ncRNA2MetS: a manually curated database for non-coding RNAs associated with metabolic syndrome

Dengju Yao [Corresp., 1, 2, 3], **Xiaojuan Zhan** [4, 5], **Xiaorong Zhan** [6], **Chee Keong Kwoh** [2], **Yuezhongyi Sun** [1, 5]

[1] School of Software and Microelectronics, Harbin University of Science and Technology, Harbin, Heilongjiang, China

[2] School of Computer Science and Engineering, Nanyang Technological University, Singapore, Singapore

[3] College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, Heilongjiang, China

[4] College of Computer Science and Technology, Heilongjiang Institute of Technology, Harbin, Heilongjiang, China

[5] School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, Heilongjiang, China

[6] Department of Endocrinology and Metabolism, the First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

Corresponding Author: Dengju Yao
Email address: ydkvictory@hrbust.edu.cn

Metabolic syndrome is a cluster of the most dangerous heart attack risk factors (diabetes and raised fasting plasma glucose, abdominal obesity, high cholesterol and high blood pressure), and has become a major global threat to human health. A number of studies have demonstrated that hundreds of non-coding RNAs, including miRNAs and lncRNAs, are involved in metabolic syndrome-related diseases such as obesity, type 2 diabetes mellitus, hypertension etc. However, these research results are distributed in a large number of literature, which is not conducive to analysis and use. There is an urgent need to integrate these relationship data between metabolic syndrome and non-coding RNA into a specialized database. To address this need, we developed a metabolic syndrome-associated non-coding RNA database (ncRNA2MetS) to curate the associations between metabolic syndrome and non-coding RNA. Currently, ncRNA2MetS contains 1068 associations between five metabolic syndrome traits and 627 non-coding RNAs (543 miRNAs and 84 lncRNAs) in four species. Each record in ncRNA2MetS database represents a pair of disease-miRNA (lncRNA) association consisting of non-coding RNA category, miRNA (lncRNA) name, name of metabolic syndrome trait, expressive patterns of non-coding RNA, method for validation, specie involved, a brief introduction to the association, the article referenced, etc. We also developed a user-friendly website so that users can access and download all data easily. In short, ncRNA2MetS is a complete and high-quality data resource for exploring the role of non-coding RNA in the pathogenesis of metabolic syndrome and seeking new treatment options. The website is freely available at http://www.biomed-bigdata.com:50020/index.html

ncRNA2MetS: a manually curated database for non-coding RNAs associated with metabolic syndrome

Dengju Yao[1, 2, 3], Xiaojuan Zhan[4, 5], Xiaorong Zhan[6], Chee Keong Kwoh[2], Yuezhongyi Sun[1, 5]

[1] School of Software and Microelectronics, Harbin University of Science and Technology, Harbin, Heilongjiang, China
[2] School of Computer Science and Engineering, Nanyang Technological University, Singapore
[3] College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, Heilongjiang, China
[4] College of Computer Science and Technology, Heilongjiang Institute of Technology, Harbin, Heilongjiang, China
[5] School of Computer Science and Technology, Harbin University of Science and Technology, Harbin, Heilongjiang, China
[6] Department of Endocrinology and Metabolism, the First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

Corresponding Author:
Dengju Yao
Harbin, Heilongjiang, China
Email address: ydkvictory@hrbust.edu.cn

Abstract

Metabolic syndrome is a cluster of the most dangerous heart attack risk factors (diabetes and raised fasting plasma glucose, abdominal obesity, high cholesterol and high blood pressure), and has become a major global threat to human health. A number of studies have demonstrated that hundreds of non-coding RNAs, including miRNAs and lncRNAs, are involved in metabolic syndrome-related diseases such as obesity, type 2 diabetes mellitus, hypertension etc. However, these research results are distributed in a large number of literature, which is not conducive to analysis and use. There is an urgent need to integrate these relationship data between metabolic syndrome and non-coding RNA into a specialized database. To address this need, we developed a metabolic syndrome-associated non-coding RNA database (ncRNA2MetS) to curate the associations between metabolic syndrome and non-coding RNA. Currently, ncRNA2MetS contains 1068 associations between five metabolic syndrome traits and 627 non-coding RNAs (543 miRNAs and 84 lncRNAs) in four species. Each record in ncRNA2MetS database represents a pair of disease-miRNA (lncRNA) association consisting of non-coding RNA category, miRNA (lncRNA) name, name of metabolic syndrome trait, expressive patterns of

non-coding RNA, method for validation, species involved, a brief introduction to the association, the article referenced, etc. We also developed a user-friendly website so that users can access and download all data easily. In short, ncRNA2MetS is a complete and high-quality data resource for exploring the role of non-coding RNA in the pathogenesis of metabolic syndrome and seeking new treatment options. The website is freely available at http://www.biomed-bigdata.com:50020/index.html

Introduction

Metabolic syndrome (MetS) is a cluster of the most dangerous heart attack risk factors: diabetes and raised fasting plasma glucose, abdominal obesity, high cholesterol and high blood pressure (Alberti et al., 2005). It is estimated that around 20-25% of the world's adult population have metabolic syndrome, making them three times more likely to have, and twice as likely to die from, a heart attack or stroke when compared to people without the syndrome (International Diabetes Federation, 2006). Metabolic syndrome has become a major threat to human health around the world. However, to date, the pathogenesis of metabolic syndrome continues to challenge experts. In recent years, a growing number of studies have suggested that many non-coding RNAs (ncRNAs), including small non-coding RNAs, particularly microRNAs (miRNAs), and long non-coding RNAs (lncRNAs), may be involved in metabolic syndrome-related diseases such as obesity, type 2 diabetes mellitus, hypertension etc. (Stoll et al., 2018; Sala et al., 2018; Esguerra et al., 2018; Cui et al, 2018; Lorente-Cebrián et al., 2019). Dysregulation of some miRNAs and lncRNAs disrupts the gene regulatory network, leading to metabolic syndrome and other related diseases. MiRNAs are ~22nt non-coding small RNAs that negatively regulate gene expression at the post-transcriptional level (Bartel, 2004). Extensive research suggests that many miRNAs, such as miR-9 (Hu el al., 2018), miR-20b-5p (Katayama el al., 2019; Gentile et al., 2019), miR-802 (Kornfeld el al., 2013), let-7f (Gentile el al., 2019), miR-33 (Rayner el al., 2010), miR-375 (Sedgeman el al., 2019), and others are involved in glucose homeostasis, diabetes mellitus, abdominal obesity and cholesterol metabolism. Furthermore, some lncRNAs, a novel class of long non-coding RNA larger than 200nt, have been reported to be involved in the pathogenesis of type 2 diabetes mellitus and metabolic syndrome (Singer, Sussel, 2018; Losko, Kotlinowski, Jura, 2016; Wang et al., 2018).

Due to the important effects of metabolic syndrome on human health, it is urgent to develop a database dedicated to various biomarkers associated with metabolic syndrome, such as genes, miRNAs and lncRNAs. In recent years, several data resources and tools have been developed for storing metabolic disease-associated biomolecules, such as T-HOD, metabolicMine, PathCaseMAW, HMA and BioM2MetDisease. T-HOD (Dai et al., 2013) is a literature-based candidate gene database currently containing 837, 835 and 821 candidate genes for hypertension, obesity and diabetes, respectively. metabolicMine (Lyne et al., 2013) is a data warehouse with a specific focus on the genomics, genetics and proteomics of common metabolic diseases. PathCaseMAW (Cicek et al., 2013) provides a database-enabled framework and web-based computational tools for browsing, querying, analyzing and visualizing stored metabolic

networks. HMA (Pornputtapong, Nookaew, Nielsen, 2015) is a human metabolic atlas website which provides information about human metabolism. These four software resources described above have provided important support for the study of the pathogenesis of metabolic diseases. However, they do not contain non-coding RNA information related to metabolic diseases. BioM2MetDisease (Xu et al., 2017) is a manually curated database containing 2681 entries of associations between 1147 biomolecules and 78 metabolic diseases. Though it is a very useful tool for studying metabolic diseases, BioM2MetDisease is not a database dedicated to metabolic syndrome. It contains miRNAs associated with 78 metabolic diseases but does not include hypertension and hypo-HDL cholesterolemia, which are two important traits of metabolic syndrome. In addition, it does not contain lncRNAs associated with metabolic syndrome and miRNAs from the last two years.

In addition to the five data resources described above, there are other resources and tools for studying human diseases. miR2Disease (Jiang et al., 2009) is a manually curated database which contains 1939 curated relationships between 299 human miRNAs and 94 human diseases. phenomiR (Ruepp, Kowarsch, Theis, 2012) provides miRNA and target relations from these studies on the association of dysregulated miRNAs and diseases. HMDB (v3.0) (Wishart et al., 2012) is a resource dedicated to the human metabolome, which includes more than 40,000 annotated metabolite entries. HMDD (v3.0) (Huang et al., 2018) manually collects a significant number of miRNA-disease association entries. All these databases have provided valuable tools for exploring the roles of these biomolecules in human diseases, but they are not designed specifically for metabolic syndrome. When faced with so many data resource options, it is difficult to find the desired data for doctors who focus on metabolic syndrome. In addition, more and more non-coding RNAs associated with metabolic syndrome have been discovered in the last two years (Saeedi et al., 2019; Zhang et al., 2019; Zhang et al., 2018; Smieszek et al., 2019; Lin et al., 2019).

So far, there is still no non-coding RNA database dedicated to metabolic syndrome. There is an urgent need for a specialized data resource containing all the latest non-coding RNAs associated with metabolic syndrome. To meet this demand, we have developed the ncRNA2MetS database which contains the latest and most complete MetS-miRNA (lncRNA) associations validated by various biological experiments. We carefully reviewed 571 articles about relationship between various metabolic syndrome traits and non-coding RNA and gained 1068 associations between five metabolic syndrome traits and 627 non-coding RNAs (543 miRNAs and 84 lncRNAs) in four species. We hope that this database can help doctors specializing in metabolic syndrome to explore the pathogenesis and treatments of metabolic syndrome.

Materials & Methods

Data collection from literature in PubMed

According to the International Diabetes Federation (IDF) definition, a person with metabolic syndrome must have central obesity (defined as waist circumference with ethnicity-specific values) plus any two of the following four factors: (1) raised triglycerides, or specific treatment for this lipid abnormality; (2) reduced HDL cholesterol, or specific treatment for this lipid abnormality; (3) raised blood pressure, or treatment of previously diagnosed hypertension; (4) raised fasting plasma glucose, or previously diagnosed type 2 diabetes (Alberti et al., 2005; International Diabetes Federation, 2006). Therefore, we divided the risk factors of metabolic syndrome into five traits: central obesity, type 2 diabetes mellitus, hypertension, hypertriglyceridemia and hypo-HDL cholesterolemia.

Referring to Xu's method (Xu et al., 2017), we manually collected and curated MetS-miRNA (lncRNA) associations from related articles in the PubMed database. First, we used 'non-coding RNA', 'ncRNA', 'microRNA', 'miRNA', 'long non-coding RNA', 'lncRNA' and each metabolic syndrome trait as search terms to search the PubMed database by Title/Abstract retrieval method. As a result, we gained more than 3000 related articles published since 2007. We filtered out a large number of irrelevant articles or reviews by reading abstracts and finally selected 571 articles that were really focused on association between metabolic syndrome trait and miRNA (lncRNA). Then, we manually extracted MetS-miRNA (lncRNA) associations by reading these selected articles in detail. In the process of extracting these associations, the detailed information about MetS-miRNA (lncRNA) association were collected, including non-coding RNA category, miRNA (lncRNA) name, name of metabolic syndrome trait, ICD-11 classification and DO (Disease Ontology) identifier for metabolic syndrome trait, method for validation (e.g. RNA-seq, luciferase report assays, gene knock-out), detected tissue (e.g. serum, adipose tissue, liver), expressive patterns (e.g. up-regulation, down-regulation, differential expression), name of the gene regulated by miRNA (lncRNA), species involved (e.g. homo sapiens, mus musculus, rattus norvegicus), referenced article (PubMed ID, title, year of publication) and a brief introduction to this association in the referenced article. Following previous research rules (Jiang et al., 2009; Xu et al., 2017; Huang et al., 2018), we only collected MetS-miRNA (lncRNA) associations validated by various biological experiments in this process. At the same time, in order to ensure the authenticity and reliability of the extracted information, each MetS-miRNA (lncRNA) association was confirmed by at least two scholars. Finally, in order to ensure consistency with other data resources, we standardized the names of miRNAs, lncRNAs and metabolic syndrome traits. We provided miRBase (Kozomara, Birgaoanu, Griffiths-Jones, 2019) identifier for miRNAs, NONCODE (Fang et al., 2017) identifier for lncRNAs, ICD-11 classification and DO identifier for metabolic syndrome trait. The process of constructing the ncRNA2MetS database is shown in Figure 1.

Database and website development

In order to facilitate users to access and use the data in the ncRNA2MetS database, we developed a user-friendly website providing data browsing, searching and downloading function. The website was implemented in Java programing language, and all data were stored in MySQL database. The website is freely available at http://www.biomed-bigdata.com:50020/index.html

158

159 # Results

160 **Database contents**

161 By April 2019, we gained 3699 potential articles from PubMed using “Title/Abstract” searching.
162 After manual screening according to the relevance of the research contents, 571 articles were
163 selected for reading in detail, and 1068 MetS-miRNA (lncRNA) associations were identified
164 finally. To describe MetS-miRNA (lncRNA) association in more detail, each record about MetS-
165 miRNA (lncRNA) association in ncRNA2MetS database consists of non-coding RNA category,
166 miRNA (lncRNA) name, miRBase identifier for miRNA, NONCODE identifier for lncRNA,
167 name of metabolic syndrome trait, ICD-11 classification and DO identifier for metabolic
168 syndrome trait, method for validation, detected tissue, expressive patterns, name of the gene
169 regulated by miRNA (lncRNA), species involved, information of cited articles and a brief
170 introduction to the associations in this referenced article. (see Materials and Methods). Currently,
171 the ncRNA2MetS database contains 1068 associations between five metabolic syndrome traits
172 (central obesity, type 2 diabetes mellitus, hypertension, hyperlipidemia and hypo-HDL
173 cholesterolemia) and 627 non-coding RNAs (543 miRNAs and 84 lncRNAs) in four species
174 (homo sapiens, mus musculus, rattus norvegicus and Sus scrofa). Among the 1068 associations,
175 the number of miRNAs related to central obesity, type 2 diabetes mellitus, hypertension,
176 hyperlipidemia and hypo-HDL cholesterolemia are 288, 207, 96, 50 and 41, respectively. In
177 addition, the number of miRNAs reported to be related to metabolic syndrome is 36 (Figure 2a).
178 The number of lncRNAs related to central obesity, type 2 diabetes mellitus, hypertension,
179 hyperlipidemia, hypo-HDL cholesterolemia and metabolic syndrome are 42, 28, 12, 1, 2 and 3,
180 respectively (Figure 2b).

181 **Database interface**

182 ncRNA2MetS database implements a user-friendly website interface through which users can
183 access all data in ncRNA2MetS conveniently and easily. The website consists of six parts,
184 namely HOME, BROWSE, SEARCH, DOWNLOAD, SUBMIT and HELP. The ‘HOME’ page
185 shows a brief introduction about metabolic syndrome, miRNA and lncRNA while ‘BROWSE’
186 page (Figure 3a) and ‘SEARCH’ page (Figure 3b) provide data query. On the ‘BROWSE’ page,
187 users can click a specific miRNA, lncRNA or metabolic syndrome trait to browse the MetS-
188 ncRNA associations. Then, the website will return all MetS-ncRNA associations that meet the
189 query criteria. If too many association entries are returned, users can specify ncRNA category,
190 species or validation method to screen for required entries. For example, if users specify ‘homo
191 sapiens’ as species, the website will return all associations related to ‘homo sapiens’ (Figure 3c).
192 The ‘SEARCH’ page provides users with faster and more accurate query method, which
193 support ‘Accurate Search’ and ‘Fuzzy Search’. For ‘Accurate Search’, users can input an
194 accurate miRNA (lncRNA) name or metabolic syndrome trait name, or both, then click ‘Search’
195 button to query the required associations. On the contrary, for ‘Fuzzy Search’, users only need to
196 input partial names of a miRNA (lncRNA) or metabolic syndrome trait to query the required

relationships. It should be noted that query keywords are case-insensitive. In addition, users can specify 'Validation Methods' to narrow query scope. For example, if the user specifies 'qRT-PCR' as the validation method, the website will return all associations that match query keywords and have been validated by qRT-PCT. Similar to 'BROWSE', in the search result page, users can also filter MetS-ncRNA associations by selecting specified ncRNA category, species or validation method. Finally, users can browse the detailed information about a specific MetS-ncRNA association by clicking a hyperlink to the 'Details' page (Figure 3d).

In addition to freely querying MetS-ncRNA associations stored in ncRNA2MetS, users can also submit novelty associations validated by their own experiments. They can do this on the 'SUBMIT' page (Figure 3e) and must provide detailed information about the new association. Our committee will regularly review new submissions. Once the submitted association is confirmed, it will be added into ncRNA2MetS. Furthermore, users can freely and easily download all MetS-ncRNA associations in the 'DOWNLOAD' page. Finally, if users encounter any difficulties or problems in using the ncRNA2MetS, they can find help information on 'HELP' page or contact us via e-mail.

Examples of using ncRNA2MetS

In this section, we will use examples to show you how to use the ncRNA2MetS database. First, users can input 'miR-155-5p' as the miRNA name on the 'SEARCH' page and then click the 'Search' button. A result page will be returned and will display all records about miR-155-5p, including those of different species and various metabolic syndrome traits. On the result page, one can easily notice that miR-155-5p is related to all five metabolic syndrome traits, which implies that it has a very important impact on metabolic syndrome. In fact, miR-155-5p has been reported to be a risk factor of metabolic syndrome. ncRNA2MetS also supports searching by the name of metabolic syndrome trait, and this can facilitate the study of pathogenesis for a certain specific metabolic syndrome trait. For example, users can input 'obesity' as the name of metabolic syndrome trait on the 'SEARCH' page and will find that a number of miRNAs and lncRNAs such as miR-21, miR-155-5p and Paral1 showed abnormal expression in human obesity. The introduction to these associations in ncRNA2MetS shows that a reduced level of miR-21 might be associated with obesity and its related metabolic traits such as hyperinsulinemia (Ghorbani et al., 2018); Obese subjects have increased expressions of miR-155-5p and miR-122, two miRNAs related to inflammation and iron metabolism, respectively, at both the systemic and sperm levels (López et al., 2018); Furthermore, a novel component of the adipogenic transcriptional regulatory network defining the lncRNA Paral1 is identified as an obesity-sensitive regulator of adipocyte differentiation and function (Firmin, et al., 2017).

As a feature, ncRNA2MetS also supports querying lncRNAs associated with metabolic syndrome. For example, by inputting 'H19', the result page will show all records of relationships between 'H19' and various metabolic syndrome traits including obesity, type 2 diabetes mellitus and hypo-HDL cholesterolemia in three species. The functional description in ncRNA2MetS shows that imprinted lncRNA H19 increases upon cold-activation and decreases in obesity in BAT (Schmidt et al., 2018); Related studies reveal a previously undescribed double-negative

feedback loop between sponge lncRNA and target miRNA that contributes to glucose regulation in muscle cells (Gao et al., 2014); A H19-miR130b pathway regulating lipid metabolism and inflammation response in ox-LDL-treated Raw264.7 cells provides new targets for atherosclerosis treatment (Han et al., 2018). Overall, ncRNA2MetS can be used as a high-quality and most complete data resource for studying the roles of miRNAs and lncRNAs involved in metabolic syndrome.

Database analysis

Currently, ncRNA2MetS provides almost all the research results related to the association between metabolic syndrome and non-coding RNA. Comprehensive analysis of the data in ncRNA2MetS can help people better explore the relationship between metabolic syndrome and non-coding RNA. For this purpose, a relational network between metabolic syndrome traits and ncRNAs (miRNAs and lncRNAs) is constructed using Cytoscape software (Figure 4). In the MetS-ncRNA association network, nodes represent metabolic syndrome traits and ncRNAs, and edges represent the relationships between them. For miRNAs and lncRNAs, degree represents the number of associated metabolic syndrome traits, and also indicates their importance for researching the pathogenesis and treatment of metabolic syndrome. Fig. 4 shows that the degrees of miR-122, miR-155-5p and miR-146a-5p (green node) were largest among the numerous miRNAs, which are related to all five metabolic syndrome traits (obesity, type 2 diabetes mellitus, hypertension, hyperlipidemia and hypo-HDL cholesterolemia), or have been reportedly involved in metabolic syndrome. This result implies that these three miRNAs play an important role in the study of metabolic syndrome. In addition, the lncRNA with the highest degree is H19, which is related to obesity, type 2 diabetes mellitus, and hypo-HDL cholesterolemia, and is reported to be involved in the pathogenesis of metabolic syndrome. Further and deeper analysis of the ncRNA2MetS data will yield more interesting results.

Discussion

Metabolic syndrome has become one of the most important diseases threatening human health worldwide. There is increasing evidence suggesting that metabolic syndrome is associated with abnormal expression of some ncRNAs, including miRNAs and lncRNAs. A database dedicated to metabolic syndrome-ncRNA association is helpful in studying the pathogenesis and treatments of metabolic syndrome. For this purpose, we developed ncRNA2MetS, a database containing almost all experimentally supported metabolic syndrome-ncRNA associations. Currently, ncRNA2MetS contains 1068 validated associations between five metabolic syndrome traits and 627 ncRNAs (543 miRNAs and 84 lncRNAs) in four species.

In recent years, some researchers have developed several high-quality databases, such as BioM2MetDisease (Xu et al., 2017) and HMDD (Huang et al., 2018), to provide metabolic disease-miRNAs associations. Nevertheless, these databases are not dedicated to metabolic syndrome and do not cover all the metabolic syndrome traits. For example, BioM2MetDisease contains 2681 entries of relationships between 524 miRNAs and 45 metabolic diseases across 14

species. This is a database with very rich storage content, but it is not specifically for metabolic syndrome. Though BioM2MetDisease contains a large number of miRNAs associated with obesity, type 2 diabetes mellitus and dyslipidemia, it lacks miRNA information related to hypertension and hypo-HDL cholesterolemia. Furthermore, although a large number of new miRNAs and lncRNAs related to metabolic syndrome have been identified and reported in the past two years, BioM2MetDisease has not been updated with the latest findings. HMDD (v3.0) is a database that curates experiment-supported evidence for human miRNA and disease associations. Currently, HMDD contains 32281 miRNA-disease association entries which include 1102 miRNA genes and 850 diseases. Similar to BioM2MetDisease, HMDD is not specifically for metabolic syndrome, and it does not contain miRNAs associated with hypo-HDL cholesterolemia. Furthermore, the number of miRNAs contained in HMDD is far less than ncRNA2MetS. To demonstrate the value of ncRNA2MetS, we comprehensively compared the amount of non-coding RNAs associated with various metabolic syndrome traits contained in BioM2MetDisease, HMDD and ncRNA2MetS. The results showed that ncRNA2MetS contained significantly more miRNAs associated with metabolic syndrome than both BioM2MetDisease and HMDD (Figure 5a). More concretely, the number of miRNAs associated with obesity, type 2 diabetes mellitus, hypertension, hypertriglyceridemia, hypo-HDL cholesterolemia and metabolic syndrome were 225, 204, 0, 24, 0 and 39 in BioM2MetDisease; 80, 128, 65, 8, 0 and 38 in HMDD; and 288, 207, 96, 50, 41 and 36 in ncRNA2MetS, respectively. Finally, ncRNA2MetS provides not only miRNAs but also lncRNAs associated with metabolic syndrome, and covers the latest research findings up to April 2019.

The prevalence of metabolic syndrome, including central obesity, type 2 diabetes mellitus, hypertension, hypertriglyceridemia and hypo-HDL cholesterolemia, is growing globally, and the amount of studies on metabolic syndrome is also increasing rapidly. To illustrate the research trend of ncRNAs related to metabolic syndrome, we counted the number of articles about non-coding RNA studies related to metabolic syndrome between 2007 and April 2019 (Figure 5b). Clearly, the amount of research about the association between ncRNAs and metabolic syndrome has increased rapidly in the past two years. Therefore, it will become a trend that more and more metabolic syndrome-ncRNA associations will be identified and validated in the future. We will keep track of the latest advances in the study of relationship between metabolic syndrome and non-coding RNA, and update ncRNA2MetS database regularly. In addition, we will focus on more types of non-coding RNA such as circular RNA, snoRNA, etc. and add associations between metabolic syndrome and these ncRNAs into ncRNA2MetS to increase its coverage. Furthermore, we will develop more powerful data analysis tools such as network visualization tool to help researchers better study the pathogenesis and treatment of metabolic syndrome in the future. In a word, we hope that ncRNA2MetS can be used as an effective tool for studying the mechanism of non-coding RNAs in metabolic syndrome.

Conclusions

A growing number of studies have suggested that many non-coding RNAs, including miRNAs and lncRNA, are involved in metabolic syndrome and its traits. In this article, we introduced ncRNA2MetS, a user-friendly web-based tool developed for curating the association between metabolic syndrome and ncRNAs (miRNA and lncRNAs). ncRNA2MetS currently contains 1068 associations between five metabolic syndrome traits and 627 ncRNAs (543 miRNAs and 84 lncRNAs) in four species. ncRNA2MetS has covered almost all relevant researches about the association between metabolic syndrome and ncRNAs between 2007 and 2019. It is expected that ncRNA2MetS will serve as a valuable data resource that will help researchers better study the pathogenesis and treatments of metabolic syndrome.

References

Alberti KGMM, Zimmet P, Shaw J. 2005. The metabolic syndrome--a new worldwide definition. *The Lancet* 366(9491):1059-1062 DOI 10.1016/S0140-6736(05)67402-8

Bartel DP. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116(2): 281-297 DOI 10.1016/S0092-8674(04)00045-5.

Cicek A E, Qi XJ, Cakmak A, Johnson SR, Han X, Alshalwi S, Ozsoyoglu ZM, Ozsoyoglu G. 2013. An online system for metabolic network analysis. *Database: The Journal of Biological Databases and Curation* 2014 DOI 10.1093/database/bau091.

Cui XW, Tan JM, Shi YJ, Sun C, Li Y, Ji CB, Wu J, Zhang Z, Chen SY, Guo XR, Liu C. 2018. The long non-coding RNA Gm10768 activates hepatic gluconeogenesis by sequestering microRNA-214 in mice. *The Journal of Biological Chemistry* 293(11): 4097-4109 DOI 10.1074/jbc.M117.812818.

Dai HJ, Wu JCY, Tsai RTH, Pan WH, Hsu WL. 2013. T-HOD: a literature-based candidate gene database for hypertension, obesity and diabetes. *Database: The Journal of Biological Databases and Curation* 2013 DOI 10.1093/database/bas061.

Esguerra JLS, Nagao M, Ofori JK, Wendt A, Eliasson L. 2018. MicroRNAs in islet hormone secretion. *Diabetes, Obesity and Metabolism* 20(S2): 11-19 DOI 10.1111/dom.13382.

Fang SS, Zhang LL, Guo JC, Niu YW, Wu Y, Li H, Zhao LH, Li XY, Teng XY, Sun XH, Sun L, Zhang MQ, Chen RS, Zhao Y. 2017. NONCODE V5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Research* 46(D1): D308-D314 DOI 10.1093/nar/gkx1107.

Firmin FF, Oger F, Gheeraert C, Dubois-Chevalier J, Vercoutter-Edouart AS, Alzaid F, Mazuy C, Dehondt H, Alexandre J, Derudas B, Dhalluin Q, Ploton M, Berthier A, Woitrain E, Lefebvre T, Venteclef N, Pattou F, Staels B, Eeckhoute J, Lefebvre P. 2017. The RBM14/CoAA-interacting, long intergenic non-coding RNA Paral1 regulates adipogenesis and coactivates the nuclear receptor PPARγ. *Scientific Reports* 7(1): 14087 DOI 10.1038/s41598-017-14570-y.

Gao Y, Wu FJ, Zhou JC, Yan L, Jurczak MJ, Lee HY, Yang LH, Mueller M, Zhou XB, Dandolo L, Szendroedi LJ, Roden M, Flannery C, Taylor H, Carmichael GG, Shulman GI, Huang YQ. 2014. The H19/let-7 double-negative feedback loop contributes to glucose

metabolism in muscle cells. *Nucleic Acids Research* 42(22): 13799-13811 DOI 10.1093/nar/gku1160.

Gentile AM, Lhamyani S, Coín-Aragüez L, Clemente-Postigo M, Olivera WO, Romero-Zerbo SY, García-Serrano S, García-Escobar E, Zayed H, Doblado E, Bermúdez-Silva FJ, Murri M, Tinahones FJ, Bekay RE. 2019. miR‐20b, miR‐296, and Let‐7f Expression in Human Adipose Tissue is Related to Obesity and Type 2 Diabetes. *Obesity* 27(2): 245-254 DOI 10.1002/oby.22363.

Ghorbani S, Mahdavi R, Alipoor B, Panahi G, Esfahani EN, Razi F, Taghikhani M, Meshkani R. 2018. Decreased serum microRNA-21 level is associated with obesity in healthy and type 2 diabetic subjects. *Archives of Physiology Biochemistry* 124(4): 300-305 DOI 10.1080/13813455.2017.1396349.

Han Y, Ma J, Wang JP, Wang LP. 2018. Silencing of H19 inhibits the adipogenesis and inflammation response in ox-LDL-treated Raw264. 7 cells by up-regulating miR-130b. *Molecular Immunology* 93: 107-114 DOI 10.1016/j.molimm.2017.11.017.

Hu DZ, Wang Y, Zhang HY, Kong DL. 2018. Identification of miR-9 as a negative factor of insulin secretion from beta cells. *Journal of Physiology and Biochemistry* 74(2): 291-299 DOI 10.1007/s13105-018-0615-3.

Huang Z, Shi JC, Gao YX, Cui CM, Zhang S, Li JW, Zhou Y, Cui QH. 2018. HMDD v3. 0: a database for experimentally supported human microRNA–disease associations. *Nucleic Acids Research* 47(D1): D1013-D1017 DOI 10.1093/nar/gky1010.

International Diabetes Federation. 2006. The IDF consensus worldwide definition of the metabolic syndrome. *Brussels: International Diabetes Federation.* 2006: 4

Jiang QH, Wang YD, Hao YY, Juan LR, Teng MX, Zhang XJ, Li MM, Wang GH, Liu YL. 2009. miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Research* 37(suppl_1): D98-D104 DOI 10.1093/nar/gkn714.

Katayama M, Wiklander OPB, Fritz T, Caidahl K, El-Andaloussi S, Zierath JR, Krook A. 2019. Circulating Exosomal miR-20b-5p Is Elevated in Type 2 Diabetes and Could Impair Insulin Action in Human Skeletal Muscle. *Diabetes* 68(3): 515-526 DOI 10.2337/db18-0470.

Kornfeld JW, Baitzel C, Könner AC, Nicholls HT, Vogt MC, Herrmanns K, Scheja L, Haumaitre C, Wolf AM, Knippschild U, Seibler J, Cereghini S, Heeren J, Stoffel M, Brüning JC. 2013. Obesity-induced overexpression of miR-802 impairs glucose metabolism through silencing of Hnf1b. *Nature* 494: 111-115 DOI 10.1038/nature11793.

Kozomara A, Birgaoanu M, Griffiths-Jones S. 2019. miRBase: from microRNA sequences to function. *Nucleic Acids Research* 47(D1): D155-D162 DOI 10.1093/nar/gky1141.

La Sala L, Micheloni S, De Nigris V, Prattichizzo F, Ceriello A. 2018. Novel insights into the regulation of miRNA transcriptional control: implications for T2D and related complications. *Acta Diabetologica* 55(10): 989-998 DOI 10.1007/s00592-018-1149-4.

López P, Castro A, Flórez M, Miranda K, Aranda P, Sánchez-González C, Llopis J, Arredondo M. 2018. miR-155 and miR-122 expression of spermatozoa in obese subjects. *Frontiers in Genetics* 9: 175 DOI 10.3389/fgene.2018.00175.

Lorente-Cebrián S, González-Muniesa P, Milagro FI, Martínez JA. 2019. MicroRNAs and other non-coding RNAs in adipose tissue and obesity: emerging roles as biomarkers and therapeutic targets. *Clinical Science* 133(1): 23-40 DOI 10.1042/CS20180890.

Losko M, Kotlinowski J, Jura J. 2016. Long noncoding RNAs in metabolic syndrome related disorders. *Mediators of Inflammation* 2016: 1-12 DOI 10.1155/2016/5365209c.

Lyne M, Smith RN, Lyne R, Aleksic J, Hu FY, Kalderimis A, Stepan R, Micklem G. 2013. metabolicMine: an integrated genomics, genetics and proteomics data warehouse for common metabolic disease research. *Database: The Journal of Biological Databases and Curation* 2013 DOI 10.1093/database/bat060.

Pornputtapong N, Nookaew I, Nielsen J. 2015. Human metabolic atlas: an online resource for human metabolism. *Database: The Journal of Biological Databases and Curation* 2015 DOI 10.1093/database/bav068.

Rayner KJ, Suárez Y, Dávalos A, Parathath S, Fitzgerald ML, Tamehiro N, Fisher EA, Moore KJ, Fernández-Hernando C. 2010. MiR-33 contributes to the regulation of cholesterol homeostasis. *Science* 328(5985): 1570-1573 DOI 10.1126/science.1189862.

Ruepp A, Kowarsch A, Theis F. 2012. PhenomiR: MicroRNAs in human diseases and biological processes. *Next-Generation MicroRNA Expression Profiling Technology*. Humana Press, 2012: 249-260 DOI 10.1007/978-1-61779-427-8_17.

Saeedi Borujeni MJ, Esfandiary E, Baradaran A, Valiani A, Ghanadian M, Codoñer-Franch P, Basirat R, Alonso-Iglesias E, Mirzaei H, Yazdani A. 2019. Molecular aspects of pancreatic β-cell dysfunction: Oxidative stress, microRNA, and long noncoding RNA. *Journal of Cellular Physiology* 234(6): 8411-8425 DOI 10.1002/jcp.27755.

Schmidt E, Dhaouadi I, Gaziano I, Oliverio M, Klemm P, Awazawa M, Mitterer G, Fernandez-Rebollo E, Pradas-Juni M, Wagner W, Hammerschmidt P, Loureiro R, Kiefer C, Hansmeier NR, Khani S, Bergami M, Heine M, Ntini E, Frommolt P, Zentis P, Ørom UA, Heeren J, Blüher M, Bilban M, Kornfeld JW. 2018. LincRNA H19 protects from dietary obesity by constraining expression of monoallelic genes in brown fat. *Nature Communications.* 9(1): 3622 DOI 10.1038/s41467-018-05933-8.

Sedgeman LR, Beysen C, Solano MAR, Michell DL, Sheng QH, Zhao SL, Turner S, Linton MF, Vickers KC. 2019. Beta cell secretion of miR-375 to HDL is inversely associated with insulin secretion. *Scientific Reports* 9(1): 3803 DOI 10.1038/s41598-019-40338-7.

Singer RA, Sussel L. 2018. Islet long noncoding RNAs: a playbook for discovery and characterization. *Diabetes* 67(8): 1461-1470 DOI 10.2337/dbi18-0001.

Smieszek A, Kornicka K, Szłapka-Kosarzewska J, Androvic P, Valihrach L, Langerova L, Rohlova E, Kubista M, Marycz K. 2019. Metformin Increases Proliferative Activity and Viability of Multipotent Stromal Stem Cells Isolated from Adipose Tissue Derived from Horses with Equine Metabolic Syndrome. *Cells* 8(2): 80 DOI 10.3390/cells8020080.

Stoll L, Sobel J, Rodriguez-Trejo A, Guay C, Lee KL, Venø MT, Kjems J, Laybutt DR, Regazzi R. 2018. Circular RNAs as novel regulators of β-cell functions in normal and disease conditions. *Molecular Metabolism* 9: 69-83 DOI 10.1016/j.molmet.2018.01.010.

Wang JP, Yang WL, Chen ZZ, Chen J, Meng YH, Feng BQ, Sun LB, Dou L, Li J, Cui QH, Yang JC. 2018. Long noncoding RNA lncSHGL recruits hnRNPA1 to suppress hepatic gluconeogenesis and lipogenesis. *Diabetes* 67(4): 581-593 DOI 10.2337/db17-0799.

Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu YF, Djoumbou Y, Mandal R, Aziat F, Dong E, Bouatra S, Sinelnikov I, Arndt D, Xia JG, Liu P, Yallou F, Bjorndahl T, Perez-Pineiro R, Eisner R, Allen F, Neveu V, Greiner R, Scalbert A. 2013. HMDB 3.0—the human metabolome database in 2013. *Nucleic Acids Research* 41(D1): D801-D807 DOI 10.1093/nar/gks1065.

Li XH, Tang SJ, Gui WW, Erik M, Tao TT, Li L, Wu F, Zhou JQ, Zheng FP, Li H. 2019. Circulating miR-143-3p inhibition protects against insulin resistance in Metabolic Syndrome via targeting of the insulin-like growth factor 2 receptor. *Translational Research* 205: 33-43 DOI 10.1016/j.trsl.2018.09.006.

Xu YJ, Yang HX, Wu T, Dong Q, Sun ZG, Shang DS, Li F, Xu YQ, Su F, Liu SY, Zhang YP, Li X. 2017. BioM2MetDisease: a manually curated database for associations between microRNAs, metabolites, small molecules and metabolic diseases. *Database: The Journal of Biological Databases and Curation* 2017 DOI 10.1093/database/bax037.

Zhang BH, Shen CA, Zhu BW, An HY, Zheng B, Xu SB, Sun JC, Sun PC, Zhang W, Wang J, Liu JY, Fan YQ. 2019. Insight into miRNAs related with glucometabolic disorder. *Biomedicine & Pharmacotherapy* 111: 657-665 DOI 10.1016/j.biopha.2018.12.123.

Zhang X, Li DY, Reilly MP. 2019. Long intergenic noncoding RNAs in cardiovascular diseases: Challenges and strategies for physiological studies and translation. *Atherosclerosis* 281: 180-188 DOI 10.1016/j.atherosclerosis.2018.09.040.

Figure 1

The flowchart of the ncRNA2MetS database design

The whole process is divided into three stages: (A) Literature retrieval; (B) Data extraction; (C) Database and website development.

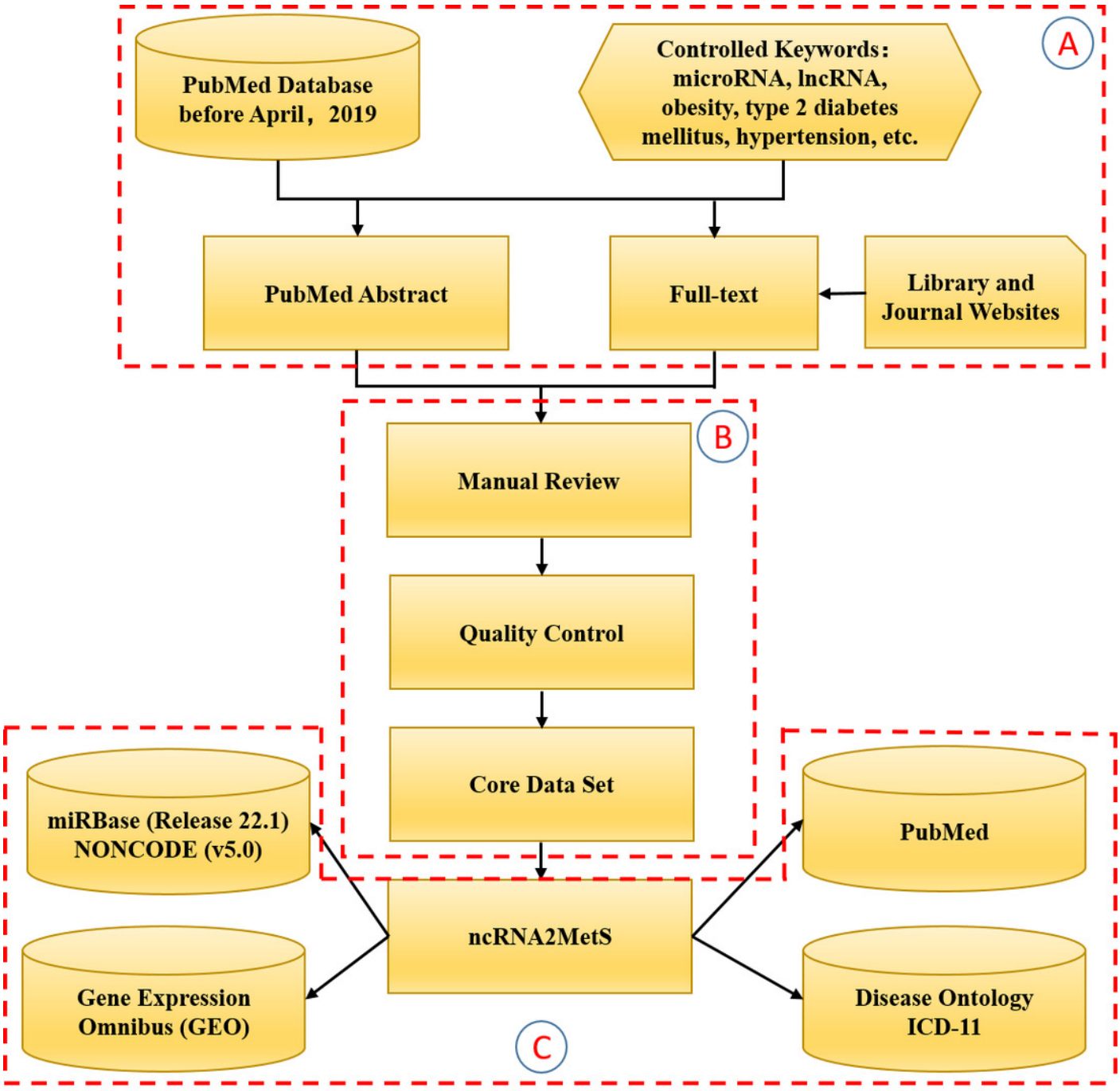

Figure 2

The statistics of ncRNAs contained in the ncRNA2MetS database

(A) The distribution of miRNAs in various metabolic syndrome traits. (B) The distribution of lncRNAs in various metabolic syndrome traits. (C) The distribution of miRNAs in different species. (D) The distribution of lncRNAs in different species.

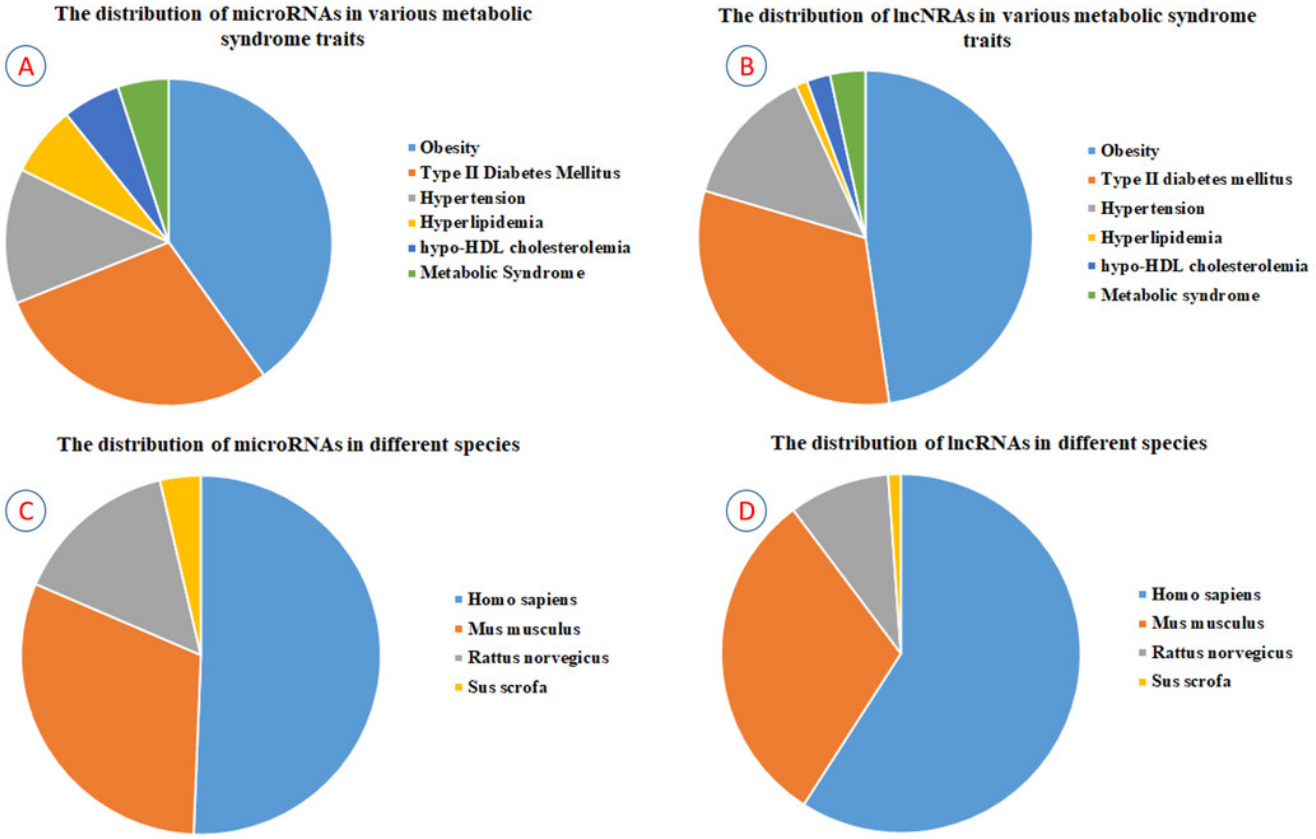

Figure 3

The schematic workflow of the ncRNA2MetS database

(A) Browse the data. (B) Search the data. (C) Browse the query results. (D) Browse the detail information about a specific MetS-ncRNA association. (E) Submit a new MetS-ncRNA association to the ncRNA2MetS database.

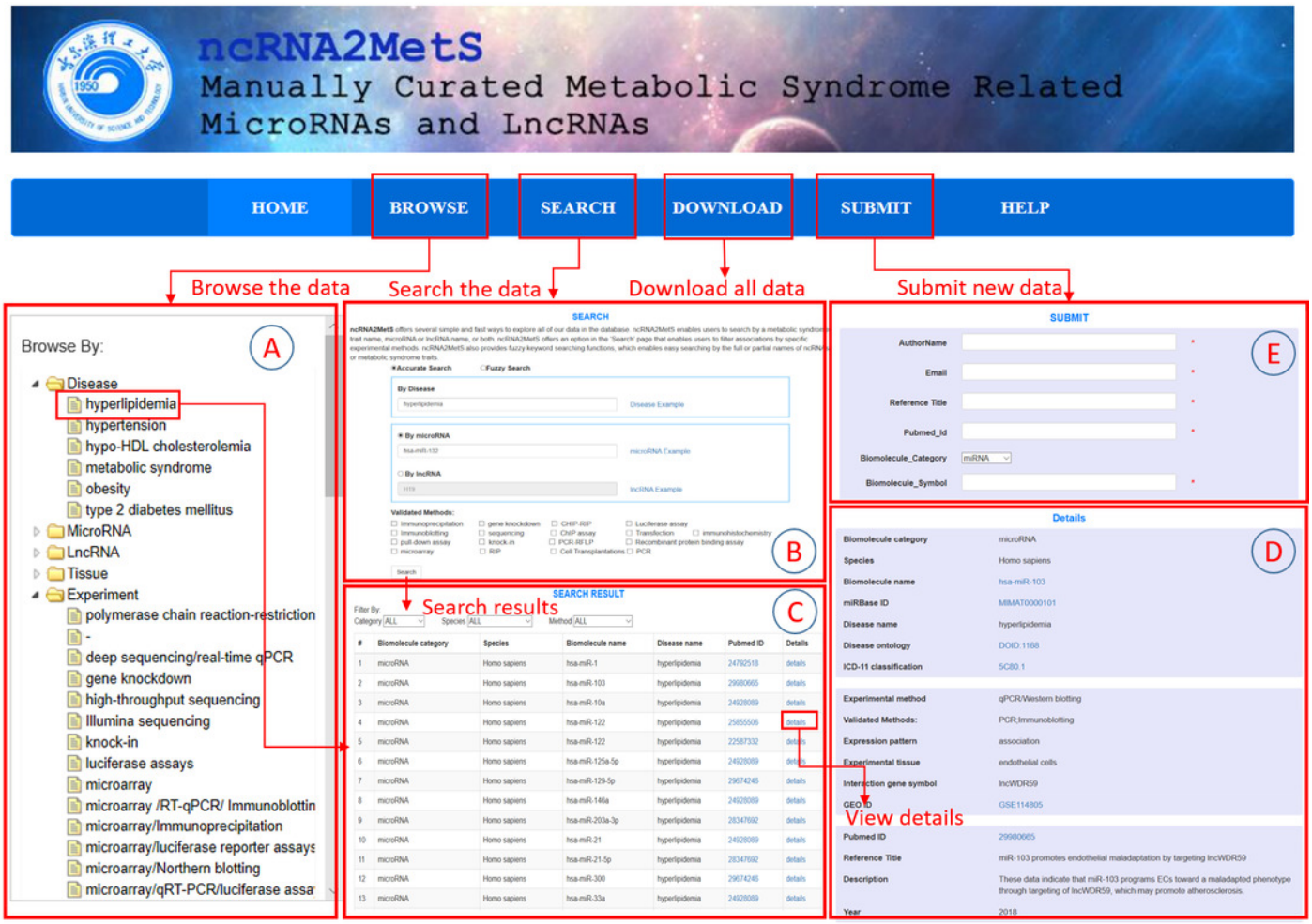

Figure 4

The MetS-ncRNA association network

Nodes correspond to ncRNAs (miRNAs and lncRNAs) and metabolic syndrome traits (central obesity, type 2 diabetes mellitus, hypertension, hyperlipidaemia and hypo-HDL cholesterolemia) and the edges correspond to experimentally supported associations. The size of the nodes corresponds to the nodes' degree.

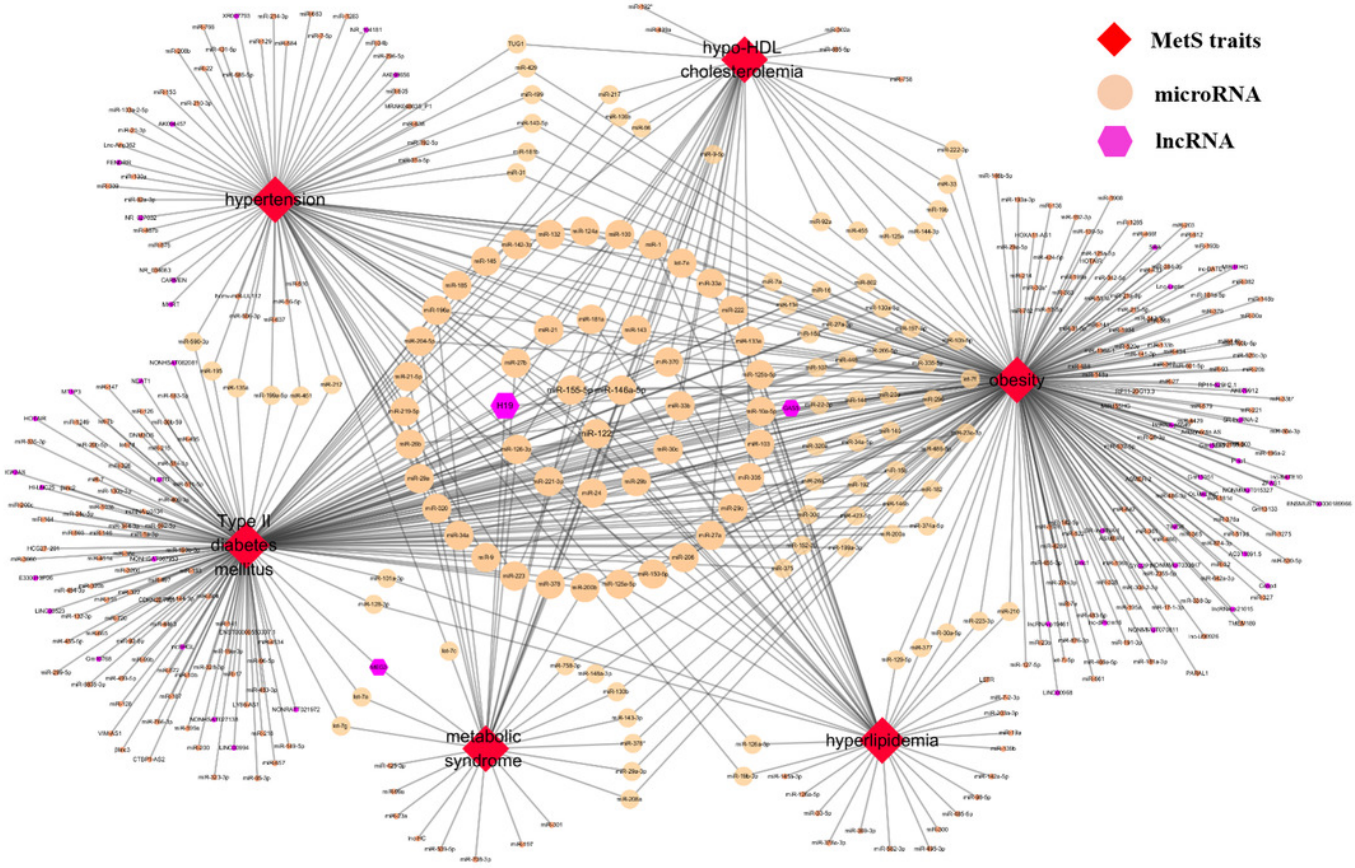

Figure 5

Comparison of the number of ncRNAs associated with metabolic syndrome in different databases

(A) Comparison of the number of ncRNAs associated with metabolic syndrome among BioM2MetDisease, HMDD and ncRNA2MetS. (B) Number of papers about ncRNAs associated with metabolic syndrome between 2007 and April 2019.