TECHNICAL ADVANCE

Open Access

Elucidation of molecular links between obesity and cancer through microRNA regulation

Haluk Dogan[1], Jiang Shu[1], Zeynep Hakguder[1], Zheng Xu[2] and Juan Cui[1*]

Abstract

Background: Obesity contributes to high cancer risk in humans and the mechanistic links between these two pathologies are not yet understood. Recent emerging evidence has associated obesity and cancer with metabolic abnormalities and inflammation where microRNA regulation has a strong implication.

Methods: In this study, we have developed an integrated framework to unravel obesity-cancer linkage from a microRNA regulation perspective. Different from traditional means of identifying static microRNA targets based on sequence and structure properties, our approach focused on the discovery of context-dependent microRNA-mRNA interactions that are potentially associated with disease progression via large-scale genomic analysis. Specifically, a meta-regression analysis and the integration of multi-omics information from obesity and cancers were presented to investigate the microRNA regulation in a dynamic and systematic manner.

Results: Our analysis has identified a total number of 2,143 unique microRNA-gene interactions in obesity and seven types of cancer. Common interactions in obesity and obesity-associated cancers are found to regulate genes in key metabolic processes such as fatty acid and arachidonic acid metabolism and various signaling pathways related to cell growth and inflammation. Additionally, modulated co-regulations among microRNAs targeting the same functional processes were reflected through the analysis.

Conclusion: We demonstrated the statistical modeling of microRNA-mediated gene regulation can facilitate the association study between obesity and cancer. The entire framework provides a powerful tool to understand multifaceted gene regulation in complex human diseases that can be generalized in other biomedical applications.

Keywords: MicroRNA, Obesity, Cancer, MicroRNA regulation

Background

Obesity is becoming one of the leading public health crises globally and has posed serious threats to human health and life quality. According to the World Health Organization (WHO), by 2030, more than 1.2 billion people globally will suffer from obesity [1]. In the United States, 35.6 percent of adults are affected by obesity. Obesity is associated with an increased risk of several common human cancers and contributes to between 14 to 20 percent of cancer-related mortality cases, making it the leading preventable cause of cancer [2]. The underlying molecular mechanism of obesity development and relapse as well as the enhancing effects of obesity on cancer potency and progression remain largely unknown. In order to gain more insights on aforementioned unascertained mechanisms, there have been strengthened efforts to study the genetic basis of obesity, as well as the regulatory and epigenetic effects on signaling process.

*Correspondence: jcui@unl.edu
[1]Systems Biology and Biomedical Informatics Lab, Department of Computer Science and Engineering University of Nebraska-Lincoln, 68588-0115 Lincoln, NE, USA
Full list of author information is available at the end of the article

A group of small non-coding RNAs, called microRNA (miRNAs), are important gene silencers in animals and plants [3–5]. For example, over 2,000 human miRNAs can bind to over 60% endogenous genes [5] in humans and fine tune numerous biological processes related to cell growth and signaling [6]. Increasing evidence has shown strong implication of miRNA in complex human diseases including obesity (e.g., elevated abundance of miR-34 in obese mice reduces NAD+ levels and results in obesity-mimetic outcomes [7, 8]; miR-27, miR-519d, and miR-30c are obesity related [9–11]); metabolic disorder (e.g., miR-33a/b, miR-107, miR-103, and miR-34a regulate different metabolism [12]); pro-inflammatory events (e.g., miR-27b, miR-21, miR-34a, miR-106b, miR-130, miR-15b, miR-155, and miR-200c); aging (e.g., downregulation of miR-24, and miR-221 in the peripheral blood mononuclear cells (PBMCs) of older individuals, and increased expression other 21 miRNAs [13]), and cancers (e.g., miR-200 and miR-205 regulates epithelial to mesenchymal transition [14], miR-21 regulates apoptosis in lung cancer [15] and drives melanoma [16]). It should be noted that such linkages were determined largely based on the alterations of miRNA expression in those diseases or the effect of miRNA manipulation in loss-of-function and gain-of-function studies; only a handful were validated by elaborate target function in vivo.

More recently, studies began to unravel miRNA's role in the obesity-cancer link. For example, it has been demonstrated that the epigenetic silencing of miR-200c is capable of targeting STAT3-G9a signaling and limiting the malignancy of obesity-related breast cancer [17]. miR-10b was found down regulated in the breast tumors of the obese subjects compared to the lean; the suppression of miR-10b in breast cancer cell line increases cell proliferation and invasion and affects inversely the expression of targets BCL2L11, PIEZO1 and NCOR2 cell [18]. As reported in Kasiappan et al., [19] over 50 miRNAs show altered expressions in either obesity or breast cancer, and miR-302b and miR-498 are differentially expressed in obesity-associated breast cancer. Other examples include a list of miRNAs potentially linking obesity and colorectal cancer, namely miR-425, -196, -155, -150, -351, -16, -34, -148, -4443, -101, -27b, and let-7 [20]. Note that most such discoveries were made based on the expression analysis and the causality of the disease development is not fully validated. Many of the dysregulated miRNAs common in obesity and cancers were evidenced to regulate conditions such as insulin resistance, low-grade inflammation, cell proliferation and survival, dysfunction of adipose tissues, and increased oxidative stress through targeting VEGF, Ras, HIF1-α, PI3K/Akt, JAK/STAT3, MAPK, ERK/p38, and NF-kβ [19, 21–25]. All those events are important elements in obesity associated cancer development. In addition to endogenous regulation, there are growing interest in studying extracellular miRNAs associated with vesicles or protein complexes, e.g., adipose-derived circulating miRNAs, as important regulators in other tissues and diseases [26]. However, a grand challenge here is the lack of systems approaches and reliable high-throughput technologies that can assess miRNA roles in human disease progression in a holistic and dynamic fashion. In this regard, the development of sequencing technologies such as cross-linking, ligation, and sequencing of hybrids (CLASH) [27] and individual-nucleotide resolution CLIP (iCLIP) [28] allows the identification of the exact binding site of the miRNA on mRNAs, however, the drawbacks of the selection bias in base pairing, low coverage, and ambiguous downregulation still hinder the broad application in this research field.

It is known that miRNAs bind to the 3'-UTR of target genes to either trigger mRNA degradation or inhibit protein translation [29, 30]. The interactions between miRNA and genes exhibits a high level of complexity since most miRNAs can bind to multiple gene targets while the genes can be targeted by different miRNAs [31–34]. Additionally, the binding process has been found to be condition-specific, which means that the same miRNA may bind to different groups of mRNAs under different biological contexts. The underlying binding mechanism involves both competition (when the binding sites of different miRNAs are in close or overlapping regions on the target genes) and cooperation (when several miRNAs bind to different regions of the same target) [31–34]. This type of dynamic feature is hardly captured by early computational tools such as PicTar [32], MicroTar [35], miTarget [36], TargetRNA2 [37], TargetScan [38], and miRwalk [39], which utilize mainly sequence and structure characteristics to build the models. The second-generation target prediction tools integrate gene and/or miRNA expression profiles into the modeling process to reflect condition-specific mechanisms, including MiRonTop [40], mirAct [41], CoSMic [42]. Negatively correlated expressions between miRNA and their target genes are largely used as a strong indicator for real interaction in those methods. Considering the complicated machinery in gene expression regulation that involves transcription factor (TFs), methylation, and genetic variations, current research is exploring systems approaches that can take into consideration of these mechanisms underlying the multifaceted gene regulation and competitive miRNA binding.

In this article, we conducted a meta-analysis of miRNA regulation on both obesity and cancer to unravel the molecular association between these two types of disease. In order to investigate context-dependent miRNA-gene interactions, a new meta-regression method was presented to integrate large-scale genomics profiles in respective diseases, such as gene and miRNA expression,

copy number variation (CNV), and DNA methylation. The inferred interactions were compared with literature and other resources. For example, we further examined liver cancer interactions through analyzing our in-house gene expression data in a liver disease mouse model. We present the association between obesity and different types of cancers through identifying common condition-associated miRNA-mRNA interactions and characterizing regulatory transition across diseases.

Methods

In this section, a detailed description of the proposed computational framework is provided, including data collection and processing, binding potential estimation, and the regression-based model for miRNA-gene interaction identification. The schematic analytical workflow is shown in Fig. 1.

Genomic expression profiles

In this study, the genomic data were collected from the following resources:

1. GSE53378 and GSE25402 on a total of 42 obese patients from the Gene Expression Omnibus (GEO);
2. A collection of 3,484 patients with 7 major types of cancer (1,279 early stage patients) from The Cancer Genome Atlas (TCGA).

Names and download links of datasets used in our experiments are listed in Table 1. All cancer types were categorized into strong and weak association groups, SAG and WAG, respectively, according to their hazard ratios of the association with obesity [43]. Table 2 contains the detailed statistics of the obesity and cancer samples used in this study. For each sample, the expression profiles of gene and miRNA were extracted while information about
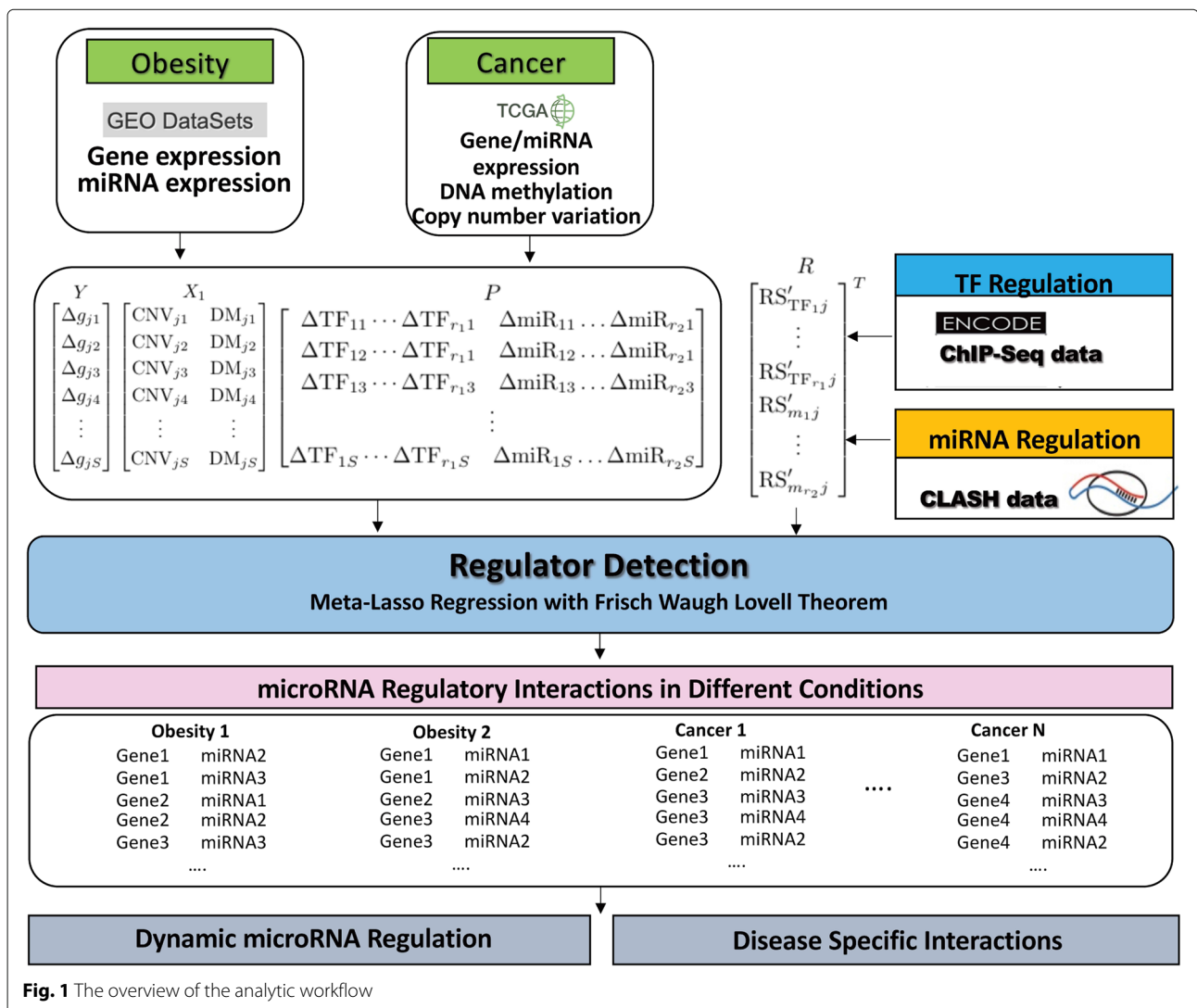

Fig. 1 The overview of the analytic workflow

Table 1 Names and download links of datasets used in our experiments

Dataset name	Download link
GSE53378	https://go.unl.edu/microarray-pre
GSE25402	https://go.unl.edu/microarray-obese
TCGA-BRCA	https://go.unl.edu/tcga-brca
TCGA-KIRC	https://go.unl.edu/tcga-kirc
TCGA-KIRP	https://go.unl.edu/tcga-kirp
TCGA-LIHC	https://go.unl.edu/tcga-lihc
TCGA-LUAD	https://go.unl.edu/tcga-luad
TCGA-LUSC	https://go.unl.edu/tcga-lusc
TCGA-UCEC	https://go.unl.edu/tcga-ucec
hg19	https://go.unl.edu/hg19
miRNA-mRNA interactions	https://go.unl.edu/clash

copy number variation and DNA methylation was also collected for each cancer sample. Note that the dataset includes both microarray and sequencing data from different disease conditions. In order to make the results meaningfully comparable, we make sure each data has its own control (either paired or unpaired with diseased samples), as well as matched miRNA and mRNA expression profiles from the same technology (either microarray or sequencing). Within each disease, normalized expression was used to calculate the expression change required for the downstream analysis.

Table 2 The detailed statistics of obesity and cancer datasets

	Obesity		Normal	Obese
	Adipose tissue (pre- and post-surgery)	OB1	16	16
	Adipose tissue (obese and lean)	OB2	30	26
	Cancer		**Normal**	**Cancer (Early Stage)**
S A G	Kidney renal clear cell carcinoma	KIRC	72	533 (251)
	Kidney renal papillary cell carcinoma	KIRP	32	290 (140)
	Liver hepatocellular carcinoma	LIHC	50	371 (164)
	Uterine corpus endometrial carcinoma	UCEC	24	176 (97)
W A G	Breast invasive carcinoma	BRCA	114	1097 (124)
	Lung adenocarcinoma	LUAD	59	515 (276)
	Lung squamous cell carcinoma	LUSC	51	502 (227)

Interactome profiles and data preprocessing

We downloaded the raw sequencing data of experimental miRNA-mRNA interactions from the CLASH analysis [27], as well as ChIP-Seq data on 150 Transcription factor (TFs) from the Encyclopedia of DNA Elements (ENCODE) Consortium [44].

First, the raw data from CLASH was re-mapped against human reference genome (hg19). Three types of reads were examined: 1) chimeric reads that contain miRNA-miRNA binding sites; 2) miRNA single reads that represent unbound miRNAs; 3) mRNA single reads that represent unbound mRNAs. After the read counts of each type were calculated, we utilized the binomial test to infer the confidence level of an identified binding site k between miRNA i and mRNA j with respect to a random observation. A p-value was calculated as follows:

$$p\text{-value}_{ijk} = \sum_{c=C_{ijk}}^{C_{\text{all}}} \binom{C_{\text{all}}}{c} \left(\frac{C_{ijk}}{C_{\text{all}}} \right)^c \left(1 - \frac{C_{ijk}}{C_{\text{all}}} \right)^{C_{\text{all}}-c}$$

where C_{all} denotes the total of supporting chimeric reads of all interactions; C_{ijk} is the count of supporting chimeric reads of interactions between miRNA i and mRNA j at the binding site k.

Binding sites with Bonferroni-multiple-test-adjusted p-value less than 0.05 were considered to be significant and included in the further analysis. For quality control, two additional filters were applied: 1) eliminating the bottom 10% of low confidence binding sites; 2) discarding the binding sites if the number of supporting chimeric reads is less than 12.

Estimation of regulatory score (RS)

As described above, we suspected that different regulators shared an equal possibility to interact with a common target. Therefore, we designed a procedure to qualitatively estimate the regulatory potentials between regulators and their targets.

RS of miRNA-gene interactions

First, we calculated the probability of a binding event between miRNA i and gene j at the binding site k through a conditional probability:

$$P_{ijk} = \frac{C_{ijk}^2}{(C_i + M_i)(C_j + M_j)}$$

where C_i and M_i are the total of chimeric reads and single reads associated with miRNA i, respectively; similarly, C_j and M_j indicate the total of interactions and free mRNA related to mRNA j, respectively.

Then, the RS of a miRNA-gene pair was calculated based on the aggregation of the binding probability and

the binding affinity, e.g., represented by the minimum free energy (MFE) [45] of the interaction, as follows:

$$RS_{ij} = \frac{1}{K} \sum_{k=1}^{K} |MFE_k| \times P_{ijk})$$

where MFE_k represents the minimum free energy of the binding site k and K is the total number of different binding sites between miRNA i and gene j.

Considering the competition among all possible targets of a miRNA regulator, RS was further normalized by the total number of interactions associated with miRNA i:

$$RS'_{ij} = \frac{RS_{ij}}{RS_i}$$

where RS_i denotes the sum of RS between miRNA i and all its target.

RS of TF-gene interactions
Unlike the RS of a miRNA-gene pair was calculated based on abundances, the RS for the TF t and gene j was estimated based on the distance, d, between the TF binding site k and the transcription start site (TSS) [46, 47], as follows:

$$RS_{tjk} = e^{-\left(0.5 + 4\left(\frac{d}{10^5}\right)\right)}$$

Then, the RSs for all binding sites between a TF and gene pair (t and j) can be aggregated by [46]:

$$RS_{tj} = 1 - \prod_{k=1}^{K} \left(1 - RS_{tjk}\right)$$

Similarly, considering that the same TF can bind to several different genes, RS was normalized as $RS'_{tj} = \frac{RS_{tj}}{RS_t}$ where RS_t denotes the sum of RS between TF t and all regulated genes.

The meta-regression model for gene regulation identification
In this study, each disease was considered as a specific condition D_t ($t = 1, 2, \ldots, 9$), which includes two conditions for obesity and seven cancers. For each condition D_t, we created four matrices for each gene as shown in Equations 1 and 2.

$$
\overset{Y}{\begin{bmatrix} \Delta g_{j1} \\ \Delta g_{j2} \\ \Delta g_{j3} \\ \Delta g_{j4} \\ \vdots \\ \Delta g_{jS} \end{bmatrix}}
\overset{X_1}{\begin{bmatrix} CNV_{j1} & DM_{j1} \\ CNV_{j2} & DM_{j2} \\ CNV_{j3} & DM_{j3} \\ CNV_{j4} & DM_{j4} \\ \vdots & \vdots \\ CNV_{jS} & DM_{jS} \end{bmatrix}}
\overset{R}{\begin{bmatrix} RS'_{TF_1 j} \\ \vdots \\ RS'_{TF_{r_1} j} \\ RS'_{m_1 j} \\ \vdots \\ RS'_{m_{r_2} j} \end{bmatrix}}^{T} \quad (1)
$$

$$
\overset{P}{\begin{bmatrix} P^1 \\ P^2 \\ P^3 \\ \vdots \\ P^S \end{bmatrix}}
\overset{P}{\begin{bmatrix} \Delta TF_{11} \cdots \Delta TF_{r_1 1} & \Delta miR_{11} \ldots \Delta miR_{r_2 1} \\ \Delta TF_{12} \cdots \Delta TF_{r_1 1} & \Delta miR_{12} \ldots \Delta miR_{r_2 1} \\ \Delta TF_{13} \cdots \Delta TF_{r_1 3} & \Delta miR_{13} \ldots \Delta miR_{r_2 3} \\ \vdots \\ \Delta TF_{1S} \cdots \Delta TF_{r_1 S} & \Delta miR_{1S} \ldots \Delta miR_{r_2 S} \end{bmatrix}} \quad (2)
$$

Expression matrix
$Y(S \times 1)$ contains the expression changes of gene j in each of the s samples associated with D_t compared to the control group. Here, the expression changes of gene j in sample s were calculated as follows:

$$\Delta g_{js} = \log_2 \frac{g_{js}}{g_{j\text{ctrl}}}$$

The background matrix
$X_1(S \times 2)$ includes values of the CNV and methylation change of the gene j in each sample versus the normal group. It must be noted that, in this study, the background matrix X_1 is only constructed for cancer conditions as CNV and DNA methylation information were not available for obesity dataset.

The regulator matrix
$P(S \times R)$ contains the expression changes of all TF and miRNA regulators of gene j in s samples versus the control group. R denotes the total number of regulators of gene j.

The regulatory score matrix
$R(1 \times R)$ contains the RS between gene j and its regulators (TFs and miRNAs). R denotes the total number of regulators of gene j.

We constructed a linear regression model [7] with a updated matrix X_2 ($[P^1 \times R, \ldots, P^S \times R]^T$, which are the products of R and each row in P) to represent the regulatory effects of TF and miRNA. Since the background matrix X_1 is not our focus, the Frisch–Waugh–Lovell (FWL) method was employed to transform the two-component regression model [7] to a standard linear regression model [8]. Through this linear transformation, the expression changes of gene (Y) were adjusted by M_{X_1}, which presents the influences of CNV and methylation on gene expression change. Last, the linear regression model

[8] was solved with variable selection by using Mallows's Cp method [48].

$$Y = X_1\beta_1 + X_2\beta_2 + \epsilon$$
$$M_{X_1}Y = M_{X_1}(X_2\beta_2 + \epsilon)$$

where $M_{X_1} = I - X_1(X_1^T X_1)^{-1} X_1^T$. The statistical significance of each regulator in the final model was assessed based on the consistency across all samples under the same condition. False discovery rate (Benjamini-Hochberg procedure, Q-value < 0.05) [49] was used to remove any insignificant regulators of corresponding gene.

Sampling for consensus regulator selection

To minimize the variation on regulator detection, we repeated the sampling process 20 times and generated 20 randomized datasets. Then, for each randomized dataset (U_i), we conducted regulator detection through the aforementioned model and obtained regulators of gene j (W_{ij}). On the other hand, the original dataset (U_{org}) was also applied on the regression-based model to obtain a set of regulator of gene j, $W_{org\text{-}j}$. The regulators which were in $W_{org\text{-}j}$ and have been consistently detected in all W_{ij}, are considered as the consensus regulators of gene j. As the final result, W_{con-j} that contains all consensus regulators of gene j is reported. For each reported interaction, a summarized coefficient and consensus measure represented by a percentage were given to reflect the selection consistency and confidence. The rationale is as follows. When regulatory interaction is formulated as a regression problem where target gene is a response variable and regulators are predictor variables, the regression coefficients, β, obtained after optimization indicate a relationship. Nonzero β values show existence of such relationship. Proportion of correctly estimated nonzero β values shows the sensitivity, and that of zero β values shows specificity. Without the ground truth of the relationships, calculation of sensitivity and specificity is not possible for method performance evaluation. Therefore, we define a consistency metric to show the confidence in our model predictions. Regression coefficients from Meta-lasso model was obtained for target gene and its regulators in each data split. Consistency is the proportion of nonzero regression coefficients in all the splits. We set a threshold at 0.7 for consistency values.

Over-sampling of the obesity dataset

Since our model considers all major regulators in this meta-regression analysis, it requires a minimal sample size for the reliable detection of gene regulators, which is challenging for the obesity datasets. To overcome this problem, an over-sampling process was implemented. The basic idea is to generate new samples by introducing random small variations to the existing samples in the original dataset to maintain the same sample distribution. As showed in [9], E_s is the whole genomic profile of one original sample s and V_σ is a randomly generated variation vector, which has the same dimension as E_s. The elements in variation vector V_σ were randomly sampled from 0.95 to 1.05. Then, the new sample E_s' was created as follows:

$$E_s' = E_s \times V_\sigma = [E_{s1}V_{\sigma 1}, E_{s2}V_{\sigma 2}, \ldots, E_{si}V_{\sigma i}, \ldots]$$

where $V_{\sigma i} \sim \text{Uniform}(0.95, 1)$

Subsequently, the same analysis was performed for every human gene g_j ($j = 1, 2, \ldots, 20531$) in obesity and seven cancers.

Results

First, through the aforementioned modeling and data analysis, we identified the differentially expressed (DE) genes and miRNAs (with Fold Change (FC) $>= 2$ and p-value < 0.05 through t-test) in each cancer and obesity, as well as the miRNA-gene interactions in each condition. The detailed statistics are summarized in (Supplementary Table 1) while the full list of interactions is at http://sbbi-panda.unl.edu:8001/obesity-regulators/.

Common DE genes and miRNA regulations reflect associations between obesity and cancers

To investigate the molecular association between obesity and different cancers, we first looked into the DE-genes and -miRNAs in each diseased condition versus the corresponding normal. In general, there are less DE-genes and -miRNAs in obesity than in cancers. For example, 405 genes and 9 miRNAs showed significant expression changes (FC $>= 2$) in obese group compared to normal while in average, there are 8,988 DE genes, and 29 DE miRNAs with FC $>= 2$ and 1,035 with FC >1 in cancers (5,425 DE-genes and 120 DE-miRNAs identified in the early stage of cancer).

Fewer DE-genes were identified in early stage cancers from SAG (Mean = 5,296, Confidence Interval = [4,242, 6,244]) than WAG (Mean = 5,599, Confidence Interval = [4,711, 6,948]). Surprisingly, even with fewer DE-genes, the SAG still consistently shares more common DE-genes with obesity than the WAG, i.e. [41.2%, 59.3%] versus [27.0%, 38.3%] as shown in Table 3. For example, 59.3% of DE-genes in obesity were also in early stage kidney cancer (KIRC), while only 27% of DE genes are common between obesity and early stage lung cancer (LUSC).

It has been known that not many miRNAs change in the adipose cells in obesity. As shown in our analysis, only 9 miRNAs were found differentially expressed in obesity, with FC $>= 2$. There is no evidence suggesting the significant commonality between SAG and obesity in terms of DE-miRNAs. Note that those reported in obesity-cancer linkage shows consistent expression patterns, e.g., let-7

Table 3 The Overview of Commonality Between Obesity and Cancers (A, B% denotes the number of overlapped genes/miRNAs/interactions between obese and cancer, and % of obese is overlapped with the corresponding cancer)

		Obesity	Strong association group				Weak association group		
			KIRC	KIRP	LIHC	UCEC	BRCA	LUAD	LUSC
Hazard ratio (HR) [32]		-	1.25		1.19	1.62	0.97	0.82	
Differential expression	Genes	405	240,	205,	168,	167,	130,	155,	109,
			59.3%	50.6%	41.5%	41.2%	32.1%	38.3%	27.0%
	miRNAs	9	2,	3,	1,	3,	2,	2,	2,
			22.2%	33.3%	11.1%	33.3%	22.2%	22.2%	22.2%
MicroRNA regulatory interactions	Interactions	1558	523,	556,	497,	568,	492,	534,	507,
			31.6%	33.6%	31.9%	36.6%	31.6%	34.3%	32.5%
	Genes	1355	806,	845,	776,	809,	770,	846,	848,
			59.5%	62.4%	57.3%	59.7%	56.8%	62.4%	62.6%
	miRNAs	180	110,	115,	115,	112,	111,	108,	107,
			61.1%	63.9%	63.9%	62.2%	61.7%	60.0%	59.4%

and miR-10b are down regulated in obesity and breast cancer). However, considering the fact that different miRNAs may co-regulate the same gene targets and lead to the same crucial functional changes, it is highly warranted to investigate the obesity-cancer associated in terms of miRNA regulatory interactions.

By examining the miRNA-gene interactions detected in each disease, we found that miRNA regulations are highly involved in both obesity and early stage cancers (consistent across all types of cancers). About 92% (out of 1,558) miRNA-gene interactions in obesity were detected in at least one type of cancer as shown in Table 3. More than half of miRNA-mediated genes in obesity were also targets of miRNAs in cancers. Observations also suggest that the same miRNAs may interact with different gene targets under different disease conditions.

miRNA regulated pathways in obesity and early-stage SAG cancers

To demonstrate that miRNA regulation is involved in the link between obesity and cancers, particularly at the early stage, we conducted gene set enrichment analysis [50] to identify the altered functional pathways based on miRNA targets in each dataset. As shown in Table 4, we listed the top five enriched pathways from three major functional categories that are critical in obesity and cancers, such as signaling pathways, metabolism pathways and inflammation pathways. We found that the patterns in some particular pathways can differentiate the SAG and WAG. For example, MAPK signaling pathway was consistently enriched in down-regulated genes in three cancers of WAG while being enriched in up-regulated genes in obesity, kidney and liver cancers. Similarly, two pathways related to fatty acid metabolism and glycine, serine and threonine metabolism also showed common

alterations among obesity, kidney cancer, and liver cancer. With respect to immune response, obesity showed higher consistency with SAG than WAG.

miRNAs co-regulate the same biological process by targeting different genes in each disease

Next, we compared obesity with early-stage SAG cancers and identified 31 common miRNA-gene interactions and 9 common critical pathways, as shown in Table 5. Taking fatty acid metabolism as an example, 8 miRNAs and 7 genes are involved in obesity as shown in the middle panel of Fig. 2. Between obesity and early-stage liver cancer, upper panel of Fig. 2, two interactions (miR-152/ECHS1 and miR-100/ALDH9A1) remained the same (in red) and another two common miRNAs (miR-193b and let-7b) regulate different gene targets in LIHC compared to obesity. Specifically, miR-193b regulates ALDH3A2 and let-7b interacts with ALDH7A1 in obesity, however, in liver cancer, miR-193b regulates ALDH7A1 and let-7b targets HADHA, they regulate different genes (miR-193b /ALDH7A1, let-7b/HADHA) that participate in the same functional process. When looking at the fatty acid metabolism in obesity and uterine cancer, three miRNA-gene interactions (miR100/ALDH9A1, miR-186/ACAA2 and miR- 193b/ALDH3A2) were shared in common. In addition, another 5 common genes and 4 miRNAs were involved with different interactions. For example, miR-615-3p interacts with CPT1C in obesity and regulates ACOX1 and ALDH1B1 in uterine cancer.

When including all cancer samples regardless of stage into the comparison, 1,147 common miRNA-gene interactions and 51 common critical pathways were identified (Supplementary Table 2). Taking two signaling pathways as examples: 193, 19, and 4 miRNAs regulate MAPK pathways in obesity, liver cancer and uterine cancer,

Table 4 Functional pathways enriched in obesity and early stage cancers in the selected categories (↑ and ↓ indicate that the pathways were identified among the up-regulated and down-regulated genes, respectively, in the corresponding condition; those in bold show consistent patterns between obesity and SAG cancers)

Pathway types	Pathways	Early stage cancers							Obesity
		KIRC	KIRP	LIHC	UCEC	BRCA	LUAD	LUSC	
Signaling Pathways	MAPK signaling pathway	↑	↑	↑		↓	↓	↓	↑
	p53 signaling pathway	↑	↑	↑	↑	↑	↑	↑	
	Calcium signaling pathway	↑	↑	↑	↓	↓	↓	↓	
	PPAR signaling pathway		↓	↑		↑			↓
	Insulin signaling pathway	↑	↑	↓	↑	↓		↑	↓
Metabolism Pathways	Linoleic acid metabolism	↑	↑	↓	↑				
	Fatty acid metabolism	↓	↓	↓		↓			↓
	Glycine, serine and threonine metabolism	↓	↓	↓	↑		↑	↑	↓
	Arachidonic acid metabolism	↓	↓	↓	↑		↓	↓	↓
	Glutathione metabolism	↓			↑		↑	↑	↑
Inflammation Pathways	Cytokine-cytokine receptor interaction	↑	↑	↑		↑	↓	↓	↑
	B cell receptor signaling pathway	↑	↑	↑				↓	
	T cell receptor signaling pathway	↑	↑	↑	↑		↓		↑
	mTOR signaling pathway	↑				↑			↑

respectively. As shown in Fig. 3A, miR-423-5p and miR-484 are the common regulator in all three diseases by targeting overlapped sets of genes (including NFKB2, GNG12, MKNK2, STMN1, FLNA) in these diseases. Similarly, in PI3K-Akt signaling pathway (Fig. 3B), miR-484 and miR-769-5p co-regulate different genes in three diseases. All these observations stressed again the modulated property of miRNA regulation as different miRNAs can regulate the similar processes by targeting the same or different genes. An expanded list of pathways revealed by the common interactions is shown in Fig. 4.

Additionally, 10 interactions were involved in the non-alcoholic fatty liver disease (NAFLD) pathway, including miR-423-3p, miR-484, miR-342-3p, miR-146b-5p, miR-193a-5p, miR-324-3p and their 10 targets. We used a set of in-house (small) RNA-Seq data generated on a liver disease mouse model (details can be found in Supplementary Materials) and examined the expression patterns related to these interactions. We found that in the mouse liver samples collected at 20 wks (with STZ+HFD induced tumor), the ITCH/miR-324-3p interaction shows negatively correlated expression between miRNA and gene (|Pearson correlation coefficient (PCC)| ≥ 0.7) while RXRA/miR-423-3p, JUN/miR-342-3p, and AKT2/miR-423-3p show weak negative correlation. Similarly, among all 721 LIHC interactions identified in the TCGA data,

328 of them show similar negatively correlated expression (Supplementary Table 3). This coupled result provides clear indication of miRNA participation in cell proliferation, differentiation, and inflammatory signaling during the progression of non-alcoholic fatty liver disease to cancer.

Discussion

In this study, we identified miRNA-gene interactions in obesity and seven major types of cancer by integrating multi-level genomic information through computational models. A list of common miRNA regulators are highly likely to be involved in the development of obesity-associated cancers in terms of growth and inflammatory signaling and metabolism. Strong evidences show that most miRNAs contribute in the same functional pathway through regulating different genes under different conditions, which indicates that miRNA regulation is a function-driven dynamic process.

We also observe the patterns among the altered functional pathways which are promising to differentiate two groups of obesity related cancers, particularly in the early stage. Note that when including all cancers regardless of stage, due to the much higher variability among the expression profiles, the detected interactions and enriched pathways are slightly different from the early

Table 5 Common miRNA Regulatory Interactions and Enriched Pathways Between Obesity and Cancer of SAG

	Strong association group
Common obesity-detected miRNA regulatory interactions	miR-320a – MLL/OTUD1/PPIAL4G/PPP1R7;
	miR-18a – CSDA/IRS4/KLHDC10;
	miR-193b – GMPR2/MYST3/ZNF71;
	miR-484 – FAM128B/NOMO3/TMEM93;
	miR-615-3p – KIAA0494/LRRC37A2/MRPL23;
	miR-17 – CBL/TRAP1;
	let-7b – SPTBN2; let-7e – EIF2C2;
	miR-106a – HDHD1A; miR-125b – GPRIN1;
	miR-149 – JMJD5; miR-186 – INPP5F;
	miR-192 – KIAA1671; miR-378 – MLF1;
	miR-935 – TALDO1; miR-625 – LASS2;
	miR-30e – EXO1; miR-671-5p – CCDC21;
	miR-1301 – ZCCHC24
Enriched pathways through miRNA regulations	Fatty Acid Metabolism
	Arachidonic Acid Metabolism
	Purine Metabolism
	MAPK Signaling Pathway
	P53 Signaling Pathway
	Axon Guidance
	Focal Adhesion
	Gap Junction
	Leukocyte Transendothelial Migration

stage cases. This is a common problem in disease-related expression analysis given the fact that many genes and miRNAs show differential expression in different disease stages in addition to the variability related to gender, age, and other factors. Based on the identified conditional miRNA-gene interactions in each disease, we observed that miRNAs regulate a few common signaling and metabolic pathways between obesity and cancers, which may imply the miR-driven linkage between these diseases. Obesity comes from a chronic imbalance between energy intake and energy expenditure, which results in changes that lead to abnormal growth. The cellular location of these changes, which potentially includes the regulation, signaling, and genomic and epigenomic systems, are reflected at the metabolic level. Therefore, metabolic abnormality as one of major hallmark of cancer also reflects the early change involved in the disease development. We believe by focusing on the metabolic aspect of disease progression, we can further study how external factors affect disease progression and understand the association between obesity and cancers at the systems level where miRNA regulation represents a key mechanism in terms of signaling from adipose tissues to remote tissues. We agree that experimental developments along the research line of miRNA tracking and targeting can help better understand miRNA's involvement in cell communication and disease progression.

It is noted that miRNA transfer and cross-talk is beyond the scope of this computational study due to the difficulty in stratifying miRNAs according to their origins. There are well-known technical challenges in designing experimental protocols to track miRNA secretion and isolate exogenous regulation in target tissues.

From a technical perspective, we have developed a new integrated framework to study dynamic miRNA-mediated regulation in human diseases and have demonstrated that such data-driven approaches with novel solutions to information fusion and computational modeling can effectively facilitate novel mechanistic discoveries and hypothesis testing in biomedical research. The comparison with other methods of the same kind has demonstrated advantageous performance of this methodology (Supplementary materials). In the meantime, we are also well aware of the challenges in data integration from heterogeneous resources. For instance, miRNA-mRNA interactions are deemed to be context-dependent and somewhat cell specific. Ideally, various levels of genetic and genomic data should be collected from the same context to avoid biased discovery. In this study, although all common interactions inferred in obesity and cancers can be validated in miRecords and mirTarBase, this prediction doesn't include interactions known in obesity and cancer linkage. This is mainly because some miRNAs (e.g., miR-10b, 302b, or -498 in obesity-breast cancer case) were not initially covered in the CLASH interactome data. This issue can be possibly addressed by compiling more complete interactome available in different cell types. For the same reason, it is also problematic to combine or compare different studies when variant context are presented, which may partially explain the current inconsistency of miRNA regulators reported in a disease by different studies in the literature. However, a higher level of conservation in terms of miRNA regulated pathways is expected in those scenarios. Furthermore, existing interactome detected in very few cell types share high level of commonality but as a whole, it is still infeasible to cover all the possible interaction patterns. To this end, future studies with well-thought out experimental design and more systematic data generation can dramatically improve the capacity of similar computational models.

Conclusion

In this paper, we have examined the association between obesity and obesity-associated cancer through studying the miRNA regulation. Particularly, a novel statistical method has been developed to discover the context-dependent miRNA-gene regulation and identify key miRNA regulators
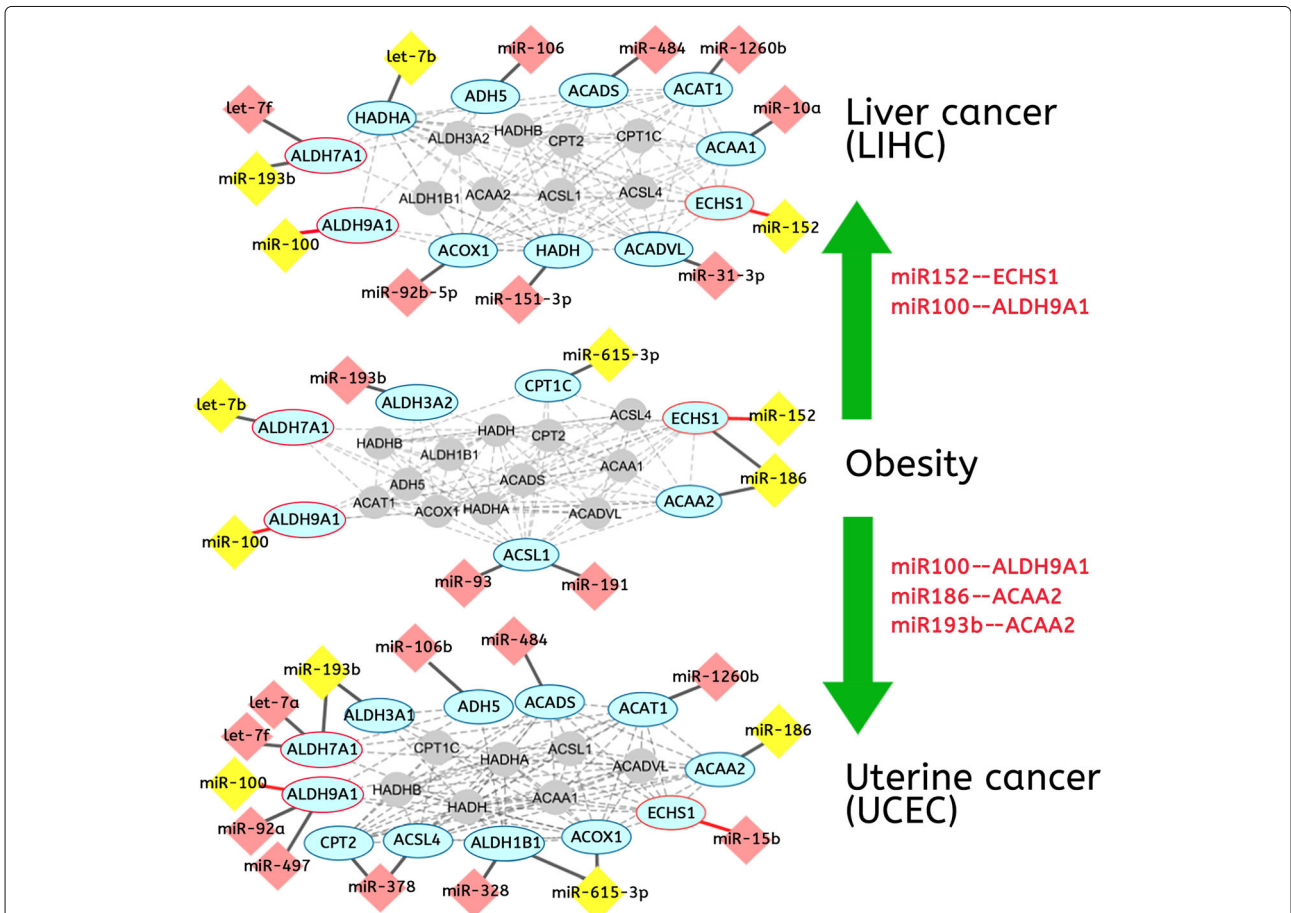
Fig. 2 Dynamic regulation of miRNA in fatty acid metabolism in obesity, early-stage liver and uterine cancers
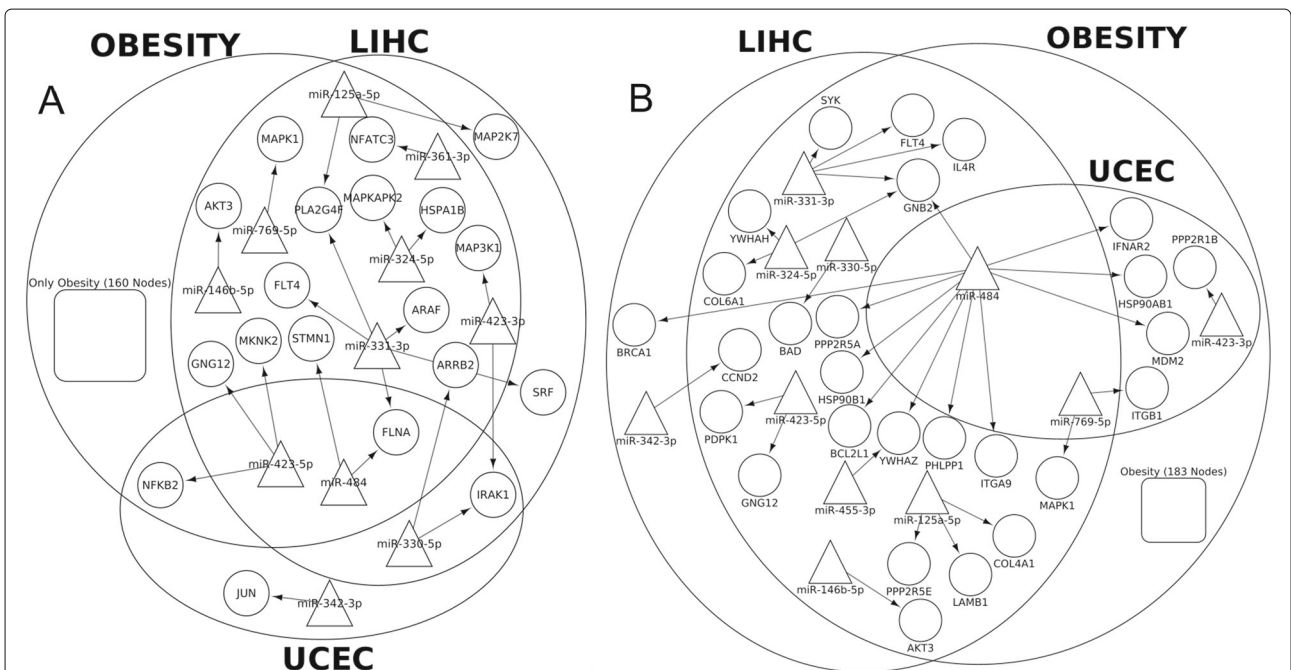

Fig. 3 MiRNA regulation in signaling processes in obesity, liver and uterine cancers. **a**) MAPK signaling pathway and **b**) PI3K-Akt signaling pathway
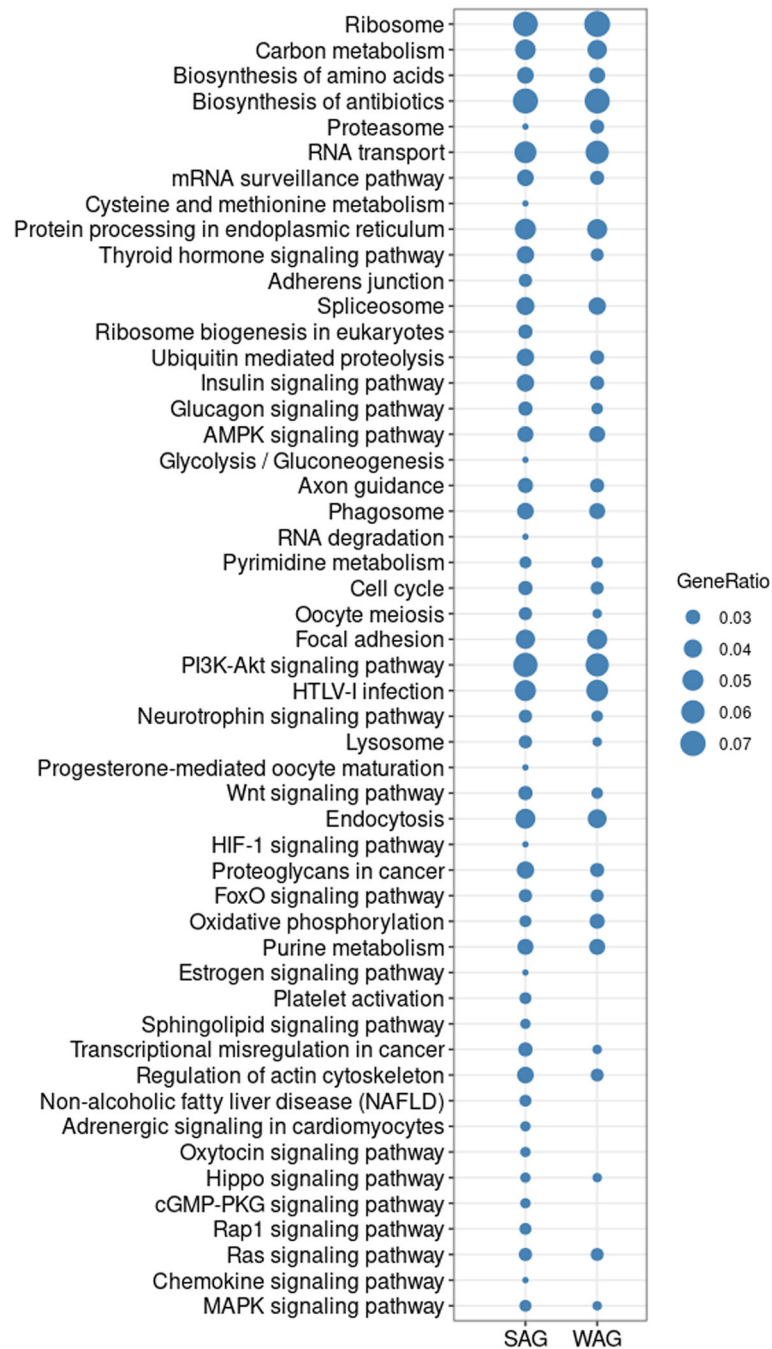
Fig. 4 Other pathways involved in miRNA regulation in SAG and WAG cancers

that are involved in the development of obesity-associated cancer. The entire framework for the detection of condition-dependent miRNA-mRNA interaction can be integrated into general dynamic gene regulation network study and be applied in similar biomedical research.

Supplementary information

Additional file 1: Supplementary Materials.

Author details
[1]Systems Biology and Biomedical Informatics Lab, Department of Computer Science and Engineering University of Nebraska-Lincoln, 68588-0115 Lincoln, NE, USA. [2]Department of Mathematics and Statistics, Wright State University, 45435 Dayton, OH, USA.

References
1. Kelly T, Yang W, Chen CS, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. Int J Obes. 2008;32(9):1431. https://doi.org/10.1038/ijo.2008.102.
2. Ligibel JA, Alfano CM, Courneya KS, Demark-Wahnefried W, Burger RA, Chlebowski RT, et al. American Society of Clinical Oncology Position Statement on Obesity and Cancer. J Clin Oncol. 2014;32(31):3568–3574. https://doi.org/10.1200/JCO.2014.58.4680.
3. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA Translation and Stability by microRNAs. Annu Rev Biochem. 2010;79(1):351–379. https://doi.org/10.1146/annurev-biochem-060308-103103.
4. Bartel DP. MicroRNAs: Target Recognition and Regulatory Functions. Cell. 2009;136(2):215–233. http://www.sciencedirect.com/science/article/pii/S0092867409000087.
5. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2013;11(42(D1)):D68–D73. https://doi.org/10.1093/nar/gkt1181.
6. Neilson JR, Sharp PA. Small RNA Regulators of Gene Expression. Cell. 2008;134(6):899–902. http://www.sciencedirect.com/science/article/pii/S0092867408011252.
7. Lee J, Padhye A, Sharma A, Song G, Miao J, Mo YY, et al. A pathway involving farnesoid X receptor and small heterodimer partner positively regulates hepatic sirtuin 1 levels via microRNA-34a inhibition. J Biol Chem. 2010;285(17):12604–12611. https://www.ncbi.nlm.nih.gov/pubmed/20185821.
8. Choi SE, Fu T, Seok S, Kim DH, Yu E, Lee KW, et al. Elevated microRNA-34a in obesity reduces NAD+ levels and SIRT1 activity by directly targeting NAMPT. Aging Cell. 2013;12(6):1062–1072. https://onlinelibrary.wiley.com/doi/abs/10.1111/acel.12135.
9. McGregor RA, Choi MS. microRNAs in the Regulation of Adipogenesis and Obesity. Curr Mol Med. 2011;11(4):304–316. http://www.eurekaselect.com/node/74129/article.
10. Karbiener M, Fischer C, Nowitsch S, Opriessnig P, Papak C, Ailhaud G, et al. microRNA miR-27b impairs human adipocyte differentiation and targets PPARγ. Biochem Biophys Res Commun. 2009;390(2):247–251. http://www.sciencedirect.com/science/article/pii/S0006291X09019159.
11. Karbiener M, Neuhold C, Opriessnig P, Prokesch A, Bogner-Strauss JG, Scheideler M. MicroRNA-30c promotes human adipocyte differentiation and co-represses PAI-1 and ALK2. RNA Biology. 2011;8(5):850–860. https://doi.org/10.4161/rna.8.5.16153.
12. Rottiers V, Näär AM. MicroRNAs in metabolism and metabolic disorders. Nat Rev Mol Cell Biol. 2012;13(4):239. https://doi.org/10.1038/nrm3313.
13. Hooten NN, Abdelmohsen K, Gorospe M, Ejiogu N, Zonderman AB, Evans MK. microRNA Expression Patterns Reveal Differential Expression of Target Genes with Age. PLoS ONE. 2010;5(5):e10724.
14. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol. 2008;10(5):593. https://doi.org/10.1038/ncb1722.
15. Seike M, Goto A, Okano T, Bowman ED, Schetter AJ, Horikawa I, et al. MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers. Proc Natl Acad Sci. 2009;106(29):12085–12090. https://www.pnas.org/content/106/29/12085.
16. Melnik BC. MiR-21: an environmental driver of malignant melanoma?. J Transl Med. 2015;13(1):202. https://doi.org/10.1186/s12967-015-0570-5.
17. Krakowsky RH, Tollefsbol TO. Impact of Nutrition on Non-Coding RNA Epigenetics in Breast and Gynecological Cancer. Front Nutr. 2015;2:16. https://www.frontiersin.org/article/10.3389/fnut.2015.00016.
18. Meerson A, Eliraz Y, Yehuda H, Knight B, Crundwell M, Ferguson D, et al. Obesity impacts the regulation of miR-10b and its targets in primary breast tumors. BMC Cancer. 2019;19(1):. https://doi.org/10.1186/s12885-019-5300-6.
19. Kasiappan R, Rajarajan D. Role of MicroRNA Regulation in Obesity-Associated Breast Cancer: Nutritional Perspectives. Adv Nutr. 2017;8(6):868–888. https://doi.org/10.3945/an.117.015800.
20. Ayers D, Boughanem H, Macías-González M. Epigenetic Influences in the Obesity/Colorectal Cancer Axis: A Novel Theragnostic Avenue. J Oncol. 2019;2019:1–10. https://doi.org/10.1155/2019/7406078.
21. Ande S, Nguyen K, Padilla-Meier G, Nyomba B, Mishra S. Expression of a mutant prohibitin from the aP2 gene promoter leads to obesity-linked tumor development in insulin resistance-dependent manner. Oncogene. 2016;35(34):4459. https://doi.org/10.1038/onc.2015.501.
22. Crisóstomo J, Matafome P, Santos-Silva D, Gomes AL, Gomes M, Patrício M, et al. Hyperresistinemia and metabolic dysregulation: a risky crosstalk in obese breast cancer. Endocrine. 2016;53(2):433–442. https://doi.org/10.1007/s12020-016-0893-x.
23. Pietrzyk L, Torres A, Maciejewski R, Torres K. Obesity and Obese-related Chronic Low-grade Inflammation in Promotion of Colorectal Cancer Development. Asian Pac J Cancer Prev. 2015;16(10):4161–4168.
24. Yang X, Jiang S, Wang G, Zhang R, Zhang J, Zhu J. Link of obesity and gastrointestinal cancer: crossroad of inflammation and oxidative stress. J Biol Regul Homeost Agents. 2015;29(4):755–760.

25. Marseglia L, Manti S, D'Angelo G, Nicotera A, Parisi E, Di Rosa G, et al. Oxidative stress in obesity: a critical component in human diseases. Int J Mol Sci. 2015;16(1):378–400.

26. Thomou T, Mori MA, Dreyfuss JM, Konishi M, Sakaguchi M, Wolfrum C, et al. Adipose-derived circulating miRNAs regulate gene expression in other tissues. Nature. 2017;542(7642):450–455.

27. Helwak A, Kudla G, Dudnakova T, Tollervey D. Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. Cell. 2013;153(3):654–665. http://www.sciencedirect.com/science/article/pii/S009286741300439X.

28. Broughton JP, Lovci MT, Huang JL, Yeo GW, Pasquinelli AE. Pairing beyond the Seed Supports MicroRNA Targeting Specificity. Molecular Cell. 2016;64(2):320–333.

29. Bagga S, Bracht J, Hunter S, Massirer K, Holtz J, Eachus R, et al. Regulation by let-7 and lin-4 miRNAs Results in Target mRNA Degradation. Cell. 2005;122(4):553–563. http://www.sciencedirect.com/science/article/pii/S0092867405008019.

30. Czech B, Hannon GJ. Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet. 2010;12(1):19–31.

31. Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell. 2004;116(2):281–297. http://www.sciencedirect.com/science/article/pii/S0092867404000455.

32. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, et al. Combinatorial microRNA target predictions. Nature Genetics. 2005;37(5):495. https://doi.org/10.1038/ng1536.

33. Cannell IG, Kong YW, Bushell M. How do microRNAs regulate gene expression?. Biochem Soc Trans. 2008;36(6):1224–1231. http://www.biochemsoctrans.org/content/36/6/1224.

34. Seitz H. Redefining MicroRNA Targets. Current Biology. 2009;19(10):870–873. http://www.sciencedirect.com/science/article/pii/S0960982209009130.

35. Thadani R, Tammi MT. MicroTar: predicting microRNA targets from RNA duplexes. BMC Bioinformatics. 2006;7:(S5).

36. Kim SK, Nam JW, Rhee JK, Lee WJ, Zhang BT. miTarget: microRNA target gene prediction using a support vector machine. BMC Bioinformatics. 2006;7(1):411.

37. Kery MB, Feldman M, Livny J, Tjaden B. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. Nucleic Acids Res. 2014;42((W1)):W124–W129.

38. Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. eLife. 2015;4:e05005.

39. Sticht C, Torre CDL, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. PLOS ONE. 2018;13(10):e0206239.

40. Brigand KL, Robbe-Sermesant K, Mari B, Barbry P. MiRonTop: mining microRNAs targets across large scale gene expression studies. Bioinformatics. 2010;26(24):3131–3132.

41. Liang Z, Zhou H, He Z, Zheng H, Wu J. mirAct: a web tool for evaluating microRNA activity based on gene expression data. Nucleic Acids Res. 2011;39((suppl_2)):W139–W144.

42. Ben-Moshe NB, Avraham R, Kedmi M, Zeisel A, Yitzhaky A, Yarden Y, et al. Context-specific microRNA analysis: identification of functional microRNAs and their mRNA targets. Nucleic Acids Res. 2012;40(21):10614–10627.

43. Bhaskaran K, Douglas I, Forbes H, dos Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. The Lancet. 2014;384(9945):755–765. http://www.sciencedirect.com/science/article/pii/S0140673614608928.

44. Consortium EP, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57. https://doi.org/10.1038/nature11247.

45. Rehmsmeier M, Steffen P, Höchsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. RNA. 2004;10(10):1507–1517. http://rnajournal.cshlp.org/content/10/10/1507.abstract.

46. Jiang P, Freedman ML, Liu JS, Liu XS. Inference of transcriptional regulation in cancers. Proc Natl Acad Sci. 2015;112(25):7731–7736. https://www.pnas.org/content/112/25/7731.

47. Wang S, Sun H, Ma J, Zang C, Wang C, Wang J, et al. Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nat Protoc. 2013;8(12):2502. https://www.ncbi.nlm.nih.gov/pubmed/24263090.

48. Ronchetti E, Staudte RG. A Robust Version of Mallows's C_p. J Am Stat Assoc. 1994;89(426):550–559. http://www.jstor.org/stable/2290858.

49. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc. 1995;57(1):289–300. http://www.jstor.org/stable/2346101.

50. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci. 2005;102(43):15545–15550. https://www.pnas.org/content/102/43/15545.

Publisher's Note