A review on computational tools for analytical visualisation and molecular interactions of sncRNAs: prospects in NDDs

Konstantinos Panagiotopoulos[1,*], Lolia Ibanibo[2,*], Benedetta Perrone[1], Caterina Marchetti[1], Martina Tumsich[1], Giovanni Traetta[1], Konstantinos Theofilatos[3,4] and Marco A. Deriu[1,*]

[1]*PolitoBIOMed Lab, Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Turin, Italy*
[2]*Instituto de investigación e innovación biomédica de Cádiz, Cadiz, Spain*
[3]*School of Cardiovascular and Metabolic Medicine & Sciences, King's College, London, United Kingdom*
[4]*Intelligent Systems Biology (InSyBio) PC, Patras Science Park building Platani, Patras, Greece*

Abstract

Neurodevelopmental disorders (NDD) including cognitive impairments, motor disabilities, and psychosocial disorders are common among infants that are born prematurely, but the molecular mechanisms behind them are still not clear. Nevertheless, recent studies have shown that there are some shared molecular pathways driving NDDs and neurodegenerative diseases with sncRNAs having a significant role in their manifestation. It is important to study and reveal the mechanism behind the development of these disorders to predict them as soon as possible, using biomarkers and allowing medical doctors to intervene early on, while neuroplasticity in newborns still allows for recovery to some extent. In this work, we examine the role of sncRNAs and some of the shared pathways in NDDs, but most importantly, we present some of the existing computational tools and databases for predicting target interactions, and tools to perform network analysis and visualization.

Keywords

Bioinformatics, Computational tools, Biological Databases, sncRNA, molecular pathways

1. Introduction

Preterm babies are considered those who have been born before the 37th week of gestation, while births given before the 32nd week, are considered very preterm [1, 2]. Premature deliveries have an average rate of more than 10% of total labors with an upward tendency

DETERMINED 2022: Neurodevelopmental Impairments in Preterm Children — Computational Advancements, August 26, 2022, Ljubljana, Slovenia
*Corresponding author.
konstantinos.panagiotopoulos@polito.it (K. Panagiotopoulos); loliaibanibo@gmail.com (L. Ibanibo); konstantinos.theofilatos@kcl.ac.uk (K. Theofilatos); marco.deriu@polito.it (M. A. Deriu)
https://www.dimeas.polito.it/it/personale/scheda/(nominativo)/konstantinos.panagiotopoulos (K. Panagiotopoulos); https://kclpure.kcl.ac.uk/portal/en/persons/konstantinos-theofilatos(cc24751e-4e2b-4471-aa12-380645c8e9b2).html (K. Theofilatos); https://www.dimeas.polito.it/en/personale/scheda/(nominativo)/marco.deriu (M. A. Deriu)
0000-0001-5034-378X (K. Panagiotopoulos); 0000-0001-6799-0553 (K. Theofilatos); 0000-0003-1918-1772 (M. A. Deriu)

worldwide [3]. From the clinical point of view, preterm infants with low birth weight have higher chances of experiencing short- or long-term neurodevelopmental disorders (NDDs) and related comorbidities [4]. Common NDDs are related to motor deficits such as cerebral palsy (CP), cognitive and speech delays, visual and hearing impairments, and some psychosocial and behavioral disorders such as Autism Spectrum Disorder (ASD) and Schizophrenia. The most common methods of assessment of the developing brains in infants are using magnetic resonance imaging (MRI), ultra-sound wave imaging, and Neuropsychological battery tests [5]. But this is a way of seeing the phenotype itself or predicting the outcome rather than finding the source of the problem. Previous works have shown that genetic factors such as copy number variations (CNVs) - which are repeated segments of DNA with higher (duplications) or lower (deletions) abundance than the reference genome - are linked to both intellectual disabilities [6] and motor impairments [7], and have a statistically significant relationship with NDDs and psychiatric comorbidities [4, 8, 9].

It is well known that even though most of the human genome (>76%) can be transcribed into RNA products, only a small fraction (∼3%) of it encodes for proteins [10]. These RNA molecules that do not follow the central dogma of molecular biology [11], are called non-coding RNAs (ncRNAs) and for many years were considered byproducts with low biological meaning. This perspective started to change, and scientists began to unravel the ncRNA mystery over the last decades with the help of advancements in sequencing methods and computational tools. Projects such as The Human Genome Project and The Encyclopedia of DNA Elements (ENCODE) [12], promoted the discovery of novel genes and shed light on functional elements encoded in the human genome, especially in non-coding areas, expanding our knowledge of their importance and their regulatory mechanisms. Studies and computational predictions suggest that even though NDDs and neuropsychiatric diseases are highly heterogenous, there are common enriched pathways and genetic factors between some of them [13, 14, 15]. As an example, ASD, Tourette syndrome (TS), and Schizophrenia share some genetic modifications that may lead to dysregulation of gene expression related to micro RNAs (miRNAs); a specific regulatory group of short non-coding RNA molecules [10].

According to their average size, ncRNAs can be categorized into two general groups: long non-coding (lncRNA) and small or short non-coding (sncRNA). LncRNAs extend to over 200 nucleotides (nt) and usually have a similar size to messenger RNAs which is more than 1000nt [16], while sncRNAs typically have a length below 200nt and they are separated into two groups based on their role in the cell; Housekeeping and Regulatory [10]. Except for other important functional roles in the cell, lncRNAs such as pseudogenes and circular RNAs can interact with some classes of sncRNAs, lowering their abundance in the free form through complementarity sequences. Housekeeping sncRNAs were discovered relatively early and are well studied, due to their abundance and their fundamental roles in the function of the cell. For example, their roles can be the amino acid transfer (tRNAs) at protein synthesis or being involved in RNA processing and splicing in the nucleus (snRNAs). Regulatory sncRNAs have drawn the attention of scientists only in the last decades when technological advancements allowed for it. Since then, their important role started to unravel and it was found that they actively interact and interfere with other molecules, regulate gene expression, and involve in important molecular pathways [17, 18, 19, 20, 21]. This control over the gene expression of the regulatory sncRNAs is important because, in many diseases dysregulation of sncRNAs sequentially causes

dysregulation of functional elements that then lead to pathological phenotypes [22, 23]. Because of the great importance of these molecules, bioinformatics tools and dedicated databases have been developed in the last decades to explore their role as biomarkers and their potential in medicine. A visual taxonomy of the classification of RNA molecules can be seen in figure 1.

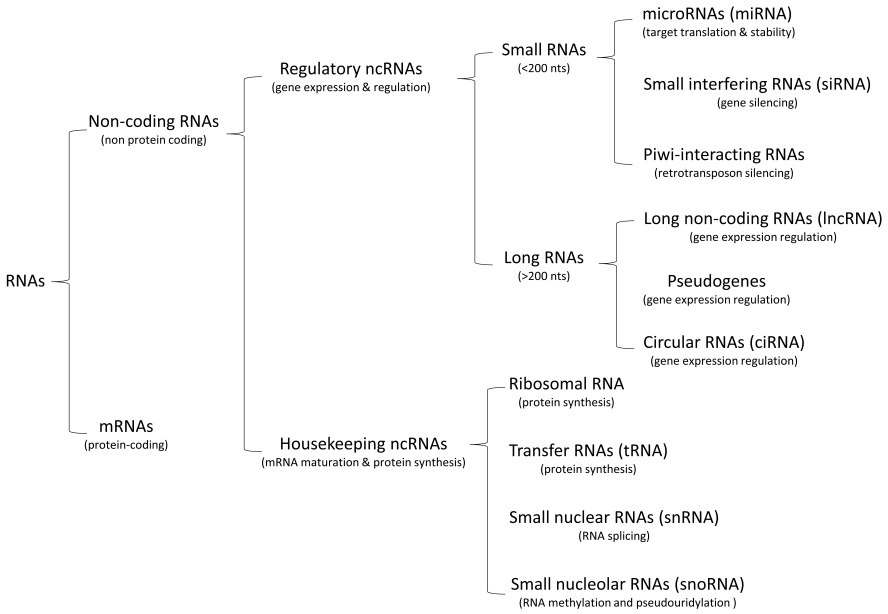

Figure 1: Classification of RNAs. The first split is based on the transcript product (coding/non-coding), followed by a discrimination on the general role of the non-coding RNAs and further on the size of these molecules. This is a modified version of the Figure 1 in the work of Gomes et al. (2020) [24].

After the systematic studies of sncRNAs, biologists clustered them by similarity and function with the most common ones being: microRNAs (miRNA) and small interfering RNA (siRNA) which regulate gene expression, small nuclear (snRNAs) that involve in RNA splicing, and piwi-interacting RNA (piRNA) that mainly interfere with transposable elements (or transposons) [10]. It is well established that sncRNAs hold a significant role also in many diseases in humans, and they can be used as biomarkers for diagnosis or prognosis, as drug targets, and as potential therapeutic methods [25]. Special attention has been given to miRNAs due to their high theoretical and experimental total number, the number of their interactions, and the role they have in both defending the homeostasis in the cell, but also related to diseases like cancer if they are dysregulated [26].

Because of the numerous interactions of sncRNAs and other molecules in the manifestation of diseases, a common approach is to handle this complexity with the use of interaction networks. Since our understanding of the underlying mechanisms is still unclear for the majority of these diseases, studying individual relationships is not enough to unravel and understand the dynamic of these pathologies. Rather than this, a more holistic view is needed with the help of multi-layer networks integrating instances belonging to different levels of complexity and domains (RNAs, proteins, diseases, functions, etc.) [27]. In this context, computational modeling can help in

reconciling the advancements in high throughput technologies with studies under the scope of systems and also explore the pathogenesis of diseases by understanding the molecular relations driving them, promoting treatments, drug discoveries, and precision medicine [28].

There are multiple tools nowadays that have been developed to predict the interactions of sncRNAs and especially miRNAs. Computational methods try to predict targets of these molecules [29, 30], pathways, and mechanisms involved in multiple diseases and disorders. Due to their interesting nature, ncRNAs have been systematically studied and there are multiple databases available where one can find experimental and computational information about them, based on their categorization. Most of these databases are open-access and publicly available. This makes the contained information accessible to everyone, helping scientists to build predictive models for diseases, discover potential biomarkers, and even design potential therapeutic targets.

Relevant studies were identified in PubMed, Scopus, ScienceDirect, and IEEE Xplore with no language restrictions. The first search from these databases was performed by the first author of this review and double-checked by the other corresponding authors. The following keywords were used: (sncRNAs OR miRNA OR siRNA OR piRNA OR RNAi), (neurodevelopmental co-morbidities OR co-occurrence of neurodevelopmental disorders), non-coding RNA Databases, (bioinformatics tools AND target prediction of sncRNA). We included only papers from January 2000 up to August 2022. Older papers were excluded, with the exception of papers explaining concepts or statistical and mathematical techniques

In this article, we review some of the most widely used molecular biology-related databases for the characterization and functionality of sncRNAs, and the state-of-the-art of computational tools for the analysis of these RNA molecules in various comorbidities, such as NDDs observed in some preterm infants [31, 32]. The main objective is to comprehensively collect in one article information on the effectiveness and usability of biological databases and databanks, as well as some computational tools for different types of bioinformatics analysis that are considered or could be considered in the future for research in the field of neurodevelopmental disorders, also considering preterm infants.

2. Features of sncRNAs

Regulatory sncRNAs can derive usually from individual genes or introns of other genes, but it is known that by the procedure of alternative splicing they may also contain some exon sequences. The most studied categories are miRNAs and siRNAs which have been found to involve in many pathologies [33] and the developmental processes [34]. The biogenesis of miRNAs has five main steps: transcription into a primary (pri-miRNA) form of a stem-loop, cleavage into a shorter stem-loop precursor (pre-miRNA) known as hairpin, transportation of the hairpin out of the nucleus, and a second cleavage followed by the unstranding of the two counterparts to produce the mature miRNA. These molecules are typically 20-24nt long [20, 35] and bind to their targets through partial complementarity -not necessarily perfect- of their seed (nucleotides 2-8), and their mRNA target sequences in the 3' untranslated region (3'UTR) which are called miRNA response elements (MREs). This leads to degradation of the mRNA molecule, or the disruption of translation by preventing the binding of ribosomes on

the mRNA. In both cases, translational inhibition results in the silencing of the target gene, a process also called RNA interference (RNAi). MREs can be found also in other types of RNAs like in lncRNAs and pseudogenes, which increases the number of targets for miRNAs since they are not strictly target-specific. This lower specificity gave rise to the idea of competitive endogenous RNAs (ceRNA). In the ceRNA field, miRNAs become the target of other competing molecules which based on their concentration, affinity, and the number of MREs regulate the abundance of available miRNAs resulting in an indirect regulation of their own expression.

Similarly, siRNAs regulate the gene expression of their target. These molecules by structure, are almost identical to miRNAs, but with the difference of having very high specificity to their target since they usually have perfect complementarity of base pairing with them [36]. The function of siRNAs lies in the interference with gene expression by degrading their transcript-targets which are far fewer targets than those of miRNAs. piRNAs on the other hand, follow a different biogenesis process which remains unclear to some extent, and also have different mechanisms of action. They are produced by a process related to the P-element induced wimpy testis or PIWI subfamily members, and recently they have been associated with cancer biology. The structure of piRNAs is single-stranded molecules of length range 26-31nt and they are known for epigenetic regulation through histone modification, but mostly for interfering with transposable elements or "genomic parasites", protecting the genome of the host [37].

The biogenesis and the mechanism of interaction of regulatory sncRNAs are important since this knowledge is also implemented in the computational tools that predict their targets. Common features on which the majority of target prediction tools are based are: the seed match, conservation sequences, free energy, and site accessibility [38].

2.1. sncRNAs in neurological disorders

sncRNAs are crucial in the maintenance of homeostasis since they coordinate the expression of genes through the RNAi process. It has been reported by many studies that sncRNAs have a linked role in neurodegenerative diseases, various types of cancers [26, 37] where the expression of sncRNAs is heavily dysregulated due to mutations, and neurodevelopmental disorders [20, 39, 40]. Specifically, sncRNAs have been found to be part of enriched molecular pathways in numerous neurodevelopmental disorders and comorbidities like Rett syndrome, ASD, Down syndrome, and others [7, 10, 39, 40]. In fact, a known commonly altered pathway in neurodevelopmental and psychiatric disorders is the mTOR pathway [41, 42]. This may indicate that there are similar mechanisms between these disorders that lead to higher probabilities of comorbidity. The role of sncRNAs in pathologies makes these molecules perfect candidates for biomarkers for early detection of diseases and mechanisms of diagnosis [43], as they are also highly ranked as therapeutic targets and in drug discovery research [44, 45]. From what is known, although there are some sncRNAs that have been identified to have different expression levels and to involve in the manifestation of NDDs, they do not show specific characteristics or significant features compared to other sncRNAs. However, mutations in the genes of the regulatory ncRNAs may be responsible for the occurrence of specific NDDs [46].

3. Computational tools to investigate molecular mechanisms and characteristics of sncRNAs

With the emergence of clinical and biological databases, as well as with the new technologies in sequencing (micro-arrays, next generation sequencing (NGS), tiling arrays, etc.), new opportunities arose for computational biology and the exploration of microscopic and macroscopic processes. Methods and tools started developing to tackle the challenges of massive amounts of data and the complication of biological systems. Results of multiple focused experiments started gathering and the findings were made easily accessible for further analysis. Additionally, databases started implementing online tools for processing information, predicting, and storing multiple-level entities because of the interoperability between different databases [47]. The ever-increasing number of databases along with the availability of data due to the new technologies of high throughput techniques led to the development of new tools, methods, and pipelines for handling the amount of available data and the extraction of new knowledge.

3.1. Biological Databases

The need for databases comes from the scattered information in literature. Having a comprehensive dataset helps researchers -especially in the clinical industry- to use the obtained knowledge from multiple and different experiments easily and find associations between instances leading to a better understanding of some conditions and processes. Biological databases can be manually or automatically curated, which means that they are constantly updated with new knowledge coming either from experiments or computational predictions. Additional to databases of linked information, there are databanks where raw data from experiments are stored. This, except for being a source of information for the databases, allows for meta-analysis of the data and merging of experiments to increase the amount of data in individual studies.

In the last decades, many efforts have been done to summarize the information about ncRNAs, as it is a game-changer in the study of cellular processes and gene regulation. Two commonly used sources of raw data are the Gene Expression Omnibus (GEO) [48] and the Sequence Read Archive (SRA) [49]. Both are from the National Institutes of Health (NIH), a part of the United States Department of Health and Human Services. In GEO, one can find collections of genomic data grouped by studies for multiple instances, and information about the protocols followed in the conducted studies. Genome browsers such as Ensembl, UCSC, and NCBI, provide interactive and comprehensive annotations of the genes on the human genome, as well as multiple tools for further bioinformatics analysis such as variant predictors and sequence comparison tools. The sequences they contain for ncRNAs are usually imported from other sources that have been created for storing information. GeneCards and HUGO Gene Nomenclature Committee (HGNC) are examples of generic databases containing information about both coding and non-coding genes. Location, aliases, description, and links to other databases can be found here, but still they only host the information contained in the sncRNA-specific databases. There are also multiple browser-based available tools for the analysis of the datasets such as gene identification tools for differential expression analysis on two or more groups. SRA is a repository of high throughput sequencing data, containing the raw sequences and alignment information, promoting reproducibility and new discoveries

through data analysis. Similar to GEO and still from the NIH is the database of Genotypes and Phenotypes (dbGaP), which provides also a controlled access space, meaning that some of the datasets stored needs authorization to get access. Finally, ArrayExpress [50] supported by the European Bioinformatics Institute (EMBL-EBI), stores high-throughput data from functional genomics experiments. The difference with the previous databanks is that ArrayExpress contains both the processed data and the raw sequences as well as links to the European Nucleotide Archive (ENA). All of these repositories contain both coding and non-coding sequences which are the building blocks for the biological databases holding information about the structure, attributes, and interactions of molecules.

General Biological Databases

A large number of biological databases for sncRNAs have been created through the years with diverse purposes such as annotation, structural information, function, interactions, location, sequence, and others. There is a large part of overlapping and redundant results contained in the databases, because of the interoperability and the information exchange between different providers. For many years now, there have been efforts to map and annotate all genes and transcripts, especially in the ever-increasing field of non-coding RNAs. The reason for non-coding-specific databases is that knowing the sequence of these molecules is the most crucial information for finding their interactions and developing computational tools for their analysis.

sncRNA sources

miRBase. The miRBase founded in 2003 [51] is among the most significant databases for miRNA sequences storage and annotation, with the latest version v22.1 (2019) containing 1917 hairpin instances and 2500 mature miRNAs for the human species alone. miRbase integrated multiple tools for sequence annotation, target prediction, and new sequence registration [52]. Additionally, it includes both experimentally verified and computationally predicted active sites and targets, and it is one of the main sources of miRNA information for other databases. Currently, there is an effort to synchronize the miRbase with Rfam; a collection of RNA families including sncRNAs with additional information about secondary structures. Both of these databases contain classifications for microRNA families but so far obtained with different methods and have a consensus of only 28% between them.

miRTarBase is a biological database that mainly provides generally validated experimentally miRNA-Target Interactions (MTI) collected in a manual way [53]. miRTarBase contains more than 4.4M interactions of about 3000 miRNAs for humans and has search filters based on specific miRNA names, their targets, and diseases.

miRCarta [54] implements the information of precursor and mature miRNAs coming from miRBase as well as predicted ones resulted from the online pipeline **miRMaster** [55]. The import of these predicted miRNAs which are based on the sequence of the sample data, results in a huge number of miRNAs in the database which is around 25k mature miRNAs and 15k precursors for the human species alone.

An interesting and recently published comprehensive database for circulating sncRNAs is **EVAtlas** [56]. It contains information for multiple families of non-coding RNAs from disease and control datasets originated from different tissues and sources. Data collection for EVAtlas is

made from 57 GEO and SRA manually reviewed registries, making it a great tool for circulating biomarker studies.

Interactions and Targets databases

miRNet [57] visualization of miRNA and other molecule interactions, can be used for multilayer network construction and ceRNA networks. It links miRNAs to coding and non-coding molecules, transcription factors, and diseases. These features make miRNet a great tool for multi-layer network reconstruction.

Pathways and enrichment analysis databases

Other than databases with structural details and interactions for ncRNAs, there are databases containing information on the involvement of ncRNAs in molecular pathways and processes, linking them in a functional role beyond their immediate first-degree interactions. The Kyoto Encyclopedia of Genes and Genome (**KEGG**) is among the most used databases for pathways, storing genomic and pathway information, and providing manually drawn maps of interactions, regulations, and signal cascading. Despite the fact that it is so well organized, KEGG has a limited amount of information about sncRNAs, and most of them are related to cancers. **Reactome**, is another generic human curated biological pathway database, that cross-references its information with NCBI, Ensebml, KEGG and others [58]. It implements online tools for analyzing and interpreting interactions and visualization of networks, but it also has relatively limited information about ncRNAs. For this reason, miRPathDB [59] has been created to indirectly link the regulatory information of miRNAs to the molecular pathways. Although **miRPathDB** [59] does not calculate the interactions, it uses context mining techniques to gather information from different enrichment analysis and pathway generic sources (KEGG, GO), linking them to information of ncRNA databases as miRBase or miRCarta.

RISE is a repository for RNA-RNA interactions coming mainly from transcriptome-wide studies [60]. Although RISE contains information about interactions between sncRNAs and other RNA molecules, it mostly focuses on lncRNA interactions. Thus, the use in sncRNA studies can be used in a validation step of a ceRNA network. **NPinter** [61] contains interactions between ncRNAs (except tRNAs and rRNAs) and biomolecules (proteins, RNAs, and DNAs) with the additional feature of visualizing the network of first-degree interactions between the query and the target. The drawback of this database is the limitation to interactions.

Lastly, a broader open-source RNA interaction database is starBase or **ENCORI** [62] which integrates information for 23 species from which it has more than 4.1 million miRNA-ncRNA interactions and 2.9 million miRNA-mRNA interactions. The data for ENCORI comes from the analysis of high throughput datasets, gene co-expression analysis, and signaling pathways sources [62]. ENCORI offers the option of searching for interactions based on the type of interaction (miRNA-Target, RNA-RNA) as well as ceRNA-Network and pathways based on KEGG terms.

3.2. Bioinformatics Tools

Bioinformatics tools are used to make the analysis of complex biological systems possible, fast and reliable. Once the sequence of sncRNAs is known through experiments and/or

prediction techniques (e.g. miRMaster), and the information of interactions is available in databases, the analysis usually proceeds with the creation of networks. Networks of single or multiple-level instances such as molecules, diseases, and pathways coexist and interact in one graph. In the case of novel transcripts, where there is no experimental evidence or previous knowledge of the targets of ncRNAs, computational tools try to predict the most probable interactions of these molecules in various ways. A list of the databases and tools discussed in this work can be found in Table 1

Target prediction tools

Binding site prediction for sncRNAs is usually referred to miRNAs and siRNA targets which are calculated based on thermodynamic criteria, anti-correlation of target genes, and miRNA/siRNA expression, but most significantly by nucleotide sequences in the target's 3'UTR MREs. Many tools developed in the last decades for this difficult task, with the most popular one being the **TargetScan**. An online computational tool for target prediction of miRNAs, based on the complementarity between the query gene transcript and the seed of the miRNA along with other multiple features related to the nucleotide sequence of the targets [63]. **DIANA** is a set of tools with the microT algorithm predicting miRNA targets in canonical (3'UTR) regions and the microT-CDS [64] algorithm for the non-canonical (coding) regions. DIANA implements also the LncBase and TarBase [65] databases for experimentally verified miRNA-target interactions with non-coding and coding transcripts respectively, and mirPath tool for identifying potential altered pathways based on miRNA expression profiles. There is a plethora of other tools and databases related to target prediction such as **miRecords** [66] or **miR2Disease** [67] which contains information about miRNAs related to specific diseases, but they are not as comprehensive or updated as the previously mentioned ones even though they are holding valuable information and are sources for databases.

Network reconstruction and visualization

The use of networks in molecular interactions is crucial to depict and tackle the complexity of biological systems. One of the uses of biological networks is the visualization of interactions, which in small networks is easy to interpret but when there are hundreds or thousands of nodes and edges it gets overwhelming for a human to handle. So, a more useful application for these systems is the analysis based on the graph theory. Metrics of centrality and affinity can be used to evaluate significant nodes and pathways, leading to important conclusions such as potential therapeutic targets [68]. Moreover, instances belonging to different categories (e.g. genes, variations, and phenotypes) can be integrated into an interactive network and help to draw conclusions about difficult problems. Tools that are used in bioinformatics for visualization of networks and analysis derive from generic network-reconstruction tools that are based on maths and the graph-theory. Functionality related to the field of biology was added through the years, mostly in the form of add-on modules that extend the basic metrics and enrich them with biological information through the available databases.

Pajek [69] is a generic, more than 20 years old, Microsoft Windows-based network visualization tool, initially implemented for social network analysis. It is also considered an immensely powerful application for analysis and visualization of massive networks because it can easily visualize a million nodes with billion connections in an average computer. For Pajek there

are available implementations that are optimized to handle faster and with a lower need for memory larger structures (Pajek-XXL or Pajek-3XL). It also implements numerous features such as Graph layout, node merging, neighborhood detection, identification of strongly connected components, clustering, and many other network analysis metrics and tools. This feature makes it a great tool for massive networks but with lower quality visualization potential.

Gephi [70] is a free offline open-source, leading visualization and exploration software and runs on all main operating systems. It is not designed specifically for biological networks, rather it is a general-purpose tool for exploratory data analysis, social networks, and biological network analysis. In Gephi there are multiple plugin modules designed for clustering of nodes and statistical analysis. It is user-friendly, allowing for customization in the visualization and due to its flexible multi-task architecture is very fast even for large datasets.

Cytoscape [71] is probably the most popular open-source desktop application for 2D network visualization in biology and health sciences. It supports all kinds of networks (e.g. weighted, unweighted, bipartite, directed, undirected, and multi-edged) and comes with an enormous library of plugins with more than 250 modules. It was initially designed for research related to biology, as its first aim was to analyze molecular interaction networks and biological pathways, integrating them with other state data such as gene expression profiles. It can handle big networks, but it requires more memory and time for clustering and layout routines than other tools which makes it less scalable, and it is recommended to run such processes in the command line and then load the results as node/edge attributes. It is a good compromise between analysis and visualization, and it comes with a great plethora of layout, clustering, and topological network analysis algorithms, such as AutoSOME, Eisen's hierarchical and k-Means clustering (in the ClusterMaker plugin), and the basic network metrics of average connectivity betweenness centrality and others. Finally, plugins for the connection of biological databases of functional enrichment, GO annotations, data retrieval, and others have been developed making it very convenient to work with.

Other solutions for network analysis may include market products or whole pipelines of processing. These solutions are usually less customizable but require less knowledge of the underlying methods and fewer resources of computational power from the user. One such example is InSyBio's suite, which implements multiple tools from the level of RNA-sequence analysis up to the network analysis by InSyBio BioNets [72] for identifying important nodes and potential biomarkers using machine learning approaches.

4. Conclusion

Since their discovery, the importance of non-coding transcripts has become clear, and they stop being considered as "Junk" DNA regions. With the advancements in technology allowing for the detection of these molecules and especially sncRNAs, a huge number of ncRNAs were discovered and got annotated. Even though some of their mechanisms of action have been decoded, their full functionality still remains to be discovered. Despite what is unknown, the focus on regulatory sncRNAs, led to significant improvements in our understanding of the molecular mechanisms driving certain diseases, and the prognosis of pathologies. It is not known if there are specific characteristics and features of the sncRNAs that are involved in

	Name	URL
Bio Databases	**Genome Browsers**	
	Ensembl	http://www.ensembl.org/index.html
	NCBI	https://www.ncbi.nlm.nih.gov/genome/51
	UCSC	https://genome.ucsc.edu/
	Non-codingRNA related	
	miRBase	https://www.mirbase.org
	miRTarBase	https://mirtarbase.cuhk.edu.cn
	miRCarta	https://mircarta.cs.uni-saarland.de
	EVAtlas	http://bioinfo.life.hust.edu.cn/EVAtlas
	miRNet	https://www.mirnet.ca
	Molecular Pathways	
	KEGG	https://www.genome.jp/kegg
	REACTOM	https://reactome.org
	miRPathDB	https://mpd.bioinf.uni-sb.de
	RISE	http://rise.life.tsinghua.edu.cn
	NPinter	http://bigdata.ibp.ac.cn/npinter4
	ENCORI	https://starbase.sysu.edu.cn
Bioinformatics Tools	**Target prediction**	
	miRMaster	https://ccb-compute.cs.uni-saarland.de/mirmaster2
	TargetScan	https://www.targetscan.org/vert_80
	DIANA	https://diana.e-ce.uth.gr/home
	miRecords	http://c1.accurascience.com/miRecords
	miR2Disease	http://www.mir2disease.org
	Network Visualization	
	Pejek	http://mrvar.fdv.uni-lj.si/pajek
	Gephi	https://gephi.org
	Cytoscape	https://cytoscape.org

Table 1
List of databases and tools discussed in the present work and the uniform resource locator (URL) for each of these sources

NDDs, so further studies are needed to understand and unravel the sources of these disorders. Computer science and Bioinformatics have a tremendous impact on systems biology, with the ever-improving development of tools helping scientists draw important conclusions from the massive amounts of available data. And this is why it is so important to have comprehensive, curated, open-source, and well-organized databases as the ones presented in this work.

The availability of datasets stored in databanks along with the interoperability and the organization of information in databases has dramatically shifted the nature of biological studies from small- to large-scale and gave rise to data-driven methods. This alternation of viewing multiple interactions and functions brought the use of multi-layer networks into the foreground as an important tool. This allowed for broader and more holistic computational approaches, which model much better the real biological systems.

To date, none of the presented tools is specific to neurodevelopmental disorders. In fact, these tools are of general use, but there is an interesting potential for application in various fields, including neurological and neurodevelopmental disorders. This derives from the fact that RNA

molecules and specifically sncRNAs have simple structures, with no particular biochemical features. Thus, the individual tools that analyse these molecules are general. Despite that, their combination -depending on the question every time- can lead to pathology-specific methods which are related to the emerged properties of more complex structures such as tissues, organs, diseases etc. The purpose of this paper is to collect the most well-known and important tools and to give an insight into their functionality and effectiveness. This is an important step towards understanding their potential in specific fields such as neurodevelopmental disorders.

Of course, except for the presented tools and databases in the current work, there are numerous others online and offline tools that could not be included because of their high number and redundancy of information. In bioinformatics, there is a continuous need for new tools and additional functionality which makes the review of new tools a hard task. As one can see, information is shared between platforms, databases, and databanks in the spirit of scientific collaboration and the pursuit of new knowledge.

In this review, we introduced tools that are needed for starting an analysis of genomic data from a high level (disease or phenotype), ending with the reconstruction of networks of interactions for ncRNAs and specifically short non-coding molecules. We did not get into methodologies of analysis of the data which is a whole field of study alone and needs special focus. The presented tools, even though not oriented only in NDDs, can be used to identify the common molecular pathways in these disorders and the comorbidity that is often present in preterm babies with NDDs.

A. Abbreviations

- NDD : neurodevelopmental disorder
- CP : ceribral palsy
- MRI : magnetic resonance imaging
- 3'UTR : 3' (prime) untranslated region
- ADS : autism disorder spectrum
- NGS : next generation sequencing
- PIWI : P-element induced wimpy testis
- mRNA : messenger RNA
- ncRNA : non-coding RNA
- sncRNA : short/small non-coding RNA
- miRNA : micro RNA
- siRNA : small interference RNA
- piRNA : P-element-induced wimpy testis-interacting RNA (piwi RNA)
- RNAi : RNA interference
- ceRNA : competitive endogenous RNA
- MRE : miRNA response element

References

[1] R. L. Goldenberg, J. F. Culhane, J. D. Iams, R. Romero, Epidemiology and causes of preterm birth, The Lancet 371 (2008) 75–84. URL: https://doi.org/10.1016/S0140-6736(08)60074-4. doi:10.1016/S0140-6736(08)60074-4, doi: 10.1016/S0140-6736(08)60074-4.

[2] H. Blencowe, S. Cousens, M. Z. Oestergaard, D. Chou, A.-B. Moller, R. Narwal, A. Adler, C. V. Garcia, S. Rohde, L. Say, J. E. Lawn, National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications, The Lancet 379 (2012) 2162–2172. URL: https://www.sciencedirect.com/science/article/pii/S0140673612608204. doi:https://doi.org/10.1016/S0140-6736(12)60820-4.

[3] H.-L. Li, L.-L. Chen, G. Qian, W.-S. Wang, X.-N. Feng, M.-Q. Yu, L.-Z. Ren, A birth cohort study of neurodevelopmental outcomes from birth to 2 years of age in preterm infants under 34 weeks of gestation, Journal of Healthcare Engineering 2022 (2022) 7496785. URL: https://doi.org/10.1155/2022/7496785. doi:10.1155/2022/7496785.

[4] D. Dewey, What is comorbidity and why does it matter in neurodevelopmental disorders?, Current Developmental Disorders Reports 5 (2018) 235–242. URL: https://doi.org/10.1007/s40474-018-0152-3. doi:10.1007/s40474-018-0152-3.

[5] D. J. Zgaljardic, R. O. Temple, Neuropsychological assessment battery (nab): Performance in a sample of patients with moderate-to-severe traumatic brain injury, Applied Neuropsychology 17 (2010) 283–288. URL: https://doi.org/10.1080/09084282.2010.525118. doi:10.1080/09084282.2010.525118, doi: 10.1080/09084282.2010.525118.

[6] L. E. L. M. Vissers, C. Gilissen, J. A. Veltman, Genetic studies in intellectual disability and related disorders, Nature Reviews Genetics 17 (2016) 9–18. URL: https://doi.org/10.1038/nrg3999. doi:10.1038/nrg3999.

[7] A. H. MacLennan, S. C. Thompson, J. Gecz, Cerebral palsy: causes, pathways, and the role of genetic variants, American Journal of Obstetrics & Gynecology 213 (2015) 779–788. URL: https://doi.org/10.1016/j.ajog.2015.05.034. doi:10.1016/j.ajog.2015.05.034, doi: 10.1016/j.ajog.2015.05.034.

[8] J. H. Thygesen, K. Wolfe, A. McQuillin, M. Viñas-Jornet, N. Baena, N. Brison, G. D'Haenens, S. Esteba-Castillo, E. Gabau, N. Ribas-Vidal, A. Ruiz, J. Vermeesch, E. Weyts, R. Novell, G. V. Buggenhout, A. Strydom, N. Bass, M. Guitart, A. Vogels, Neurodevelopmental risk copy number variants in adults with intellectual disabilities and comorbid psychiatric disorders, British Journal of Psychiatry 212 (2018) 287–294. doi:DOI:10.1192/bjp.2017.65.

[9] M. K. Licari, A. Finlay-Jones, J. E. Reynolds, G. A. Alvares, A. J. Spittle, J. Downs, A. J. O. Whitehouse, H. Leonard, K. L. Evans, K. Varcin, The brain basis of comorbidity in neurodevelopmental disorders, Current Developmental Disorders Reports 6 (2019) 9–18. URL: https://doi.org/10.1007/s40474-019-0156-7. doi:10.1007/s40474-019-0156-7.

[10] P. Zhang, W. Wu, Q. Chen, M. Chen, Non-coding RNAs and their integrated networks, Journal of Integrative Bioinformatics 16 (2019) 20190027. URL: https://doi.org/10.1515/jib-2019-0027. doi:doi:10.1515/jib-2019-0027.

[11] F. CRICK, Central dogma of molecular biology, Nature 227 (1970) 561–563. URL: https://doi.org/10.1038/227561a0. doi:10.1038/227561a0.

[12] I. D. et al., An integrated encyclopedia of DNA elements in the human genome, Nature

489 (2012) 57–74. URL: https://doi.org/10.1038/nature11247. doi:10.1038/nature11247.

[13] A. S. Cristino, S. M. Williams, Z. Hawi, J.-Y. An, M. A. Bellgrove, C. E. Schwartz, L. da F Costa, C. Claudianos, Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system, Molecular Psychiatry 19 (2014) 294–301. URL: https://doi.org/10.1038/mp.2013.16. doi:10.1038/mp.2013.16.

[14] R. A. Clarke, S. Lee, V. Eapen, Pathogenetic model for tourette syndrome delineates overlap with related neurodevelopmental disorders including autism, Translational Psychiatry 2 (2012) e158–e158. URL: https://doi.org/10.1038/tp.2012.75. doi:10.1038/tp.2012.75.

[15] M. Sahin, M. Sur, Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders, Science 350 (2015) aab3897. URL: https://www.science.org/doi/abs/10.1126/science.aab3897. doi:10.1126/science.aab3897. arXiv:https://www.science.org/doi/pdf/10.1126/science.aab3897.

[16] I. V. Novikova, S. P. Hennelly, K. Y. Sanbonmatsu, Sizing up long non-coding RNAs: do lncRNAs have secondary and tertiary structure?, Bioarchitecture 2 (2012) 189–199. URL: https://pubmed.ncbi.nlm.nih.gov/23267412https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3527312/. doi:10.4161/bioa.22592.

[17] A. Mahfouz, M. N. Ziats, O. M. Rennert, B. P. F. Lelieveldt, M. J. T. Reinders, Shared pathways among autism candidate genes determined by co-expression network analysis of the developing human brain transcriptome, Journal of molecular neuroscience : MN 57 (2015) 580–594. URL: https://pubmed.ncbi.nlm.nih.gov/26399424https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4644211/. doi:10.1007/s12031-015-0641-3.

[18] S. F. Zhang, J. Gao, C. M. Liu, The role of non-coding RNAs in neurodevelopmental disorders, 2019. doi:10.3389/fgene.2019.01033.

[19] S. K. Fineberg, K. S. Kosik, B. L. Davidson, MicroRNAs potentiate neural development, Neuron 64 (2009) 303–309. URL: https://www.sciencedirect.com/science/article/pii/S0896627309008502. doi:https://doi.org/10.1016/j.neuron.2009.10.020.

[20] C. N. Watson, A. Belli, V. D. Pietro, Small non-coding RNAs: New class of biomarkers and potential therapeutic targets in neurodegenerative disease, Frontiers in Genetics 10 (2019). URL: https://www.frontiersin.org/articles/10.3389/fgene.2019.00364. doi:10.3389/fgene.2019.00364.

[21] J. O'Brien, H. Hayder, Y. Zayed, C. Peng, Overview of microRNA biogenesis, mechanisms of actions, and circulation, Frontiers in Endocrinology 9 (2018). URL: https://www.frontiersin.org/article/10.3389/fendo.2018.00402. doi:10.3389/fendo.2018.00402.

[22] M. Muñoz-Culla, H. Irizar, M. Sáenz-Cuesta, T. Castillo-Triviño, I. Osorio-Querejeta, L. Sepúlveda, A. L. de Munain, J. Olascoaga, D. Otaegui, SncRNA (microRNA & snoRNA) opposite expression pattern found in multiple sclerosis relapse and remission is sex dependent, Scientific Reports 6 (2016) 20126. URL: https://doi.org/10.1038/srep20126. doi:10.1038/srep20126.

[23] O. Noronha, L. Mesarosovo, J. J. Anink, A. Iyer, E. Aronica, J. D. Mills, Differentially expressed miRNAs in age-related neurodegenerative diseases: A meta-analysis, Genes 13 (2022). URL: https://www.mdpi.com/2073-4425/13/6/1034. doi:10.3390/genes13061034.

[24] C. P. da Costa Gomes, B. Schroen, G. M. Kuster, E. L. Robinson, K. Ford, I. B. Squire, S. Heymans, F. Martelli, C. Emanueli, Y. Devaux, null null, Regulatory rnas in heart

failure, Circulation 141 (2020) 313–328. URL: https://doi.org/10.1161/CIRCULATIONAHA.119.042474. doi:10.1161/CIRCULATIONAHA.119.042474, doi: 10.1161/CIRCULATIONAHA.119.042474.

[25] E. S. Smith, E. Whitty, B. Yoo, A. Moore, L. F. Sempere, Z. Medarova, Clinical applications of short non-coding RNA-based therapies in the era of precision medicine, Cancers 14 (2022). URL: https://www.mdpi.com/2072-6694/14/6/1588. doi:10.3390/cancers14061588.

[26] Y. Peng, C. M. Croce, The role of microRNAs in human cancer, Signal Transduction and Targeted Therapy 1 (2016) 15004. URL: https://doi.org/10.1038/sigtrans.2015.4. doi:10.1038/sigtrans.2015.4.

[27] R. Bardini, G. Politano, A. Benso, S. D. Carlo, Multi-level and hybrid modelling approaches for systems biology, Computational and Structural Biotechnology Journal 15 (2017) 396–402. URL: https://www.sciencedirect.com/science/article/pii/S2001037017300314. doi:https://doi.org/10.1016/j.csbj.2017.07.005.

[28] Z. Ji, K. Yan, W. Li, H. Hu, X. Zhu, Mathematical and computational modeling in complex biological systems, BioMed Research International 2017 (2017) 5958321. URL: https://doi.org/10.1155/2017/5958321. doi:10.1155/2017/5958321.

[29] A. Oulas, N. Karathanasis, A. Louloupi, G. Pavlopoulos, P. Poirazi, K. Kalantidis, I. Iliopoulos, Prediction of miRNA targets, 2015. doi:10.1007/978-1-4939-2291-8_13.

[30] A. Korfiati, K. A. Theofilatos, D. Kleftogiannis, C. E. Alexakos, S. D. Likothanassis, S. Mavroudi, Predicting human miRNA target genes using a novel computational intelligent framework, Inf. Sci. 294 (2015) 576–585.

[31] M. N. Ziats, O. M. Rennert, Identification of differentially expressed microrna across the developing human brain, Molecular Psychiatry 19 (2014) 848–852. URL: https://doi.org/10.1038/mp.2013.93. doi:10.1038/mp.2013.93.

[32] H. J. Kang, Y. I. Kawasawa, F. Cheng, Y. Zhu, X. Xu, M. Li, A. M. M. Sousa, M. Pletikos, K. A. Meyer, G. Sedmak, T. Guennel, Y. Shin, M. B. Johnson, Željka Krsnik, S. Mayer, S. Fertuzinhos, S. Umlauf, S. N. Lisgo, A. Vortmeyer, D. R. Weinberger, S. Mane, T. M. Hyde, A. Huttner, M. Reimers, J. E. Kleinman, N. Šestan, Spatio-temporal transcriptome of the human brain, Nature 478 (2011) 483–489. URL: https://doi.org/10.1038/nature10523. doi:10.1038/nature10523.

[33] P. Paul, A. Chakraborty, D. Sarkar, M. Langthasa, M. Rahman, M. Bari, R. K. S. Singha, A. K. Malakar, S. Chakraborty, Interplay between miRNAs and human diseases, Journal of Cellular Physiology 233 (2018) 2007–2018. URL: https://doi.org/10.1002/jcp.25854. doi:https://doi.org/10.1002/jcp.25854, https://doi.org/10.1002/jcp.25854.

[34] H. Lv, Q. Wang, S. Wu, L. Yang, P. Ren, Y. Yang, J. Gao, L. Li, Neonatal hypoxic ischemic encephalopathy-related biomarkers in serum and cerebrospinal fluid, Clinica Chimica Acta 450 (2015) 282–297. doi:10.1016/j.cca.2015.08.021.

[35] M. Ha, V. N. Kim, Regulation of microRNA biogenesis, Nature Reviews Molecular Cell Biology 15 (2014) 509–524. URL: https://doi.org/10.1038/nrm3838. doi:10.1038/nrm3838.

[36] J. K. W. Lam, M. Y. T. Chow, Y. Zhang, S. W. S. Leung, siRNA versus miRNA as therapeutics for gene silencing, Molecular Therapy - Nucleic Acids 4 (2015). URL: https://doi.org/10.1038/mtna.2015.23. doi:10.1038/mtna.2015.23, doi: 10.1038/mtna.2015.23.

[37] C. B. Assumpção, D. Q. Calcagno, T. M. T. Araújo, S. E. B. dos Santos, Ândrea Kely Campos Ribeiro dos Santos, G. J. Riggins, R. R. Burbano, P. P. Assumpção, The role of piRNA

and its potential clinical implications in cancer, Epigenomics 7 (2015) 975–984. URL: https://doi.org/10.2217/epi.15.37. doi:10.2217/epi.15.37, doi: 10.2217/epi.15.37.

[38] S. Peterson, J. Thompson, M. Ufkin, P. Sathyanarayana, L. Liaw, C. B. Congdon, Common features of microRNA target prediction tools, Frontiers in Genetics 5 (2014). URL: https://www.frontiersin.org/articles/10.3389/fgene.2014.00023. doi:10.3389/fgene.2014.00023.

[39] K. T. Thomas, S. S. Zakharenko, MicroRNAs in the onset of schizophrenia, Cells 10 (2021). URL: https://www.mdpi.com/2073-4409/10/10/2679. doi:10.3390/cells10102679.

[40] S. Chang, S. Wen, D. Chen, P. Jin, Small regulatory RNAs in neurodevelopmental disorders, Human Molecular Genetics 18 (2009) R18–R26. URL: https://doi.org/10.1093/hmg/ddp072. doi:10.1093/hmg/ddp072.

[41] M. Costa-Mattioli, L. M. Monteggia, mTOR complexes in neurodevelopmental and neuropsychiatric disorders, Nature Neuroscience 16 (2013) 1537–1543. URL: https://doi.org/10.1038/nn.3546. doi:10.1038/nn.3546.

[42] I. Parenti, L. G. Rabaneda, H. Schoen, G. Novarino, Neurodevelopmental disorders: From genetics to functional pathways, Trends in Neurosciences 43 (2020) 608–621. URL: https://doi.org/10.1016/j.tins.2020.05.004. doi:10.1016/j.tins.2020.05.004, doi: 10.1016/j.tins.2020.05.004.

[43] G. Jain, A. Stuendl, P. Rao, T. Berulava, T. P. Centeno, L. Kaurani, S. Burkhardt, I. Delalle, J. Kornhuber, M. Hüll, W. Maier, O. Peters, H. Esselmann, C. Schulte, C. Deuschle, M. Synofzik, J. Wiltfang, B. Mollenhauer, W. Maetzler, A. Schneider, A. Fischer, A combined miRNA–piRNA signature to detect alzheimer's disease, Translational Psychiatry 9 (2019) 250. URL: https://doi.org/10.1038/s41398-019-0579-2. doi:10.1038/s41398-019-0579-2.

[44] M. A. Obeid, A. A. A. Aljabali, W. Alshaer, N. B. Charbe, D. K. Chellappan, K. Dua, S. Satija, M. M. Tambuwala, Chapter 35 - targeting siRNAs in cancer drug delivery, 2021. URL: https://www.sciencedirect.com/science/article/pii/B9780323855037000274. doi:https://doi.org/10.1016/B978-0-323-85503-7.00027-4.

[45] L. H. Madkour, Chapter 5 - therapeutic applications of siRNA gene delivery systems, 2019. URL: https://www.sciencedirect.com/science/article/pii/B9780128197776000056. doi:https://doi.org/10.1016/B978-0-12-819777-6.00005-6.

[46] T. N. Turner, E. E. Eichler, The role of de novo noncoding regulatory mutations in neurodevelopmental disorders, Trends in Neurosciences 42 (2019) 115–127. URL: https://www.sciencedirect.com/science/article/pii/S0166223618302960. doi:https://doi.org/10.1016/j.tins.2018.11.002.

[47] A. D. Baxevanis, The importance of biological databases in biological discovery, Current Protocols in Bioinformatics 34 (2011) 1.1.1–1.1.6. URL: https://doi.org/10.1002/0471250953.bi0101s34. doi:https://doi.org/10.1002/0471250953.bi0101s34, https://doi.org/10.1002/0471250953.bi0101s34.

[48] T. Barrett, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L. Robertson, N. Serova, S. Davis, A. Soboleva, NCBI GEO: archive for functional genomics data sets—update, Nucleic Acids Research 41 (2013) D991–D995. URL: https://doi.org/10.1093/nar/gks1193. doi:10.1093/nar/gks1193.

[49] N. R. Coordinators, Database resources of the national center for biotechnology information, Nucleic Acids Research 44 (2016) D7–D19. URL: https://doi.org/10.1093/nar/gkv1290. doi:10.1093/nar/gkv1290.

[50] A. Athar, A. Füllgrabe, N. George, H. Iqbal, L. Huerta, A. Ali, C. Snow, N. A. Fonseca, R. Petryszak, I. Papatheodorou, U. Sarkans, A. Brazma, ArrayExpress update – from bulk to single-cell expression data, Nucleic Acids Research 47 (2018) D711–D715. URL: https://doi.org/10.1093/nar/gky964. doi:10.1093/nar/gky964.

[51] S. Griffiths-Jones, The microRNA registry, Nucleic Acids Research 32 (2004) D109–D111. URL: https://doi.org/10.1093/nar/gkh023. doi:10.1093/nar/gkh023.

[52] S. Griffiths-Jones, H. K. Saini, S. van Dongen, A. J. Enright, miRBase: tools for microRNA genomics, Nucleic Acids Research 36 (2008) D154–D158. URL: https://doi.org/10.1093/nar/gkm952. doi:10.1093/nar/gkm952.

[53] H.-Y. Huang, Y.-C.-D. Lin, J. Li, K.-Y. Huang, S. Shrestha, H.-C. Hong, Y. Tang, Y.-G. Chen, C.-N. Jin, Y. Yu, J.-T. Xu, Y.-M. Li, X.-X. Cai, Z.-Y. Zhou, X.-H. Chen, Y.-Y. Pei, L. Hu, J.-J. Su, S.-D. Cui, F. Wang, Y.-Y. Xie, S.-Y. Ding, M.-F. Luo, C.-H. Chou, N.-W. Chang, K.-W. Chen, Y.-H. Cheng, X.-H. Wan, W.-L. Hsu, T.-Y. Lee, F.-X. Wei, H.-D. Huang, miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database, Nucleic Acids Research 48 (2020) D148–D154. URL: https://doi.org/10.1093/nar/gkz896. doi:10.1093/nar/gkz896.

[54] C. Backes, T. Fehlmann, F. Kern, T. Kehl, H.-P. Lenhof, E. Meese, A. Keller, miRCarta: a central repository for collecting miRNA candidates, Nucleic Acids Research 46 (2018) D160–D167. URL: https://doi.org/10.1093/nar/gkx851. doi:10.1093/nar/gkx851.

[55] T. Fehlmann, F. Kern, O. Laham, C. Backes, J. Solomon, P. Hirsch, C. Volz, R. Müller, A. Keller, mirmaster 2.0: multi-species non-coding rna sequencing analyses at scale, Nucleic Acids Research 49 (2021) W397–W408. URL: https://doi.org/10.1093/nar/gkab268. doi:10.1093/nar/gkab268.

[56] C.-J. Liu, G.-Y. Xie, Y.-R. Miao, M. Xia, Y. Wang, Q. Lei, Q. Zhang, A.-Y. Guo, EVAtlas: a comprehensive database for ncRNA expression in human extracellular vesicles, Nucleic Acids Research 50 (2022) D111–D117. URL: https://doi.org/10.1093/nar/gkab668. doi:10.1093/nar/gkab668.

[57] L. Chang, G. Zhou, O. Soufan, J. Xia, miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology, Nucleic Acids Research 48 (2020) W244–W251. URL: https://doi.org/10.1093/nar/gkaa467. doi:10.1093/nar/gkaa467.

[58] A. Fabregat, K. Sidiropoulos, P. Garapati, M. Gillespie, K. Hausmann, R. Haw, B. Jassal, S. Jupe, F. Korninger, S. McKay, L. Matthews, B. May, M. Milacic, K. Rothfels, V. Shamovsky, M. Webber, J. Weiser, M. Williams, G. Wu, L. Stein, H. Hermjakob, P. D'Eustachio, The reactome pathway knowledgebase, Nucleic Acids Research 44 (2016) D481–D487. URL: https://doi.org/10.1093/nar/gkv1351. doi:10.1093/nar/gkv1351.

[59] T. Kehl, F. Kern, C. Backes, T. Fehlmann, D. Stöckel, E. Meese, H.-P. Lenhof, A. Keller, miRPathDB 2.0: a novel release of the miRNA pathway dictionary database, Nucleic Acids Research 48 (2020) D142–D147. URL: https://doi.org/10.1093/nar/gkz1022. doi:10.1093/nar/gkz1022.

[60] J. Gong, D. Shao, K. Xu, Z. Lu, Z. J. Lu, Y. T. Yang, Q. C. Zhang, Rise: a database of rna interactome from sequencing experiments, Nucleic Acids Research 46 (2018) D194–D201.

URL: https://doi.org/10.1093/nar/gkx864. doi:10.1093/nar/gkx864.

[61] X. Teng, X. Chen, H. Xue, Y. Tang, P. Zhang, Q. Kang, Y. Hao, R. Chen, Y. Zhao, S. He, NPInter v4.0: an integrated database of ncRNA interactions, Nucleic Acids Research 48 (2020) D160–D165. URL: https://doi.org/10.1093/nar/gkz969. doi:10.1093/nar/gkz969.

[62] J.-H. Li, S. Liu, H. Zhou, L.-H. Qu, J.-H. Yang, starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data, Nucleic Acids Research 42 (2014) D92–D97. URL: https://doi.org/10.1093/nar/gkt1248. doi:10.1093/nar/gkt1248.

[63] V. Agarwal, G. W. Bell, J.-W. Nam, D. P. Bartel, Predicting effective microRNA target sites in mammalian mRNAs, eLife 4 (2015) e05005. URL: https://doi.org/10.7554/eLife.05005. doi:10.7554/eLife.05005.

[64] M. D. Paraskevopoulou, G. Georgakilas, N. Kostoulas, I. S. Vlachos, T. Vergoulis, M. Reczko, C. Filippidis, T. Dalamagas, A. G. Hatzigeorgiou, DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows, Nucleic Acids Research 41 (2013) W169–W173. URL: https://doi.org/10.1093/nar/gkt393. doi:10.1093/nar/gkt393.

[65] D. Karagkouni, M. D. Paraskevopoulou, S. Chatzopoulos, I. S. Vlachos, S. Tastsoglou, I. Kanellos, D. Papadimitriou, I. Kavakiotis, S. Maniou, G. Skoufos, T. Vergoulis, T. Dalamagas, A. G. Hatzigeorgiou, DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA–gene interactions, Nucleic Acids Research 46 (2018) D239–D245. URL: https://doi.org/10.1093/nar/gkx1141. doi:10.1093/nar/gkx1141.

[66] F. Xiao, Z. Zuo, G. Cai, S. Kang, X. Gao, T. Li, miRecords: an integrated resource for microRNA–target interactions, Nucleic Acids Research 37 (2009) D105–D110. URL: https://doi.org/10.1093/nar/gkn851. doi:10.1093/nar/gkn851.

[67] Q. Jiang, Y. Wang, Y. Hao, L. Juan, M. Teng, X. Zhang, M. Li, G. Wang, Y. Liu, miR2Disease: a manually curated database for microRNA deregulation in human disease, Nucleic Acids Research 37 (2009) D98–D104. URL: https://doi.org/10.1093/nar/gkn714. doi:10.1093/nar/gkn714.

[68] P. N. Deepthi, R. Anitha, Applications of network analysis in bioinformatics, Springer International Publishing, 2020, pp. 79–84.

[69] A. Mrvar, V. Batagelj, Analysis and visualization of large networks with program package pajek, Complex Adaptive Systems Modeling 4 (2016) 6. URL: https://doi.org/10.1186/s40294-016-0017-8. doi:10.1186/s40294-016-0017-8.

[70] M. Bastian, S. Heymann, M. Jacomy, Gephi: An open source software for exploring and manipulating networks, Proceedings of the International AAAI Conference on Web and Social Media 3 (2009) 361–362. URL: https://ojs.aaai.org/index.php/ICWSM/article/view/13937.

[71] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks, Genome Research 13 (2003) 2498–2504. doi:10.1101/gr.1239303.

[72] K. Theofilatos, C. Dimitrakopoulos, C. Alexakos, A. Korfiati, S. Likothanassis, S. Mavroudi, InSyBio BioNets: an efficient tool for network-based biomarker discovery, EMBnet.journal 22 (2016) 871. doi:10.14806/ej.22.0.871.