Afman, M. Müller, Nutrigenomics: from molecular nutrition to prevention of disease, J. Acad. Nutr. Diet. 106 (4) (2006) 569–576.
[5] E.W. Ebomoyi, Nutrigenomics and nutrigenetics and the medicinal values of vegetables and fruits. Food Science and Nutrition: Breakthroughs in Research and Practice: Breakthroughs in Research and Practice, (2018), p. 52.
[6] B. Quintanilha, et al., Nutrimiromics: role of microRNAs and nutrition in modulating inflammation and chronic diseases, Nutrients 9 (11) (2017) 1168.
[7] B.P. Lewis, et al., Prediction of mammalian microRNA targets, Cell 115 (7) (2003) 787–798.
[8] L. Zhang, et al., Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA, Cell Res. 22 (1) (2012) 107–126.
[9] P. Sarkies, E.A. Miska, Is there social RNA? Science 341 (6145) (2013) 467–468.
[10] K. Ražná, et al., MicroRNA (miRNA) in food resources and medicinal plant, Potravinarstvo 10 (1) (2016) 188–194.
[11] B. Dickinson, et al., Lack of detectable oral bioavailability of plant microRNAs after feeding in mice, Nat. Biotechnol. 31 (11) (2013) 965.
[12] H. Huang, C.D. Davis, T.T. Wang, Extensive degradation and low bioavailability of orally consumed corn miRNAs in mice, Nutrients 10 (2) (2018) 215.
[13] H. See, et al., Targeted therapy for epithelial ovarian cancer: current status and future prospects, Int. J. Gynecol. Canc. 13 (6) (2003) 701–734.
[14] M. Yarmarkovich, K.D. Hirschi, Digesting dietary miRNA therapeutics, Oncotarget 6 (16) (2015) 13848.
[15] K.C. Vickers, et al., Corrigendum: MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins, Nat. Cell Biol. 17 (1) (2015) 104.
[16] A. Turchinovich, et al., Characterization of extracellular circulating microRNA, Nucleic Acids Res. 39 (16) (2011) 7223–7233.
[17] H. Iguchi, N. Kosaka, T. Ochiya, Secretory microRNAs as a versatile communication tool, Commun. Integr. Biol. 3 (5) (2010) 478–481.
[18] K.W. Witwer, K.D. Hirschi, Transfer and functional consequences of dietary microRNAs in vertebrates: concepts in search of corroboration, Bioessays 36 (4) (2014) 394–406.
[19] M.S. Duxbury, S.W. Ashley, E.E. Whang, RNA interference: a mammalian SID-1 homologue enhances siRNA uptake and gene silencing efficacy in human cells, Biochem. Biophys. Res. Commun. 331 (2) (2005) 459–463.
[20] K.W. Witwer, XenomiRs and miRNA homeostasis in health and disease: evidence that diet and dietary miRNAs directly and indirectly influence circulating miRNA profiles, RNA Biol. 9 (9) (2012) 1147–1154.
[21] E. Huntzinger, E. Izaurralde, Gene silencing by microRNAs: contributions of translational repression and mRNA decay, Nat. Rev. Genet. 12 (2) (2011) 99–110.
[22] R. Parker, H. Song, The enzymes and control of eukaryotic mRNA turnover, Nat. Struct. Mol. Biol. 11 (2) (2004) 121–127.
[23] J. Rehwinkel, et al., A crucial role for GW182 and the DCP1: DCP2 decapping complex in miRNA-mediated gene silencing, RNA 11 (11) (2005) 1640–1647.
[24] R.F. Place, et al., MicroRNA-373 induces expression of genes with complementary promoter sequences, Proc. Natl. Acad. Sci. Unit. States Am. 105 (5) (2008) 1608–1613.
[25] M.W. Jones-Rhoades, D.P. Bartel, Computational identification of plant microRNAs and their targets, including a stress-induced miRNA, Mol. Cell 14 (6) (2004) 787–799.
[26] M.W. Rhoades, et al., Prediction of plant microRNA targets, Cell 110 (4) (2002) 513–520.
[27] X. Yi, et al., PNRD: a plant non-coding RNA database, Nucleic Acids Res. 43 (D1) (2014) D982–D989.

[28] K.W. Witwer, et al., Real-time quantitative PCR and droplet digital PCR for plant miRNAs in mammalian blood provide little evidence for general uptake of dietary miRNAs: limited evidence for general uptake of dietary plant xenomiRs, RNA Biol. 10 (7) (2013) 1080–1086.

[29] W. Kang, et al., Survey of 800+ data sets from human tissue and body fluid reveals xenomiRs are likely artifacts, RNA 23 (4) (2017) 433–445.

[30] K.D. Hirschi, New foods for thought, Trends Plant Sci. 17 (3) (2012) 123–125.

[31] J. Yang, et al., Detection of dietary plant-based small RNAs in animals, Cell Res. 25 (4) (2015) 517.

[32] R. Ashby, et al., MicroRNAs in honey bee caste determination, Sci. Rep. 6 (2016) 18794.

[33] A.R. Chin, et al., Cross-kingdom inhibition of breast cancer growth by plant miR159, Cell Res. 26 (2) (2016) 217–228.

[34] S. Baier, MicroRNAs Are Absorbed in Biologically Meaningful Amounts from Nutritionally Relevant Doses of Cow's Milk and Chicken Eggs and Affect Gene Expression in Peripheral Blood Mononuclear Cells, Cell Cultures, and Mouse Livers, The University of Nebraska-Lincoln, 2015.

[35] S.R. Baier, et al., MicroRNAs are absorbed in biologically meaningful amounts from nutritionally relevant doses of cow milk and affect gene expression in peripheral blood mononuclear cells, HEK-293 kidney cell cultures, and mouse livers, J. Nutr. 144 (10) (2014) 1495–1500.

[36] M. Toth, A. Sohail, R. Fridman, Assessment of gelatinases (MMP-2 and MMP-9) by gelatin zymography, Metastasis Research Protocols, Springer, 2012, pp. 121–135.

[37] J.-P. Tassan, et al., Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C, Proc. Natl. Acad. Sci. Unit. States Am. 92 (19) (1995) 8871–8875.

[38] U. Ueberham, A. Hessel, T. Arendt, Cyclin C expression is involved in the pathogenesis of Alzheimer's disease, Neurobiol. Aging 24 (3) (2003) 427–435.

[39] J.H. Chung, et al., A PTCH1 homolog transcriptionally activated by p53 suppresses Hedgehog signaling, J. Biol. Chem. 289 (47) (2014) 33020–33031.

[40] T. André, et al., Vegf, vegf-B, vegf-C and their receptors KDR, FLT-1 and FLT-4 during the neoplastic progression of human colonic mucosa, Int. J. Canc. 86 (2) (2000) 174–181.

[41] H.S. Parekh, G. Liu, M.Q. Wei, A new dawn for the use of traditional Chinese medicine in cancer therapy, Mol. Canc. 8 (1) (2009) 21.

[42] T. Efferth, From ancient herb to modern drug: artemisia annua and artemisinin for cancer therapy, Seminars in Cancer Biology, Elsevier, 2017.

[43] N.P. Singh, H.C. Lai, Artemisinin induces apoptosis in human cancer cells, Anticancer Res. 24 (4) (2004) 2277–2280.

[44] N.P. Singh, H.C. Lai, Synergistic cytotoxicity of artemisinin and sodium butyrate on human cancer cells, Anticancer Res. 25 (6B) (2005) 4325–4331.

[45] F. Troncoso, et al., Gestational diabetes mellitus is associated with increased promigratory activation of vascular endothelial growth factor receptor 2 and reduced expression of vascular endothelial growth factor receptor 1, PLoS One 12 (8) (2017) e0182509.

[46] T. Wang, et al., A MYLK variant regulates asthmatic inflammation via alterations in mRNA secondary structure, Eur. J. Hum. Genet. 23 (6) (2015) 874.

[47] L.E. Connell, D.M. Helfman, Myosin light chain kinase plays a role in the regulation of epithelial cell survival, J. Cell Sci. 119 (11) (2006) 2269–2281.

[48] Y. Mi, et al., COPB2 is upregulated in prostate cancer and regulates PC-3 cell proliferation, cell cycle, and apoptosis, Arch. Med. Res. 47 (6) (2016) 411–418.

[49] T. Yuasa, et al., Polycystin-1L2 is a novel G-protein-binding protein, Genomics 84 (1) (2004) 126–138.

[50] H. Xu, The Functional Role of CD36 Involved in Fatty Acid Transduction, (2014).

[51] F.E. Mackenzie, et al., Upregulation of PKD1L2 provokes a complex neuromuscular disease in the mouse, Hum. Mol. Genet. 18 (19) (2009) 3553–3566.

[52] Y. Yamada, Y. Seino, Physiology of GIP-a lesson from GIP receptor knockout mice, Horm. Metab. Res. 36 (11/12) (2004) 771–774.

[53] E. Joo, et al., Inhibition of gastric inhibitory polypeptide receptor signaling in adipose tissue reduces insulin resistance and hepatic steatosis in high-fat diet–fed mice, Diabetes 66 (4) (2017) 868–879.

[54] A.M. Snijders, et al., FAM83 family oncogenes are broadly involved in human cancers: an integrative multi-omics approach, Mol. Oncol. 11 (2) (2017) 167–179.

[55] R. Cipriano, et al., Conserved oncogenic behavior of the FAM83 family regulates MAPK signaling in human cancer, Mol. Canc. Res. 12 (8) (2014) 1156–1165.

[56] M.Y. Hein, et al., A human interactome in three quantitative dimensions organized by stoichiometries and abundances, Cell 163 (3) (2015) 712–723.

[57] H.-J. Warnatz, et al., The BTB and CNC homology 1 (BACH1) target genes are involved in the oxidative stress response and in control of the cell cycle, J. Biol. Chem. 286 (26) (2011) 23521–23532.

[58] M. Gupta, et al., TSPAN5, ERICH3 and selective serotonin reuptake inhibitors in major depressive disorder: pharmacometabolomics-informed pharmacogenomics, Mol. Psychiatr. 21 (12) (2016) 1717.

[59] Y. Harada, et al., Cell-permeable peptide DEPDC1-ZNF224 interferes with transcriptional repression and oncogenicity in bladder cancer cells, Cancer Res. 70 (14) (2010) 5829–5839.

[60] L. Huang, et al., DEPDC1 promotes cell proliferation and tumor growth via activation of E2F signaling in prostate cancer, Biochem. Biophys. Res. Commun. 490 (3) (2017) 707–712.

[61] R. Foti, et al., Parkinson disease-associated DJ-1 is required for the expression of the glial cell line-derived neurotrophic factor receptor RET in human neuroblastoma cells, J. Biol. Chem. 285 (24) (2010) 18565–18574.

[62] W. Xia, Y. Liu, J. Jiao, GRM7 regulates embryonic neurogenesis via CREB and YAP, Stem Cell. Rep. 4 (5) (2015) 795–810.

[63] A. Congrains, et al., ANRIL: molecular mechanisms and implications in human health, Int. J. Mol. Sci. 14 (1) (2013) 1278–1292.

[64] K.L. Yap, et al., Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a, Mol. Cell 38 (5) (2010) 662–674.

[65] I. Yeh, B.C. Bastian, Genome-wide associations studies for melanoma and nevi, Pigm. Cell. Melanoma Res. 22 (5) (2009) 527.