Chapter 4
Transcriptomics of Neonatal and Infant Human Thymus

Carlos Alberto Moreira-Filho, Silvia Yumi Bando, Fernanda Bernardi Bertonha, and Magda Carneiro-Sampaio

4.1 Introduction

The thymus is a primary lymphoid organ where immunocompetent and self-tolerant T cells are produced (Kondo et al. 2019; Miller 2020). Anatomic descriptions of the thymus date back to Rufus of Ephesus and Galen, in 1st–2nd c A.D (Laios 2018). Yet, it was only in the 1970s and 1980s – with the discovery of the basic cellular and molecular mechanisms for generating B and T cell antigen receptor diversity, and for the elimination of self-reactive T cells in the thymic stroma – that our current understanding of thymic functions became established (reviewed in Geenen 2021). Presently, remarkable progresses in genomics and systems biology – mainly on single-cell RNA sequencing (scRNA-seq) methodologies and on computational analysis of genomic and proteomic big data – have allowed the construction of detailed atlases of thymic organogenesis and development (Kernfeld et al. 2018; Ohigashi et al. 2019; Park et al. 2020; Hao et al. 2021). By using scRNA-seq and spatial cell localization techniques, Park et al. (2020) constructed a temporal scenario of the human thymus cell state dynamic changes, spanning from organ development to pediatric and adult life. The human thymus involutes very early in life (Steinmann 1986; Rezzani et al. 2014) and its functioning suffers the influence of sex hormones (Dragin et al. 2016; Berrih-Aknin et al. 2018; Merrheim et al. 2020). Therefore, the thymus early programming – sexual dimorphism, dynamics of thymocyte populations, and thymic microenvironment change – influences immune activity throughout life. Here, we used transcriptome analysis to address two key issues in early thymus development: the thymic sexual dimorphism during the first

C. A. Moreira-Filho (✉) · S. Y. Bando · F. B. Bertonha · M. Carneiro-Sampaio
Departamento de Pediatria, Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
e-mail: carlos.moreira@hc.fm.usp.br; silvia.bando@hc.fm.usp.br; fernanda.bernardi@fm.usp.br; magdacs@usp.br

G. A. Passos (ed.), *Transcriptomics in Health and Disease*,
https://doi.org/10.1007/978-3-030-87821-4_4

6 months of age (i.e., during minipuberty), and the genomic mechanisms governing cellular and molecular processes involved in the functioning of the neonate thymus and in the onset of thymic decline.

4.2 Methodology

Fresh thymic explants (corticomedullar sections) from karyotypically normal neonates and infants who underwent cardiac surgery were collected at surgery room and immediately preserved for total RNA extraction. The resected thymic specimens were preserved in formalin and paraffin-embedded for histological analyses. DNA microarray technology was employed to obtain mRNA and miRNA expression profiles. Transcriptomics was performed in whole thymic tissue to avoid gene expression artifacts caused by mechanic/enzymatic tissue dissociation. Whole thymic tissue transcriptome datasets were interpreted through modular repertoire identification, a community detection network analysis (Barabási and Oltvai 2004; Zhu et al. 2007; Barabási et al. 2011; Chaussabel and Baldwin 2014; Gaiteri et al. 2014; van Dam et al. 2018) used, for instance, in the investigation of immune responses in vivo following the administration of vaccines (Nakaya et al. 2011; Obermoser et al. 2013) and that has proved to be a suitable strategy to circumvent tissue microdissection and cell separation (Moreira-Filho et al. 2015, 2016; Bando et al. 2021).

For the investigation of thymic sexual dimorphism, we adopted a network-based approach for gene co-expression (GCN) analysis that allows the identification of modular transcriptional repertoires (communities) and the interactions between all the system's constituents through community detection (Chaussabel and Baldwin 2014; Moreira-Filho et al. 2016). MiRNA-target analysis was used to investigate how the abundantly expressed thymic miRNAs modulate the expression of highly connected genes (hubs) in the GCNs. In the study of neonatal and infant thymus we employed a weighted gene co-expression network analysis (WGCNA) (Langfelder and Horwath 2008) for describing correlation patterns among genes across microarray datasets that allows: (i) the identification of transcriptional modules (van Dam et al. 2018) and their association with specific age groups; (ii) the identification of highly connected genes (hubs) and of significant genes (HGS genes) for the trait of interest (age). This analysis was complemented by an integrative mRNA–miRNA–transcription factor (TF) co-expression analysis encompassing mRNAs from hubs and HGS genes, the abundantly expressed miRNAs, and the TFs that covaried with hubs and/or HGS genes. The methodological framework employed these two studies is detailed in chapter 6. The workflow of experimental approaches and bioinformatic analyses is depicted in Fig. 4.1.

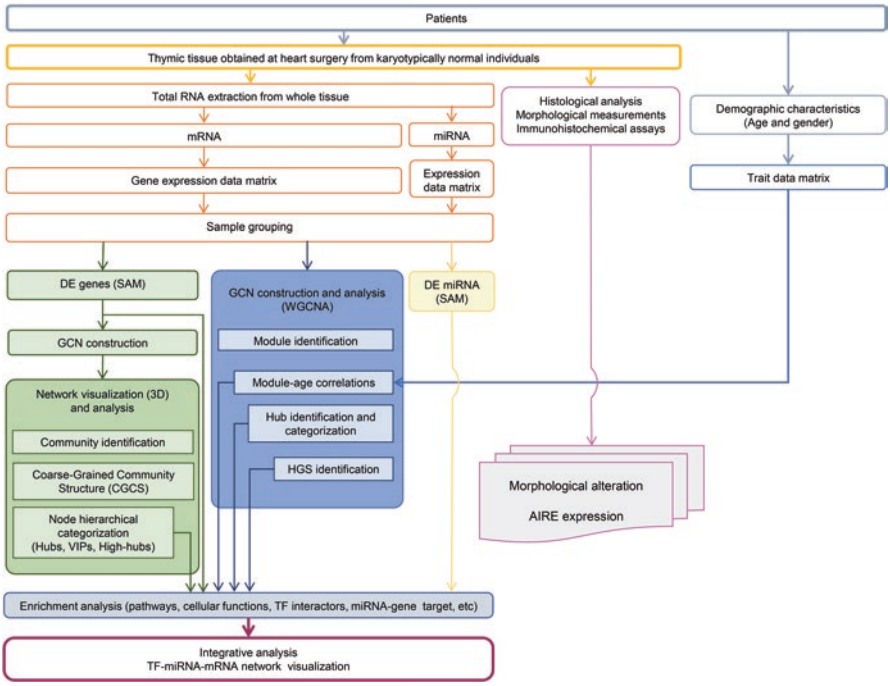

Fig. 4.1 Systems biology workflow used for generating and integrating data on neonatal and infant thymus development

4.3 Human Thymus During Minipuberty

Sexual dimorphism in the immune system is well documented in humans – as well as in other mammals and birds – encompassing sex differences in responses to self and foreign antigens: women usually mount stronger immune responses to infections and vaccination but have higher susceptibility to autoimmune diseases than men (Klein and Flanagan 2016). Regarding autoimmune diseases, it is striking that in USA almost 80% of autoimmune patients are women (Billi et al. 2019). Autoimmunity results from a tolerance breakdown and essentially involves the thymus, the site of T cell selection (Cheng and Anderson 2018). T cell selection depends on the ectopic thymic transcription of thousands of genes coding for tissue-specific antigens, which is induced by the autoimmune regulator gene *AIRE* (Passos et al. 2018; Perniola 2018). Despite our incomplete knowledge on the biological processes responsible for autoimmunity, it would be reasonable to assume that sex hormones impact the genomic mechanisms governing *AIRE* functions and T cell selection. An important experimental evidence regarding this assumption came from the work of Dumont-Lagacé et al. (2015), who showed in a murine model that sex hormones have pervasive effects on thymic epithelial cells (TEC) – antigen presenting cells that regulate T cell repertoire and tolerance – and that androgens have a greater impact on TEC

transcriptome than estrogens. Interestingly, the authors observed that sex steroids repressed the expression of tissue-restricted antigens but did not alter the expression of *Aire*. Just after this work, Dragin et al. (2016) demonstrated that estrogen mediates the downregulation of *AIRE* in human pubescent and adult thymic tissues, thus indicating that the reduced expression of AIRE protein in women may be related to autoimmunity susceptibility. However, this study did not cover infants along the first 6 months of age, i.e., during minipuberty (Kuiri-Hänninen et al. 2014; Becker and Hesse 2020), a period when sex hormones conceivably act on thymic tissue.

To further investigate the presumptive sexual dimorphism induced by minipuberty on infant thymus, we performed comparative genomic and immunohistochemical studies on thymic surgical explants obtained from karyotypically normal male (M) and female (F) infants during minipuberty, here termed MM and MF groups, and from karyotypically normal M and F non-puberty (N) infants aged 7–31 months, the NM and NF groups. Analyses included gene co-expression networks (GCN) for differentially expressed genes, miRNA-target analyses, *AIRE*-centered gene–gene interaction networks encompassing the genes coding for AIRE interactors, quantitative RT-qPCR and immunohistochemical measurements of AIRE expression, and comparative thymic histomorphometry. GCN analysis was performed for the identification of high-hierarchical genes, modular transcriptional repertoires (communities), and the interactions between all the system's constituents through community detection (Moreira-Filho et al. 2016). MiRNA-target analysis was used to investigate how the abundantly expressed thymic miRNAs modulate the expression of highly connected genes (hubs) in the GCNs. This study is fully described in Moreira-Filho et al. (2018).

4.4 Thymic Gene Expression in Minipuberty and Non-puberty Infants: mRNA, miRNA, and *AIRE* Interactors

4.4.1 Sample Grouping

Thymic samples from 17 patients aged up to 6 months were classified as minipuberty (M) – ten males and seven females, here termed MM and MF, respectively – and 17 samples from patients aged 7–17 months were classified as non-puberty (N), being nine males and eight females, here termed NM and NF, respectively.

4.4.2 Global Gene Expression and miRNA Expression in Minipuberty and Non-puberty Groups

DNA microarray technology was used to obtain mRNA and miRNA expression profiles in minipuberty (M) and non-puberty (N) groups. The mRNA expression data was analyzed by SAM test (Tusher et al. 2001) for determining the

differentially expressed (DE) genes. In the MM *vs.* MF group comparison 494 DE genes were identified, all being hyper-expressed in the MM group. No DE genes were found in the NM *vs.* NF group comparison. Hence, we conducted DE network analysis only for MM and MF groups. Statistical analysis (t-test) for miRNA expression data revealed 16 abundantly expressed miRNAs in the M group, being all hyper-expressed in the MF group (female infants). In the non-puberty groups, 20 abundantly expressed miRNAs (15 of which were also abundantly expressed in the minipuberty groups) and all hyper-expressed in the NM group (male infants). The abundantly expressed miRNAs for minipuberty and non-puberty groups were selected after analyzing miRNA expression value distribution through a scatter dot plot and adopting abundant expression cut-off values for MM and MF groups, and for NM and NF groups, respectively. The fold change values were calculated as the ratio of the average expression value of each abundantly expressed miRNA to the average expression value of the non-abundantly expressed miRNAs for each group.

4.4.3 Gene Co-expression Network Construction and Analysis

Differentially expressed (DE) GO annotated gene co-expression networks were constructed for MM and MF groups based on gene–gene Pearson's correlation method and the Networks 3D software developed by Luciano Costa's Research Group, Institute of Physics at São Carlos, University of São Paulo (Bando et al. 2013). This package allowed the categorization of network nodes according to distinct hierarchical levels of gene–gene connections: hubs are highly connected nodes, VIPs have low node degree but connect only with hubs, and high-hubs have VIP status and high overall number of connections. We classified network nodes as VIPs, hubs, or high-hubs by obtaining the node degree, k_0, and the first level concentric node degree, k_1, which takes into account all node connections leaving from its immediate neighborhood, then projecting all node values in a k_0 vs k_1 graphic (see Chap. 6 and Bando et al. 2013).

The GCN topologies were then analyzed for identifying their community structure. Communities can encompass complex mechanisms that work together to maintain the cellular processes across different conditions (Barabasi et al. 2011; Gaiteri et al. 2014; van Dam et al. 2018). For example, community structure analysis of gene co-expression networks obtained from skeletal muscle cells' transcriptome revealed different biological pathways for Duchenne muscular dystrophy patients comparatively to normal individuals (Narayanan and Subramaniam 2013). The same modular approach has successfully been used for investigating immune response to infections and vaccines using whole blood transcriptome data sets (reviewed in Chaussabel and Baldwin 2014), and for characterizing thymic gene dysregulation in 21 trisomy patients (Moreira-Filho et al. 2016).

Community detection in complex networks is usually accomplished by discovering the network modular structure that optimizes the modularity measurement. Modularity considers the relationship between the number of links inside a

community compared to connections between nodes in distinct communities (Newman and Girvan 2004; Newman 2010). A diverse range of optimization techniques exists to optimize the modularity. Here we applied the method proposed by Blondel et al. (2008), which attains good modularity values and at the same time presents excellent performance.

Most of the methods for community detection generate hierarchical structures. The Newman-Girvan method uses the edge betweenness centrality measurement as a criterion for removing edges and obtaining connected components that correspond to each network partition. This builds a tree of communities with branches occurring every time a component is divided into two. Agglomerative methods start from a set of communities, where each node corresponds to a different community, which are progressively merged according to a similarity criterion or to directly maximize the change of modularity (Clauset et al. 2004). In both cases, a dendrogram of the partition hierarchy is obtained. The optimal set of communities is then obtained by a cut for the highest value of modularity.

Figure 4.2a, b depicts the two minipuberty networks, MM-DE and MF-DE, their gene communities (modules), and the high-hierarchy genes for each network. Different node colors identify the distinct gene communities in each network. Modularity values and the number of communities in each network were quite close: 0.728 and 15 communities in the MM-DE and 0.649 and 16 communities in MF-DE (Moreira-Filho et al. 2018). Coarse-grained community structure (CGCS) was obtained for each DE network, disclosing the relationships between each community in the network (Fig. 4.2c, d) for MM-DE and MF-DE, respectively). Communities having the highest node strength (total probability for community's nodes to connect to distinct communities) hold the most significant functional interactions in the network (Chaussabel and Baldwin 2014).

The integrative network analyses between abundantly expressed miRNAs and target high-hierarchy genes (HH) from MM-DE and MF-DE networks appear in Fig. 4.2a, b. It is worth to note that all miRNAs interacting with HH genes in the MM-DE and MF-DE networks play important roles in the regulation of immune processes, and particularly in the thymic environment. Let-7 miRNAs regulate NKT cell differentiation (Pobezinsky et al. 2015). The cluster miR15/16 enhances the induction of regulatory T-cells by regulating the expression of Rictor and TOR

Fig. 4.2 (continued) High hierarchy genes are identified by their node border color: green for high-hubs, red for VIPs, and blue for hubs. Abundantly expressed miRNAs are depicted as vee nodes. Gray lines indicate gene–gene links, whereas miRNA-gene validated interactions are indicated by blue lines. The vees filled with red or green colors indicate, respectively, hyper- or hypo-expressed miRNAs. Gene communities in both network diagrams are distinguished by different node colors. In CGCS, the communities are identified by different colors and the edge width and intensity are proportional to the connection weight of edges linking distinct communities. In the networks, the node size is proportional to the number of gene–gene links. In CGCS diagrams, the node size is proportional to the number of nodes/genes in each community. In the MM-DE network, the communities harboring high hierarchy genes are identified by the following colors: A, blue; B, orange; D, red; F, brown; G, pink; and I, olive green. In the MF-DE network the communities and their respective colors are: A, blue; B, orange; C, green; D, red; and E, purple

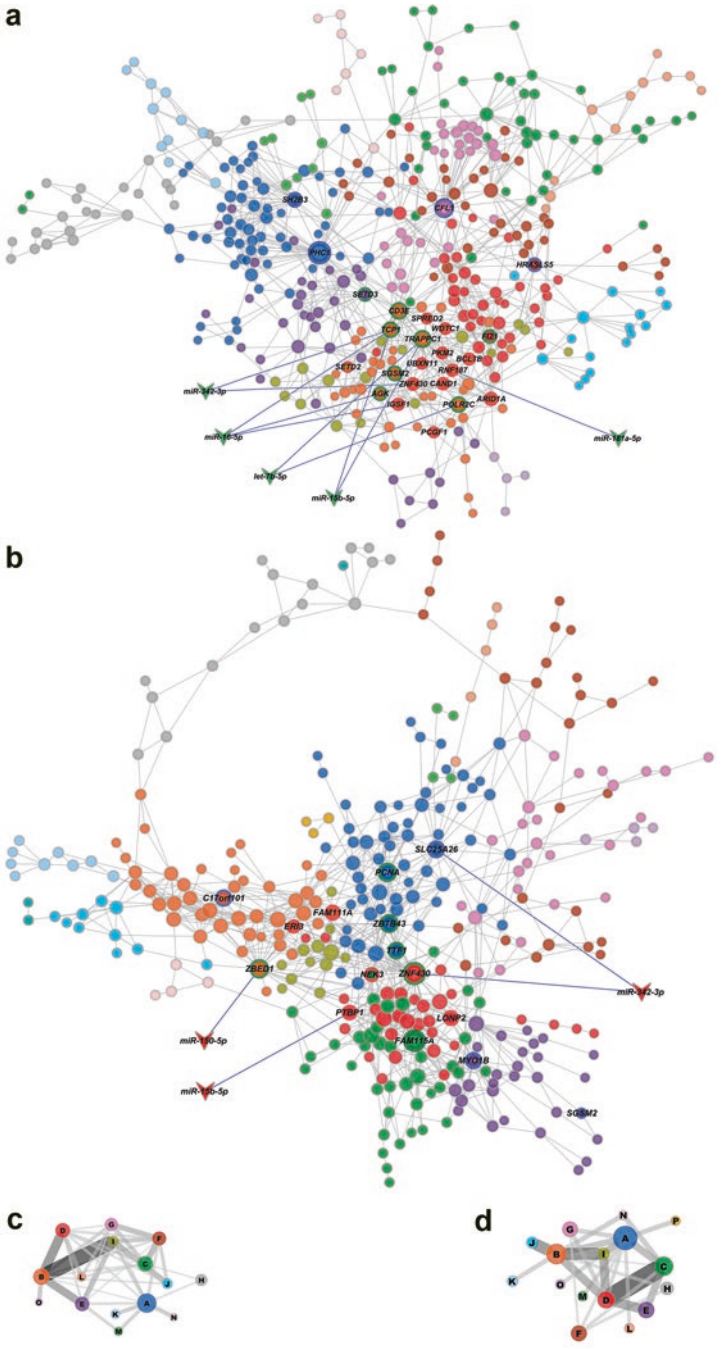

Fig. 4.2 DE networks with their respective gene communities (modules), miRNA–target interactions, and coarse-grained community structure (CGCS) diagrams. Network topology and community structure for minipuberty DE networks (**a** for MM and **b** for MF), and CGCSs for minipuberty DE networks (**c** for MM and **d** for MF) considering 15 and 16 communities per network, respectively.

(Singh et al. 2015). MiR-150 controls the Notch pathway and influences T-cell development and physiology (Ghisi et al. 2011). MiR-181 enhances cell proliferation in medullary thymic epithelial cells via regulating TGF-β signaling (Guo et al. 2016) and is involved in the positive and negative selection of T-cells (Fu et al. 2014). MiR-342-3p is a well-known regulator of the NF-κB pathway (Zhao and Zhang 2015), whose activation was shown to be necessary for the thymic expression of Aire in mice (Zhu et al. 2006; Haljasorg et al. 2015).

In the MM-DE network (Fig. 4.2a) community B harbors most of the HH genes (17 out of 24) and all the interactions between HH genes and abundantly expressed miRNAs. Moreover, all the HH genes in community B are VIPs (11 genes) or high-hubs (six genes), which means that these genes play relevant roles regarding the network functioning and robustness (van Dam et al. 2018). Indeed, VIPs connect different gene communities (Bando et al. 2013) and high-hubs are essential for the maintenance of network robustness (Azevedo and Moreira-Filho 2015). Network biology studies have shown that GCNs can be effectively used to associate highly connected genes (i.e., GCN hubs) with biological functions/processes in cells and tissues (Zhu et al. 2007; Gaiteri et al. 2014). Targeted hub attacks in protein–protein and gene–gene networks have been used to disclose relevant functional genes in health and disease (Gaiteri et al. 2014; Azevedo and Moreira-Filho 2015; Farooqui et al. 2018). Therefore, GCN hubs are relevant for both network topology and cell functioning.

Noteworthy, miRNA-target interactions involved only VIPs and high-hubs in MM-DE network. One of these high-hubs, TCP1, which codes for a molecular chaperone required for the transition of double negative to double positive T cells in the thymus (Cao et al. 2008), has interactions with three abundantly expressed miRNAs, all exerting known regulatory roles in the immune system. Functionally, most of the HH genes in MM-DE network are related to DNA and chromatin binding, DNA repair, histone modification, and ubiquitination (Moreira-Filho et al. 2018). CGCS analysis shows clearly that community B holds the highest connection weights, thus evidencing its importance for network functioning (Fig. 4.2c).

In the MF-DE network (Fig. 4.2b), the HH genes are quite evenly distributed among five gene communities: A (three high-hubs and one hub), B (two VIPs, one high-hub, and one hub), C (one high-hub and one VIP), D (two VIPs and one high-hub), and E (two hubs). Abundantly expressed miRNAs were found to interact with two high-hubs, one VIP, and one hub. The genes involved in these interactions were related to DNA binding (two genes), alternative mRNA splicing (one gene), and transmembrane (mitochondrial) transporter activity (one gene). The most represented molecular functions and biological processes among HH genes in MF-DE network are related to DNA binding, control of gene expression, and DNA repair and replication. CGCS analysis shows that the five gene communities harboring HH genes are also the ones presenting the highest connection weights (Fig. 4.2d).

GCN analyses (Fig. 4.2a, b) clearly show that abundantly expressed miRNAs interact almost exclusively with high-hubs and VIPs, i.e., with genes that are essential for network robustness (high-hubs) and for connecting gene communities

(VIPs). Altogether, these results indicate that testosterone and estradiol surges in minipuberty are related to significant changes in HH genes in MM and MF networks, respectively, and that these changes are under tight control by abundantly expressed miRNAs interacting with high-hubs and VIPs. In fact, relevant thymic functions, such as the induction of regulatory T cells, are regulated by abundantly expressed miRNAs (Singh et al. 2015). Noteworthy, all miRNAs interacting with HH genes in both networks play important roles in the regulation of immune processes, and particularly in the thymic environment.

4.4.4 AIRE *Interactors*

Since *AIRE* expression assessment by microarray analysis, RT-qPCR, and immunohistochemistry revealed no significant differences between male and female groups (Moreira-Filho et al. 2018), we constructed four other GCNs for investigating *AIRE* interactors' gene–gene expression relationships for minipuberty (MM and MF) and non-puberty groups (NM and NF). These *AIRE* interactors networks included *AIRE* and other 34 genes (34 genes in the minipuberty group and 33 genes in the non-puberty group), which code for proteins that are associated, directly or indirectly, with AIRE and exert impact on its functions (Abramson et al. 2010; Abramson and Goldfarb 2016). *AIRE* interactors were classified according to their molecular function and represented by different node colors in the networks (Fig. 4.3). Gene–gene expression relationships of *AIRE* interactors presenting a Pearson's correlation coefficient value ≥ 0.70 at least in one group across minipuberty and non-puberty samples were termed high interactors. We found 14 high-interactors distributed among minipuberty and non-puberty groups and, noteworthy, distinctive profiles of AIRE interactors' gene–gene relationships for each group (Fig. 4.3). The MM group encompassed more high interactors (seven) than the other three groups. These data suggest that sex hormones and genomic background exert their influence on AIRE interactors' gene–gene expression relationship during and after minipuberty.

Interestingly, neither the sex steroids surge during minipuberty nor the XY or XX background were found to promote any significant gender-related histomorphometric changes – average cortical thickness; average diameter of the medullary region; total area of the lobule; area of the medullary region; and medullary area/lobule area (%) – in neonatal and infant thymus (Moreira-Filho et al. 2018), thus corroborating previous data (Steinmann et al. 1985; Steinmann 1986).

In conclusion, these results suggest that genomic mechanisms and postnatal hormonal influences probably act synergistically in shaping thymic sexual dimorphism along the first 6 months of life, but this process does not involve changes in *AIRE* expression, although may involve differences – perhaps long-lasting differences – in the interactions of AIRE with its partners.

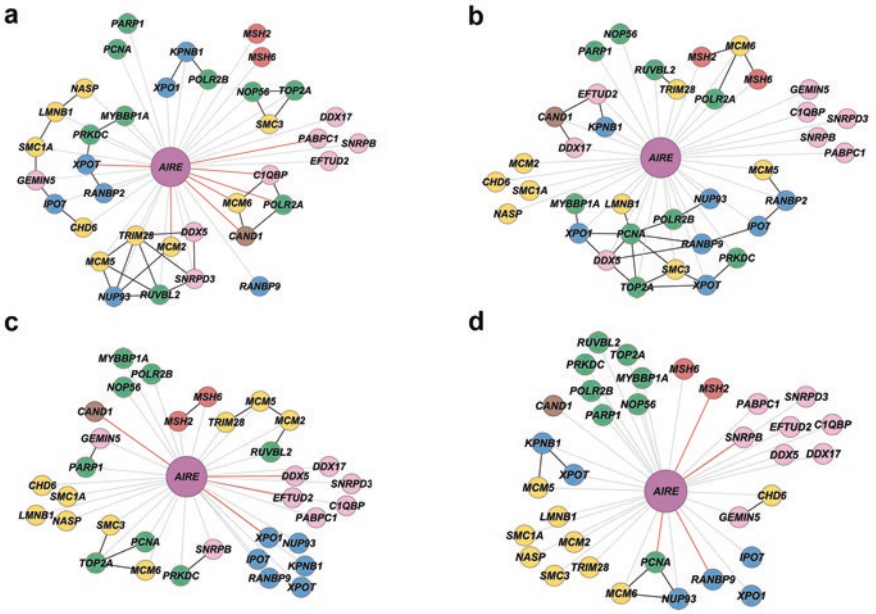

Fig. 4.3 AIRE interactors' gene–gene expression relationships. Gene–gene expression relationship networks for MM (**a**), MF (**b**), NM (**c**), and NF (**d**) groups. Nodes are colored according to their molecular function (GO): green for transcription, yellow for chromatin binding/structure, blue for nuclear transport, brown for ubiquitination, pink for pre-mRNA processing, red for DNA repair, and purple for *AIRE*. *AIRE*–gene expression correlation values <|0.70| are depicted with gray links; *AIRE*–gene expression correlation values ≥|0.70| are depicted with red links; gene–gene expression correlation values ≥|0.90| are depicted with black links

4.5 Age-Related Transcriptional Modules and TF–miRNA–mRNA Interactions in Neonatal and Infant Human Thymus

The human thymus grows only during the first year of life and its steady involution begins thereafter (Steinmann 1986; Thapa and Farber 2019). Moreover, the human thymus presents some functional peculiarities in the neonatal period and along the first 6 months of age, i.e., during minipuberty, when a transient surge in gonadal hormones takes place (Kuiri-Hänninen et al. 2014; Becker and Hesse 2020). In the neonatal thymus occurs a transient involution marked by severe depletion of double-positive (DP) thymocytes, which is later compensated by increased levels of primitive T-cell precursors. Concomitantly, there is a reinforcement of the subcapsular epithelial cell layer and an increase of the intralobular extracellular matrix network, leading to augmented thymic permeability and to the recirculation of primitive precursors and mature T-cells in the neonatal thymus (Varas et al. 2000). After the first year of life the total amount of lymphatic thymic tissue declines 5% per year until

the 10th year and at progressively slower rates afterwards (Steinmann 1986). The histomorphological features of thymic postnatal growth and of infant and adult thymic aging (lymphatic tissue decline, lipomatous atrophy) were quite well studied (Steinmann et al. 1985; Steinmann 1986; Chinn et al. 2012; Gui et al. 2012; Cowan et al. 2020), but the genomic mechanisms underlying this process remain largely unknown.

4.5.1 Transcriptome Analysis

In order to further investigate the genomic mechanisms involved in thymic growth and early involution, we performed a comparative transcriptome analysis of whole thymic tissue from human neonates and from infants grouped according to sequential age intervals (6 months) up to the first 2½ years of life. Thymic tissue samples were obtained from 57 karyotypically normal patients who underwent cardiac surgery. For genomic analyses samples were classified according to patients' age in five sequential age groups: A, neonates up to 30 days; B, infants aged 31 days to 6 months; C, infants aged 7–12 months; D, infants aged 13–18 months; and E, patients aged 19–31 months. DNA microarray technology was employed to obtain mRNA and miRNA expression profiles.

Whole thymic tissue transcriptome datasets were interpreted through modular repertoire identification. Here we employed a weighted gene co-expression network analysis (WGCNA) (Langfelder and Horvath 2008) for describing correlation patterns among genes across microarray datasets that allows: (i) the identification of transcriptional modules (Chaussabel and Baldwin 2014) and their association with particular age groups; (ii) the identification of highly connected genes (hubs) and of significant genes (HGS genes) for the trait of interest (age). This analysis was complemented by an integrative mRNA–miRNA–transcription factor (TF) co-expression analysis encompassing mRNAs from hubs and HGS genes, the abundantly expressed miRNAs, and the TFs that covaried with hubs and/or HGS genes.

The mRNA expression data matrix was used for GCN construction employing the WGCNA package (Langfelder and Horvath 2008). After dynamic tree cut, the hierarchical clustering dendrogram identified 15 distinct gene modules (Fig. 4.4a, b) containing from 85 (midnight blue module) to 403 genes (turquoise module). Genes not clustered in any module were grouped in the grey module. Subsequently, each age group was correlated with all the co-expression modules. This module-trait correlation analysis revealed three modules – tan, green yellow, and brown – that were significantly ($p < 0.05$) associated with at least one age group (Fig. 4.4c). The green yellow module was positively correlated with group E (MS = 0.41, $p = 0.003$); the brown module was negatively correlated with group E (MS = −0.34, $p = 0.02$); while the tan module was negatively correlated with group A (MS = −0.31, $p = 0.03$), and it was positively and significantly correlated with group E (MS = 0.30, $p = 0.03$). None of the modules were significantly correlated with gender or with the age groups ranging from 31 days to 18 months.

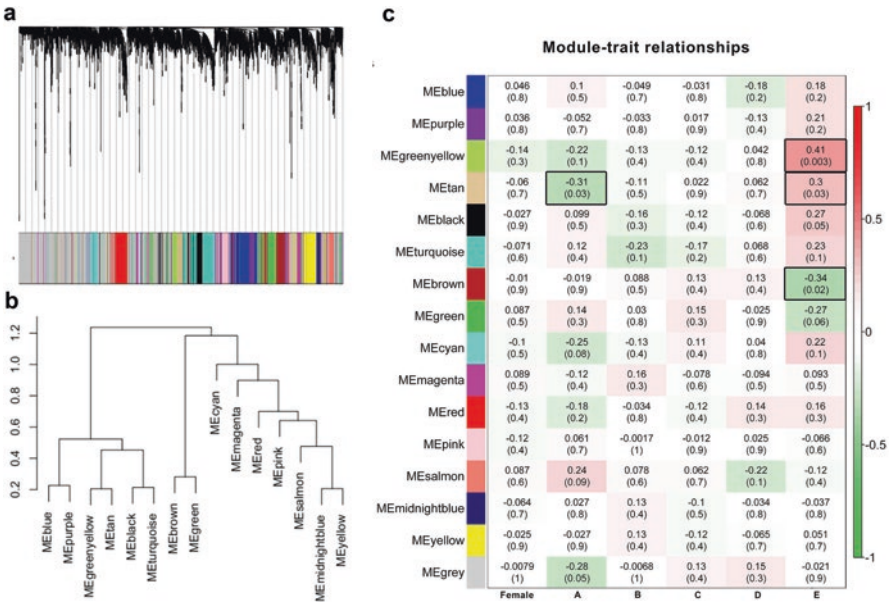

Fig. 4.4 WGCNA analysis. Gene dendrogram and gene clustering analysis for module identification (**a**). Hierarchical clustering dendrogram of the module-eigengenes (**b**). Module-trait relationships (**c**). The modules' names correspond to their colors (rows). Each column indicates a specific trait. The numbers inside each colored box are the module significance (MS) correlation values for gender and age groups, with p-value between parentheses. The more intense the box color, the more negatively (green) or positively (red) correlated is the module with the trait (MS values are indicated at the right color bar). Black-border boxes highlight the significant module-trait relationships

Hub genes identification in each significant module was accomplished through intramodular connectivity measures, i.e., the network nodes presenting high *k*Within values. A total of 34 hubs were found and assessed by an enrichment analysis. The Enrichr online web-based tool (Chen et al. 2013; Kuleshov et al. 2016) was used to identify significantly over-represented terms on GO Biological Process, Transcription Factor–PPIs Database, and miRTarBase. The TF–miRNA–mRNA regulatory network was then visualized by using the Cytoscape software, version 3.8.2 (Shannon et al. 2003). The hubs were mostly related to cellular/metabolic processes or related to T-cell development. The tan module (negatively associated with group A and positively associated with group E) encompasses a total of nine hubs. Two of them – *CAND1* and *ZNF675* – are related to medullary thymic epithelial cells (mTECs). The green yellow module (positively associated with group E) has a total of seven hubs. Three of them – *SNX17*, *MTMR4*, and *NKIRAS2* – are related to T-cell receptor (TCR) and thymic stromal functions. The brown module (negatively associated with group E) harbors a total of 18 hubs. Six of them – *CHMP5*, *PIK3CA*, *ARL8B*, *RNF138*, *NMI*, and *NRIP1* – are involved in T-cell development and antigen presentation-related functions.

The three age-related transcriptional modules here identified are correlated with two distinct and characteristic time intervals in human thymic evolution during the first 2½ years of life: the neonatal period (age group A) and the fourth/fifth semester period (age group E). In the neonatal period a transient thymic involution takes place, marked by severe cortical DP thymocyte depletion, and pronounced changes in the extracellular matrix network (Varas et al. 2000), whereas in the fourth semester of life the thymic involution becomes histomorphologically patent through the initial decline of the total amount of lymphatic tissue (Steinmann 1986). Interestingly, no transcriptional module was correlated with any interval in the 31 days to 18 months period, along which the thymus reaches its maximal growth (Steinmann 1986). Our TMA-IHC data on thymic cell subpopulations reflects this scenario, showing a continuous and moderate increase of thymocyte and B-cell numbers. These findings indicate that a genomic mechanism may act, synergistically with physiological and environmental stimuli (Moreira-Filho et al. 2018; Gui et al. 2012), on early thymic evolution/involution programming.

Additionally, we were able to identify the high gene significance (HGS) genes of the three modules significantly associated with age groups A (tan) and/or E (tan, green yellow, and brown). This categorization was accomplished according to module membership (MM) and gene significance (GS) values for groups A or E. Among the 50 HGS genes, a set of 37 were found to be DE: 19 genes were hyper-expressed in group A and 18 in group E. Moreover, 34 of these DE genes significantly varied their expression across all age groups. Among the hyper-expressed genes in group A, six are related to T-cell development and 13 to other cellular and metabolic processes. The hypo-expressed genes encompassed one gene related to T-cell development, six related to signaling pathways, and 11 related to other cellular and metabolic processes. It is interesting to mention that three of the hyper-expressed genes in the group A presented high fold-change values (>2.0): *CD5* and *CAND1*, in the tan module, and *SCML4*, in the green yellow module.

The integrative analysis of hubs and HGS genes, abundantly expressed miRNAs, and transcription factors (TFs) was accomplished by Pearson's correlation. For network construction we first obtained a gene expression data matrix for the above-mentioned genes, miRNA, and TFs. Subsequently, (i) for miRNA expression data we identified those abundantly expressed in at least one age group, and (ii) for TF expression data we used Enrichr online web-based tool (Chen et al. 2013; Kuleshov et al. 2016) to search TFs that have protein–protein interactions with hubs and HGS genes.

An integrative TF–miRNA–mRNA co-expression network of the abundantly expressed miRNAs covarying with hubs, HGS genes, and TFs was subsequently constructed (Fig. 4.5). It shows that most of the hub–hub or HGS–HGS gene links have positive correlations, while many hub–HGS gene links present negative correlations. Moreover, there are more hub–hub links than hub–HGS genes or HGS–HGS links. This result indicates that hubs are related to network module robustness and the HGS genes – which are differentially expressed genes – are either bridges between modules or border genes.

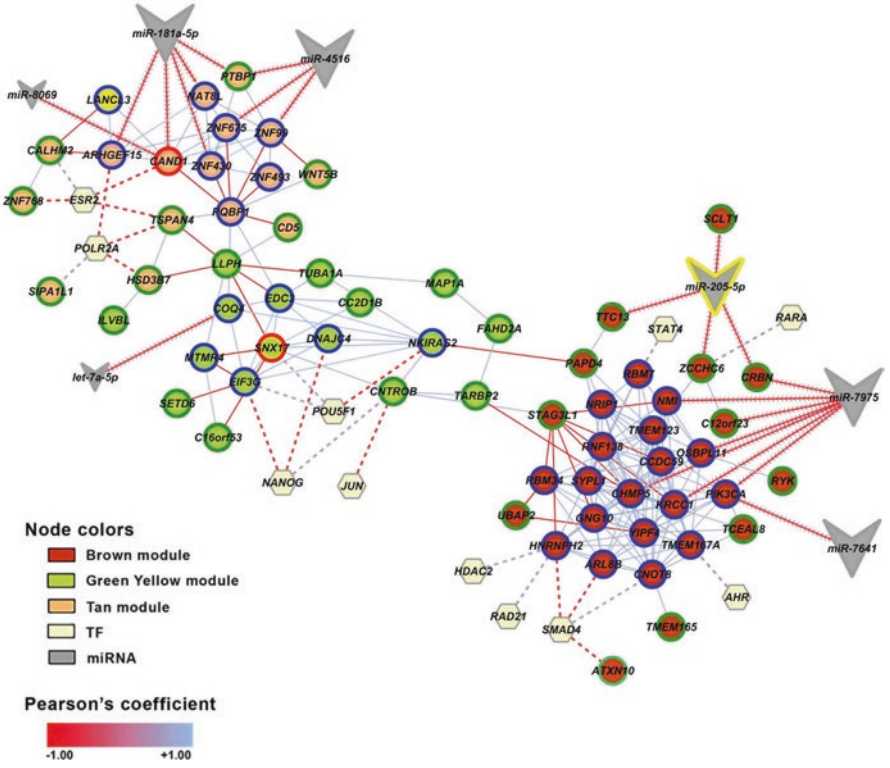

Fig. 4.5 Integrative TF–miRNA–mRNA co-expression subnetwork of the abundantly expressed miRNAs covarying with hubs, HGS genes, and transcription factors (TFs). Only co-expression covariance values of ≥|0.70| between gene–gene (solid lines), ≤ −0.50 between gene-miRNA (arrowed lines), and ≥|0.50| gene-TFs (dashed lines) were considered. Abundantly expressed miRNAs are depicted by gray vees; abundantly expressed and DE miRNAs are highlighted with a yellow border; HGS genes are depicted by green border nodes; hubs are depicted by blue border nodes; the two HGS genes that are also a hub gene are depicted by red border nodes; TFs are depicted by light yellow hexagons; positive and negative co-expression interactions are depicted by blue and red links, respectively

This integrative analysis clearly showed that the three age-related modules and their respective hubs are regulated by different and quite specific sets of abundantly expressed miRNAs and TF–hub interactions. The same situation prevails for the HGS genes, though it should be noted that just three TFs and several abundant miRNAs interact with the hyper-expressed genes in the age group A, whereas six different TFs but no abundantly expressed miRNA interact with the hypo-expressed genes in this group. The validated TF–miRNA interactions occurred more frequently with miR-150-5p, miR-181a-5p, and miR-205-5p.

Altogether, our results (Bertonha et al. 2020) show a genomic mechanism differentially governing the cellular and molecular processes involved in the functioning of the neonate thymus and, later, in the beginning of thymic decline. Along the first 2 years of age, this mechanism is tightly regulated by the differential expression of HGS genes and by TF-miRNA-hub/HGS interactions.

Acknowledgments This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) research grants 2015/22308-2 (CAM-F) and 2014/50489-9 (MC-S); and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) grant 3068936/2018-5 (CAM-F).

References

Abramson J, Goldfarb Y (2016) AIRE: From promiscuous molecular partnerships to promiscuous gene expression. Eur J Immunol 46:22–33

Abramson J, Giraud M, Benoist C et al (2010) AIRE's partners in the molecular control of immunological tolerance. Cell 140:123–135

Azevedo H, Moreira-Filho CA (2015) Topological robustness analysis of protein interaction networks reveals key targets for overcoming chemotherapy resistance in glioma. Sci Rep 5:16830

Bando SY, Silva FN, Costa L et al (2013) Complex network analysis of CA3 transcriptome reveals pathogenic and compensatory pathways in refractory temporal lobe epilepsy. PLoS One 8:e79913

Bando SY, Bertonha FB, Pimentel-Silva LR et al (2021) Hippocampal CA3 transcriptional modules associated with granule cell alterations and cognitive impairment in refractory mesial temporal lobe epilepsy patients. Sci Rep 11(1):10257

Barabási AL, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. Nat Rev Genet 5:101–113

Barabási AL, Gulbahce N, Loscalzo J (2011) Network medicine: a network-based approach to human disease. Nat Rev Genet 13:56–68

Becker M, Hesse V (2020) Minipuberty: why does it happen? Horm Res Paediatr 93:76–84

Berrih-Aknin S, Panse RL, Dragin N (2018) AIRE: a missing link to explain female susceptibility to autoimmune diseases. Ann N Y Acad Sci 1412:21–32

Bertonha FB, Bando SY, Ferreira LR et al (2020) Age-related transcriptional modules and TF-miRNA-mRNA interactions in neonatal and infant human thymus. PLoS One 15:e0227547

Billi AC, Kahlenberg JM, Gudjonsson JE (2019) Sex bias in autoimmunity. Curr Opin Rheumatol 31:53–61

Blondel VD, Guillaume JL, Lambiotte R et al (2008) Fast unfolding of communities in large networks. J Stat Mech P10008

Cao S, Carlesso G, Osipovich AB et al (2008) Subunit 1 of the prefoldin chaperone complex is required for lymphocyte development and function. J Immunol 181:476–484

Chaussabel D, Baldwin N (2014) Democratizing systems immunology with modular transcriptional repertoire analyses. Nat Rev Immunol 14:271–280

Chen EY, Tan CM, Kou Y et al (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinf 14:128

Cheng M, Anderson MS (2018) Thymic tolerance as a key brake on autoimmunity. Nat Immunol 19:659–664

Chinn IK, Blackburn CC, Manley NR et al (2012) Changes in primary lymphoid organs with aging. Semin Immunol 24:309–320

Clauset A, Newman MEJ, Moore C (2004) Finding community structure in very large networks. Phys Rev E70:066111

Cowan JE, Takahama Y, Bhandoola A et al (2020) Postnatal involution and counter-involution of the thymus. Front Immunol 11:897

Dragin N, Bismuth J, Cizeron-Clairac G et al (2016) Estrogen-mediated downregulation of AIRE influences sexual dimorphism in autoimmune diseases. J Clin Invest 126:1525–1537

Dumont-Lagacé M, St-Pierre C, Perreault C (2015) Sex hormones have pervasive effects on thymic epithelial cells. Sci Rep 5:12895

Farooqui A, Tazyeen S, Ahmed MM et al (2018) Assessment of the key regulatory genes and their Interologs for Turner Syndrome employing network approach. Sci Rep 8:10091
Fu G, Rybakin V, Brzostek J et al (2014) Fine-tuning T cell receptor signaling to control T cell development. Trends Immunol 35:311–318
Gaiteri C, Ding Y, French B et al (2014) Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. Genes Brain Behav 13:13–24
Geenen V (2021) The thymus and the science of self. Semin Immunopathol 43:5–14
Ghisi M, Corradin A, Basso K et al (2011) Modulation of microRNA expression in human T-cell development: targeting of NOTCH3 by miR-150. Blood 117:7053–7062
Gui J, Mustachio LM, Su DM et al (2012) Thymus size and age-related thymic involution: early programming, sexual dimorphism, progenitors and stroma. Aging Dis 3:280–290
Guo D, Ye Y, Qi J et al (2016) MicroRNA-181a-5p enhances cell proliferation in medullary thymic epithelial cells via regulating TGF-β signaling. Acta Biochim Biophys Sin Shanghai 48:840–849
Haljasorg U, Bichele R, Saare M et al (2015) A highly conserved NF-κB-responsive enhancer is critical for thymic expression of Aire in mice. Eur J Immunol 45:3246–3256
Hao Y, Hao S, Andersen-Nissen E et al (2021) Integrated analysis of multimodal single-cell data. Cell S0092-8674:00583–00583
Kernfeld EM, Genga RMJ, Neherin K et al (2018) A single-cell transcriptomic atlas of thymus organogenesis resolves cell types and developmental maturation. Immunity 48:1258–1270.e6
Klein SL, Flanagan KL (2016) Sex differences in immune responses. Nat Rev Immunol 16:626–638
Kondo K, Ohigashi I, Takahama Y (2019) Thymus machinery for T-cell selection. Int Immunol 31:119–125
Kuiri-Hänninen T, Sankilampi U, Dunkel L (2014) Activation of the hypothalamic-pituitary-gonadal axis in infancy: minipuberty. Horm Res Paediatr 82:73–80
Kuleshov MV, Jones MR, Rouillard AD et al (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 44(W1):W90–W97
Laios K (2018) The thymus gland in ancient Greek medicine. Hormones (Athens) 17:285–286
Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinf 9:559
Merrheim J, Villegas J, Van Wassenhove J et al (2020) Estrogen, estrogen-like molecules and autoimmune diseases. Autoimmun Rev 19:102468
Miller JFAP (2020) The function of the thymus and its impact on modern medicine. Science 369(6503):eaba2429
Moreira-Filho CA, Bando SY, Bertonha FB et al (2015) Community structure analysis of transcriptional networks reveals distinct molecular pathways for early- and late-onset temporal lobe epilepsy with childhood febrile seizures. PLoS One 10(5):e0128174
Moreira-Filho CA, Bando SY, Bertonha FB et al (2016) Modular transcriptional repertoire and MicroRNA target analyses characterize genomic dysregulation in the thymus of Down syndrome infants. Oncotarget 7:7497–74533
Moreira-Filho CA, Bando SY, Bertonha FB et al (2018) Minipuberty and sexual dimorphism in the infant human thymus. Sci Rep 8:13169
Nakaya HI, Wrammert J, Lee EK et al (2011) Systems biology of vaccination for seasonal influenza in humans. Nat Immunol 12:786–795
Narayanan T, Subramaniam S (2013) Community structure analysis of gene interaction networks in Duchenne muscular dystrophy. PLoS One 8:e67237
Newman MEJ (2010) Networks: an introduction. Oxford University Press, New York
Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69:026113
Obermoser G, Presnell S, Domico K et al (2013) Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines. Immunity 38:831–844

Ohigashi I, Tanaka Y, Kondo K et al (2019) Trans-omics impact of thymoproteasome in cortical thymic epithelial cells. Cell Rep 29:2901–2916.e6
Park JE, Botting RA, Domínguez Conde C et al (2020) A cell atlas of human thymic development defines T cell repertoire formation. Science 367(6480):eaay3224
Passos GA, Speck-Hernandez CA, Assis AF et al (2018) Update on Aire and thymic negative selection. Immunology 153:10–20
Perniola R (2018) Twenty years of AIRE. Front Immunol 9:98
Pobezinsky LA, Etzensperger R, Jeurling S et al (2015) Let-7 microRNAs target the lineage-specific transcription factor PLZF to regulate terminal NKT cell differentiation and effector function. Nat Immunol 16:517–524
Rezzani R, Nardo L, Favero G et al (2014) Thymus and aging: morphological, radiological, and functional overview. Age (Dordr) 36:313–351
Shannon P, Markiel A, Ozier O et al (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504
Singh Y, Garden OA, Lang F et al (2015) MicroRNA-15b/16 enhances the induction of regulatory T cells by regulating the expression of rictor and mTOR. J Immunol 195:5667–5677
Steinmann GG (1986) Changes in the human thymus during aging. Curr Top Pathol 75:43–88
Steinmann GG, Klaus B, Müller-Hermelink HK (1985) The involution of the ageing human thymic epithelium is independent of puberty. A morphometric study. Scand J Immunol 22:563–575
Thapa P, Farber DL (2019) The role of the thymus in the immune response. Thorac Surg Clin 29:123–131
Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98:5116–21
van Dam S, Võsa U, van der Graaf A et al (2018) Gene co-expression analysis for functional classification and gene-disease predictions. Brief Bioinform 19:575–592
Varas A, Jiménez E, Sacedón R et al (2000) Analysis of the human neonatal thymus: evidence for a transient thymic involution. J Immunol 164:6260–6267
Zhao L, Zhang Y (2015) miR-342-3p affects hepatocellular carcinoma cell proliferation via regulating NF-κB pathway. Biochem Biophys Res Commun 457:370–377
Zhu M, Chin RK, Christiansen PA et al (2006) NF-kappaB2 is required for the establishment of central tolerance through an Aire-dependent pathway. J Clin Invest 116:2964–2971
Zhu X, Gerstein M, Snyder M (2007) Getting connected: analysis and principles of biological networks. Genes Dev 21:1010–1024