Research Article

Portrait of Tissue-Specific Coexpression Networks of Noncoding RNAs (miRNA and IncRNA) and mRNAs in Normal Tissues

Claudia Cava, Gloria Bertoli, and Isabella Castiglioni

Institute of Molecular Bioimaging and Physiology, National Research Council (IBFM-CNR), Segrate (Mi), Italy

Correspondence should be addressed to Claudia Cava; claudia.cava@ibfm.cnr.it

Received 10 January 2019; Revised 1 August 2019; Accepted 10 August 2019; Published 3 September 2019

Academic Editor: Nadia A. Chuzhanova

Copyright © 2019 Claudia Cava et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Genes that encode proteins playing a role in more than one biological process are frequently dependent on their tissue context, and human diseases result from the altered interplay of tissue- and cell-specific processes. In this work, we performed a computational approach that identifies tissue-specific co-expression networks by integrating miRNAs, long-non-coding RNAs, and mRNAs in more than eight thousand human samples from thirty normal tissue types. Our analysis (1) shows that long-non-coding RNAs and miRNAs have a high specificity, (2) confirms several known tissue-specific RNAs, and (3) identifies new tissue-specific co-expressed RNAs that are currently still not described in the literature. Some of these RNAs interact with known tissue-specific RNAs or are crucial in key cancer functions, suggesting that they are implicated in tissue specification or cell differentiation.

1. Introduction

The portion of the genome encoding proteins constitutes only 1.5% of the human genome, while the transcribed RNAs not yielding coding proteins, named non-coding RNAs (ncRNAs) [1], are the major part.

ncRNAs are capable of controlling the expression of many genes, thus can simultaneously control multiple cellular functions regulating a variety of physiological and pathological processes [2, 3].

The family of ncRNAs includes small ncRNAs such as miRNAs, piRNAs, and siRNAs and long ncRNAs (lncRNAs). miRNAs, approximately 18–25 nucleotides in length, are highly conserved and involved in post-transcriptional gene silencing by binding complementary sequences in the 3′ untranslated region (3′ UTR) of messenger RNAs (mRNAs) [4]. Significant evidences in the last few years showed a fundamental role of miRNAs in determining cell fate, in pattern formation in embryonic development, and in controlling cell proliferation, cell differentiation, and cell death [4]. It has also been shown that aberrations in miRNA expression levels contribute to the onset and progression of many types of cancer [5–7].

Published studies have identified some tissue-specific (TS) or developmental-stage-specific miRNAs, suggesting their potential role in maintaining tissue identity and function [8–10]. The first comprehensive analyses of miRNAs across different human tissues have been reported in 2007, by Landgraf et al. [11] analyzing 340 distinct mature miRNAs in 26 human tissues and by Liang et al. [12] analyzing 345 distinct mature miRNAs in 40 human tissues. More recently, in 2016, Ludwig et al. profiled 1997 different mature miRNAs in 61 human tissues [13]. These studies showed that few miRNAs are expressed with a high degree of tissue specificity [11–13].

lncRNAs are transcribed RNA molecules bigger than 200 nucleotides in length, poorly conserved. They can regulate gene expression by different mechanisms that are not yet fully understood [3, 4]. Scientific evidences show they have roles in DNA replication, chromosomal dynamics, telomere biology, and subcellular structural organization [14–16].

Some works have found that lncRNAs tend to show even more TS expressions than mRNAs [14, 15]. For example, many lncRNAs are expressed only in limited developmental contexts or many of them are expressed during embryonic stem cell
Computational and Mathematical Methods in Medicine
differentiation and in the brain (in a mouse model) [9]. However, few TS IncRNAs have been well characterized to date.

Although all human tissues carry out common processes, each tissue is characterized by specific gene expression patterns that are essential for the conservation and survival of the tissue environment in physiological conditions. According to the Human Protein Atlas [17], approximately 40% of human mRNAs show a higher gene level expression in one single tissue. However, these studies did not focus on TS-miRNAs, TS-IncRNAs, and TS-mRNAs interactions.

The common and TS processes are ultimately controlled by gene regulatory networks that can integrate the complex interactions between mRNAs, miRNAs, and IncRNAs. Since different tissues have different regulatory networks, mapping regulatory networks among TS-RNAs could represent a significant step towards the knowledge, on molecular basis, of tissues and on how tissue specificity is determined.

The most widely used network model is based on gene co-expression network [18, 19]. Co-expression networks of biological elements (TS-mRNAs, TS-miRNAs, and TS-IncRNAs) are interesting since co-expressed elements are regulated by the same, functionally related transcriptional program or by members of the same pathway or protein complex. However, most of the few published co-expression studies between mRNAs, miRNAs, and IncRNAs [11, 15, 20] are focused on a single tissue or on a small number of different tissues.

Furthermore, recent studies showed that those ncRNAs, proposed as cancer biomarkers, were not tissue specific or were not expressed in the tissues where the disease is supposed to originate [21, 22]. Thus, it is unexpected to think about these ncRNAs as cancer biomarkers specific for a tissue. Indeed, a robust cancer biomarker of a specific tissue should show enrichment in that tissue in which the tumor develops. These findings highlight the need of developing new computational models to identify disease-related miRNAs/IncRNAs based on the tissue-specific co-expression networks. For example, it is not sufficient to predict the association of some miRNAs to a disease only based on the similarity expression, but it is also necessary to include the tissue specificity of non-coding RNAs. Adding this analysis into several existing tools [23, 24] could lead to a more comprehensive knowledge and an improved precision in the prediction of the association between non-coding RNA and disease.

In our study, focusing on the co-expression network between TS-mRNAs, TS-miRNAs, and TS-IncRNAs, we generated a large overview of TS-distribution of RNAs by using thirty different human normal tissues and more than eight thousand human samples.

Our portrait may help to understand how genes, whose functions depend on their tissue context, are expressed and regulated in a normal condition.

2. Materials and Methods

2.1. Data. GTEx data [25] (2016 release) provided the expression levels of RNAs in thirty different human normal tissues: adipose tissue, adrenal gland, bladder, blood, blood vessel, brain, breast, cervix uteri, colon, esophagus, fallopian tube, heart, kidney, liver, lung, muscle, nerve, ovary, pancreas, pituitary, prostate, salivary gland, skin, small intestine, spleen, stomach, testis, thyroid, uterus, and vagina, specifically 1368 miRNAs, 10167 IncRNAs (as reported from lncRNome database [26]), and 30441 mRNAs (as reported from HUGO database [27]). The subjects enrolled for the study range from 6 to 1259 for a total of 8553 samples based on the examined tissue (Table 1). For each tissue, and for each sample, we filtered mRNA profiles by removing pseudogenes, and we calculated median values of RNA expression levels. We thus created an $m \times N$ expression matrix, being $m$ the number of RNAs and $N$ the number of tissues (1–30). As RNA expression levels, we considered the median values over all samples for that tissue. We removed RNAs with median expression values $= 0$.

2.2. Gini Index. The Gini index (GI) [28] was used to identify TS-RNAs (miRNAs, IncRNAs, and mRNAs). Being $m$ the number of RNAs and $N$ the number of tissues (1–30), GI is defined as

$$GI = \frac{1}{N} \left( N + 1 - 2 \left( \sum_{j=1}^{N} \left( N + 1 - j \times x_i \right) \right) / \left( \sum_{j=1}^{N} x_i \right) \right),$$ (1)

where $x_i$ is the median expression of the RNA $i$ in tissue $j$ and $x_i$ is a vector in the n descending order ($i = 1, \ldots, m$; $j = 1, \ldots, N$).

GI measures the specificity of expression of an RNA with respect to different tissues and ranges from 0 to 1 indicating RNAs expressed in many tissues (housekeepers and constitutive) for scores of GI close to 0, and RNAs expressed in only one specific tissue for scores of GI close to 1 [28].

2.3. Gini Index Selection of TS RNAs. We considered a TS-RNA $i$ if $GI \geq 0.85$ and a nonspecific tissue RNA $i$ if $GI \leq 0.15$ [13, 28]. For RNA $i$ with $GI \geq 0.85$, we choose the TS where RNA expression profile has the highest value. Thus, an RNA can be TS only for one tissue but a tissue can have more TS RNAs.

2.4. Tissue-Specific miRNA-IncRNA-mRNA Regulatory Co-Expression Network. For each tissue, we generated a co-expression network: two RNAs are linked if their expression levels are correlated. We considered a co-expression interaction between two mRNAs if the Pearson correlation index was $>0.7$ or $<-0.7$. The correlation was calculated between expression levels of TS-miRNAs and TS-IncRNAs of the samples for each tissue. The same procedure was performed between TS-miRNAs and TS-mRNAs and between TS-miRNAs and TS-IncRNAs.

3. Results

3.1. Tissue-Specific RNAs. We present the results obtained from the Gini index used to measure tissue specificity.
Table 1: Number of samples for each normal tissue from GTEX.

<table>
<thead>
<tr>
<th>Tissue</th>
<th># of normal samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Adipose tissue</td>
<td>577</td>
</tr>
<tr>
<td>(2) Adrenal gland</td>
<td>145</td>
</tr>
<tr>
<td>(3) Bladder</td>
<td>11</td>
</tr>
<tr>
<td>(4) Blood</td>
<td>511</td>
</tr>
<tr>
<td>(5) Blood vessel</td>
<td>689</td>
</tr>
<tr>
<td>(6) Brain</td>
<td>1259</td>
</tr>
<tr>
<td>(7) Breast</td>
<td>214</td>
</tr>
<tr>
<td>(8) Cervix uteri</td>
<td>11</td>
</tr>
<tr>
<td>(9) Colon</td>
<td>345</td>
</tr>
<tr>
<td>(10) Esophagus</td>
<td>686</td>
</tr>
<tr>
<td>(11) Fallopian tube</td>
<td>6</td>
</tr>
<tr>
<td>(12) Heart</td>
<td>412</td>
</tr>
<tr>
<td>(13) Kidney</td>
<td>32</td>
</tr>
<tr>
<td>(14) Liver</td>
<td>119</td>
</tr>
<tr>
<td>(15) Lung</td>
<td>320</td>
</tr>
<tr>
<td>(16) Muscle</td>
<td>430</td>
</tr>
<tr>
<td>(17) Nerve</td>
<td>304</td>
</tr>
<tr>
<td>(18) Ovary</td>
<td>97</td>
</tr>
<tr>
<td>(19) Pancreas</td>
<td>171</td>
</tr>
<tr>
<td>(20) Pituitary</td>
<td>103</td>
</tr>
<tr>
<td>(21) Prostate</td>
<td>106</td>
</tr>
<tr>
<td>(22) Salivary gland</td>
<td>57</td>
</tr>
<tr>
<td>(23) Skin</td>
<td>890</td>
</tr>
<tr>
<td>(24) Small intestine</td>
<td>88</td>
</tr>
<tr>
<td>(25) Spleen</td>
<td>104</td>
</tr>
<tr>
<td>(26) Stomach</td>
<td>192</td>
</tr>
<tr>
<td>(27) Testis</td>
<td>172</td>
</tr>
<tr>
<td>(28) Thyroid</td>
<td>323</td>
</tr>
<tr>
<td>(29) Uterus</td>
<td>83</td>
</tr>
<tr>
<td>(30) Vagina</td>
<td>96</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>8553</strong></td>
</tr>
</tbody>
</table>

RNA selection led to 240 miRNAs, 9297 lncRNAs, and 19028 mRNAs. GI was calculated from these selected miRNAs, lncRNAs, and mRNAs (Figure 1). Only 4 miRNAs (miR-302b, miR-611, miR-1282, and miR-4721) (1.6%) were ubiquitously expressed with a GI < 0.15. Ninety-five miRNAs (39.7%) showed a TS expression with GI > 0.85 (Figure 1). The largest group of all miRNAs (58.7%) have 0.15 < GI < 0.85. Table 2 shows the partition of 95 miRNAs in the tissues.

RNA selection led to 240 miRNAs, 9297 lncRNAs, and 19028 mRNAs. GI was calculated from these selected miRNAs, lncRNAs, and mRNAs (Figure 1). Only 4 miRNAs (miR-302b, miR-611, miR-1282, and miR-4721) (1.6%) were ubiquitously expressed with a GI < 0.15. Ninety-five miRNAs (39.7%) showed a TS expression with GI > 0.85 (Figure 1). The largest group of all miRNAs (58.7%) have 0.15 < GI < 0.85. Table 2 shows the partition of 95 miRNAs in the tissues.

No TS-miRNA was found for the adipose tissue, blood vessel, breast, colon, esophagus, heart, pancreas, stomach, uterus, and vagina. Testis contains the highest number of TS-miRNAs.

Only ~1% lncRNAs were ubiquitously expressed with a GI < 0.15. The largest group (51%), including 4687 lncRNAs (Figure 1), showed a TS expression with GI > 0.85. However, 48.2% of all lncRNAs have 0.15 < GI < 0.85.

Table 3 shows the distribution of the 4687 lncRNAs for each tissue. TS-lncRNAs range from 1 in the esophagus to 3264 in the testis, representing the largest group. Supplementary Table 1 shows the list of TS-lncRNAs in detail.

About 9% mRNAs were ubiquitously expressed with GI < 0.15. A TS expression with GI > 0.85 was shown in 3668 mRNAs (Figure 1), and 72% of all mRNAs have 0.15 < GI < 0.85, representing the largest group.

Since the TS-mRNAs obtained from our analyses are only 19.3% of the total mRNAs, it appears that the regulation of the transcription by miRNAs and lncRNAs seems to be fundamental for the differentiation of each tissue.

Table 4 shows 3668 mRNAs for each tissue. TS-mRNAs range from 2 in the esophagus to 1563 in the testis, representing the largest group.

### 3.2. Tissue-Specific miRNA-LNC-mRNA Regulatory Networks

For each tissue type considered in this work, we analyzed the number of co-expression interactions between TS-miRNAs, TS-lncRNAs, and TS-mRNAs, and we present the results in Table 5.

The most numerous co-expression networks involve the testis, brain, small intestine, skin, and liver with 184041, 1512, 1197, 711, and 476 interactions, respectively. No TS-gene regulatory network was found in the blood vessel, esophagus, and uterus. In total, we found 189530 interactions and 3751 TS-RNAs (Figure 1).

The flow chart of the whole data selection and results is shown in Figure 1, for better synthesis of methodology.

### 4. Discussion

In the past decades, functional studies have focused on the analysis of protein-coding transcripts (mRNA expression) to characterize biological processes such as metabolism, cell differentiation, immune response, and tumorigenesis. The recent growth of next-generation sequencing technology has brought the discovery and characterization of a new class of non-coding RNA transcripts, such as miRNAs and lncRNAs. Evidences from high-throughput genomic platforms suggest the importance of ncRNAs in the regulation of multiple major biological processes such as cell development, differentiation, and metabolism [2].

However, despite successful studies highlighting important roles of miRNAs and lncRNAs in different tissues and diseases, no comprehensive portrait across various tissue types has been reported.

To assess the relation between mRNAs, miRNAs, and lncRNAs in several normal tissues, by using a computational approach, we investigated tissue specificity of miRNAs, lncRNAs, and mRNAs across 30 normal tissue types, obtaining 95 TS-miRNAs, 4687 TS-lncRNAs, and 3668 TS-mRNAs.

Our analysis on mRNAs suggests that they are not the determinants of tissue phenotype since only 19.3% of the total mRNAs is tissue specific. On the contrary, our study confirms the high specificity of lncRNAs in various tissues, consistently with previous studies [15, 29]. Indeed, we found that 51% of lncRNAs show high tissue specificity, highlighting the role of lncRNAs as regulators of gene expression in specific cellular phenotype. Thus, the identification and characterization of human lncRNAs with TS expression become crucial in order to determine their relevant functions.

Similarly to what is observed with lncRNAs, we found that 39.7% of miRNAs is tissue specific, also in this case proving the high tissue specificity of miRNAs and their...
critical role as regulators of mRNA expression, and thus, tissue phenotype.

We speculated that lncRNAs could (1) be controlled by miRNAs, reducing lncRNA stability and altering the cellular response in physiologic and pathologic processes; (2) act as endogenous sponges, harbouring similar microRNA target sequences and sequestering microRNAs; and (3) be precursors for the generation of miRNAs to silence target mRNAs [30].

For these reasons, the exploration of the regulatory relationships between miRNAs, lncRNAs, and mRNAs, specifically expressed in the same tissue, could offer useful information to shed light on how miRNAs, lncRNAs, and mRNAs play a role in tissue specification or cell differentiation [31].

However, the limit of our work is that we focused our analysis on a single dataset for each considered normal tissue. Future studies should validate our results on an independent dataset. Furthermore, we did not assess the influence on gene expression of some risk factors, such as age, smoking status, gender, and body mass index; thus, further studies on datasets including phenotype risk data could improve our methodology and results.

Despite these limitations and although several studies found that non-coding RNAs have important roles in diverse physiology and disease processes, we proposed a unique study that comprehensively compared the expression levels of tissue-specific non-coding RNAs with those of tissue-specific mRNAs.

Since a great number of TS-mRNAs are also key biomarkers in specific cancer diseases, the identification of TS co-expression network could be useful to deepen TS function of RNAs at molecular basis of cancer.

4.1. Adipose Tissue. TS-co-expression gene regulatory network in adipose tissue involves 5 mRNAs (ADIPOQ, CIDEA, CIDEC, PLIN1, and TUSC5) and 1 lncRNA (ADIPOQ-AS1), sharing 5 interactions. Although all 5 mRNAs in previous studies have showed a principal role in adipose tissue, no study has revealed an interaction between them.

Adiponectin, C1Q and collagen domain containing (ADIPOQ) is expressed exclusively in adipose tissue and is involved in the control of fat metabolism and insulin sensitivity [32].

Cell death-inducing DFFA-like effector A (CIDEA), along with CIDEB and CIDEC, constitutes the CIDE family of proteins. They are important modulators of several lipid metabolic pathways, including lipolysis, fatty acid oxidation, VLDL lipolization, and lipid droplet growth in adipocytes and hepatocytes [33]. Perilipin 1 (PLIN1) is a modulator of adipocyte lipid metabolism, lipolysis, and triglyceride levels. It plays a role in lipid formation by regulating CIDEC. The CIDEC-PLIN1 interaction promotes lipid droplet proteins acting on fat metabolism in adipocytes [34].

Tumor suppressor candidate 5 (TUSC5) was found as an adipose TS protein that is abundantly expressed in white [35] and brown adipose tissue [36], and TUSC5 mRNA expression increases during adipogenesis [36].

The interactions between them involve a lncRNA, ADIPOQ-AS1, that shows a positive correlation with other mRNAs. ADIPOQ-AS1 is located within the genomic location of ADIPOQ.

4.2. Adrenal Gland. TS-co-expression gene regulatory network in the adrenal gland involves 6 mRNAs (CYP11A1, CYP11B1, FAM166B, MRAP, NR5A1, and VWA5B2) 2 lncRNAs (AP000266.7 and RP11-320G24.1) and 1 miRNA (miR-1224), sharing 6 interactions.

Cytochrome P450 family 11 subfamily A member 1 (CYP11A1), subfamily B member 1 (CYP11B1), and family with sequence similarity 166 member B (FAM166B) showed a high expression in the adrenal gland but no study demonstrated a particular role in the adrenal gland to date [17, 37].
Table 2: Number of tissue-specific (TS) miRNAs for each tissue as obtained from the Gini index.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>TS-miRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Adipose tissue</td>
<td>—</td>
</tr>
<tr>
<td>(2) Adrenal gland</td>
<td>miR-4263, miR-1224, miR-4709, miR-33b</td>
</tr>
<tr>
<td>(3) Bladder</td>
<td>miR-3972, miR-5091, miR-93, miR-4469</td>
</tr>
<tr>
<td>(4) Blood vessel</td>
<td>miR-144, miR-3909, miR-223</td>
</tr>
<tr>
<td>(5) Brain</td>
<td>—</td>
</tr>
<tr>
<td>(6) Breast</td>
<td>miR-3682, miR-4477b, miR-3665, miR-628</td>
</tr>
<tr>
<td>(7) Cervix uteri</td>
<td>miR-320b2, miR-4269, miR-5481, miR-593</td>
</tr>
<tr>
<td>(8) Colon</td>
<td>—</td>
</tr>
<tr>
<td>(9) Esophagus</td>
<td>—</td>
</tr>
<tr>
<td>(10) Fallopian tube</td>
<td>miR-3671, miR-3125, miR-3622a, miR-23b, miR-3613, miR-369, miR-1225</td>
</tr>
<tr>
<td>(11) Heart</td>
<td>miR-1229a, miR-4783, miR-135a1, miR-1236, miR-4482-1, miR-4751</td>
</tr>
<tr>
<td>(12) Kidney</td>
<td>miR-4635, miR-4668</td>
</tr>
<tr>
<td>(13) Liver</td>
<td>miR-1295a, miR-4783, miR-135a1, miR-1236, miR-4482-1, miR-4751</td>
</tr>
<tr>
<td>(14) Lung</td>
<td>miR-133b</td>
</tr>
<tr>
<td>(15) Muscle</td>
<td>miR-133b, miR-133b, miR-133a2</td>
</tr>
<tr>
<td>(16) Nerve</td>
<td>miR-378b, miR-4500h, miR-5093, miR-657</td>
</tr>
<tr>
<td>(17) Ovary</td>
<td>miR-54812, miR-3960, miR-433, miR-1257</td>
</tr>
<tr>
<td>(18) Pancreas</td>
<td>miR-137hg, miR-339, miR-177, miR-32, miR-7-3hg, miR-7-3, miR-659</td>
</tr>
<tr>
<td>(19) Pituitary</td>
<td>miR-224</td>
</tr>
<tr>
<td>(20) Prostate</td>
<td>miR-642a</td>
</tr>
<tr>
<td>(21) Salivary gland</td>
<td>miR-205hg, miR-567, miR-936, miR-203</td>
</tr>
<tr>
<td>(22) Skin</td>
<td>miR-559, miR-3944, miR-192, miR-194-2, miR-147b</td>
</tr>
<tr>
<td>(23) Small intestine</td>
<td>miR-589, miR-378i, miR-1587</td>
</tr>
<tr>
<td>(24) Spleen</td>
<td>miR-1302-11, miR-4258, miR-4436a, miR-1302-3, miR-663b, miR-219-1, miR-3943, miR-5707, miR-3150b, miR-1302-3, miR-let7dg, miR-3689b, miR-202, miR-26a2, miR-5188, miR-4313, miR-3180-3, miR-632, miR-3615, miR-4526, miR-181c, miR-639, miR-519a2, miR-37b, miR-663a</td>
</tr>
<tr>
<td>(25) Stomach</td>
<td>miR-3907, miR-135a2</td>
</tr>
<tr>
<td>(26) Testis</td>
<td>—</td>
</tr>
<tr>
<td>(27) Thyroid</td>
<td>—</td>
</tr>
<tr>
<td>(28) Uterus</td>
<td>—</td>
</tr>
<tr>
<td>(29) Vagina</td>
<td>—</td>
</tr>
<tr>
<td>Total</td>
<td>95</td>
</tr>
</tbody>
</table>

Table 3: Number of tissue-specific (TS) IncRNAs for different tissues.

<table>
<thead>
<tr>
<th>Tissue</th>
<th># TS-IncRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Adipose tissue</td>
<td>16</td>
</tr>
<tr>
<td>(2) Adrenal gland</td>
<td>42</td>
</tr>
<tr>
<td>(3) Bladder</td>
<td>29</td>
</tr>
<tr>
<td>(4) Blood</td>
<td>31</td>
</tr>
<tr>
<td>(5) Blood vessel</td>
<td>23</td>
</tr>
<tr>
<td>(6) Brain</td>
<td>172</td>
</tr>
<tr>
<td>(7) Breast</td>
<td>8</td>
</tr>
<tr>
<td>(8) Cervix uteri</td>
<td>14</td>
</tr>
<tr>
<td>(9) Colon</td>
<td>10</td>
</tr>
<tr>
<td>(10) Esophagus</td>
<td>1</td>
</tr>
<tr>
<td>(11) Fallopian tube</td>
<td>39</td>
</tr>
<tr>
<td>(12) Heart</td>
<td>46</td>
</tr>
<tr>
<td>(13) Kidney</td>
<td>116</td>
</tr>
<tr>
<td>(14) Liver</td>
<td>108</td>
</tr>
<tr>
<td>(15) Lung</td>
<td>34</td>
</tr>
<tr>
<td>(16) Muscle</td>
<td>35</td>
</tr>
<tr>
<td>(17) Nerve</td>
<td>48</td>
</tr>
<tr>
<td>(18) Ovary</td>
<td>32</td>
</tr>
<tr>
<td>(19) Pancreas</td>
<td>52</td>
</tr>
<tr>
<td>(20) Pituitary</td>
<td>159</td>
</tr>
<tr>
<td>(21) Prostate</td>
<td>54</td>
</tr>
<tr>
<td>(22) Salivary gland</td>
<td>56</td>
</tr>
<tr>
<td>(23) Skin</td>
<td>44</td>
</tr>
<tr>
<td>(24) Small intestine</td>
<td>61</td>
</tr>
<tr>
<td>(25) Spleen</td>
<td>94</td>
</tr>
<tr>
<td>(26) Stomach</td>
<td>14</td>
</tr>
<tr>
<td>(27) Testis</td>
<td>326</td>
</tr>
<tr>
<td>(28) Thyroid</td>
<td>53</td>
</tr>
<tr>
<td>(29) Uterus</td>
<td>11</td>
</tr>
<tr>
<td>(30) Vagina</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>4687</td>
</tr>
</tbody>
</table>

Melanocortin 2 receptor accessory protein (MRAP) encodes a protein that controls protein trafficking and the functions of its receptor in the adrenal gland [38]. It also associated with adrenal disease and was detected in human adrenal and brain tissue [39].

Nuclear receptor subfamily 5 group A member 1 (NR5A1) is a nuclear receptor involved in gonadal and adrenal development. Mutations in this gene are a reason of XY sex reversal without a direct link with adrenal malfunctions [40].

von Willebrand factor A domain containing 5B2 (VWA5B2) was found expressed more than five fold in the human adrenal gland compared with other normal tissue types [41] but no study to date demonstrated a particular role in the adrenal gland. VWA5B2 showed a positive correlation with miR-1124.

AP000266.7 showed a positive correlation with MRAP and CYP11A1. AP000266.7 and MRAP are neighbor genes localized in chromosome 21.

RP11-320G24.1 showed a positive correlation with FAM166B, CYP11B1, and NR5A1. No relationship seems to be between them in literature.

C-type lectin domain family 3 member A (CLEC3A) was originally identified as a cartilage protein from shark and bovine [42], but recently it was found for the first time to be associated to the urinary bladder in [43] along with dehydrogenase/reductase 2 (DHRS2) and peptidyl arginine deiminase 3 (PADI3).

Several of identified genes have been associated with the progression of bladder cancer, such as S100 calcium-binding protein P (S100P), belonging to a family of calcium-binding proteins,[44], tachykinin receptor 3 (TACR3)[45], and UDP glucuronosyltransferase family 1 member A1 (UGT1A1)[46].

Transient receptor potential cation channel subfamily A member 1 (TRPA1) is expressed in the urinary bladder [47] as well as uroplakins (UPK1A, UPK1B, UPK2, and UPK3A) with a role in regulating the membrane permeability [48].

Among lncRNAs, we found POU6F2-AS2, which showed a positive correlation with PADI3, UPK1B, S100P, TACR3, TRPA1, OR5P2, and TMPRSS5. Although POU6F2-AS2 has not yet been reported in the bladder, previous studies showed its role in the initiation and progression of esophageal squamous cell carcinoma [49]. No information has been reported in the literature about other lncRNAs.


Apolipoprotein B MRNA editing enzyme catalytic subunit 3A (APOBEC3A) induces RNA editing in monocytes and macrophages [50]. Aquaporin 9 (AQP9) was found expressed in endothelial cells of neo-vessels [51] and along with S100 calcium-binding protein (S100A) they were identified as potential biomarkers of infective endocarditis.

For several found genes, a role has been reported in blood such as (1) colony-stimulating factor 3 receptor (CSF3R) that encodes a receptor protein for colony-stimulating factor 3, a cytokine that regulates the assembly, differentiation, and role of granulocytes [52], (2) cystatin F (CST7) known to be expressed by CD8+ T-cells [53], and (3) Fc fragment of IgA receptor (FCAR) that is a member of the immunoglobulin gene superfamily localized on the surface of, e.g., neutrophils, monocytes, macrophages, and eosinophils, with a role in immunologic responses to pathogens [54]. In particular,
using the Reactome pathways, 12 of 27 mRNAs are involved in neutrophil degranulation (i.e., matrix metalloproteinase 9 (MMP9), C-X-C motif chemokine receptor 1 (CXCR1), formyl peptide receptor 1 (FPR1), Selectin L (SELL), and leukocyte immunoglobulin-like receptor A5 (LILRA5)) [55]. Blood IncRNAs have not yet been reported in the literature.

miR-144, which correlates positively with Kruppel-like factor 1 (KLF1) and glycophrin A (MNS blood group) (GYPA), regulates hematopoiesis and vascular development by inhibiting expression of MEIS1 [56].

miR-223, which correlates positively with 16 mRNAs (APOBEC3A, AQP9, CSF3R, CST7, CXCR1, CXCR2, FFAR2, FPR1, FPR2, LILRA5, NFE2, S100A12, S100A9, SELL, TREML2, and VNN2) and 2 IncRNAs (AOAH-IT1 and AC008984.2), is a highly conserved miRNA [57] linked to granulopoiesis [58].

4.5. Brain. TS-co-expression gene regulatory network in the brain constitutes one of the biggest networks with 195 mRNAs, 102 IncRNAs, and 1 miRNA (miR-3665) sharing 1512 interactions.

Using the Reactome pathways, 38 of 195 mRNAs are involved in neuronal system (i.e., gamma-aminobutyric acid-type A receptor gamma 3), 29 of 195 mRNAs in transglutamate decarboxylase 2, protein kinase C gamma, and G protein subunit gamma 3), 21 of 195 mRNAs in neurotransmitter receptors and postsynaptic signal. miR-3665 is positively correlated with 3 mRNAs (contactin 2, myelin oligodendrocyte glycoprotein, and myelin-associated glycoprotein) and 1 IncRNA (RP11-116N8.1).

4.6. Breast. TS-co-expression gene regulatory network in the breast involves 1 mRNA (ANKRD30A) and 2 IncRNAs (RP11-20F24.2 and CTD-3032H12.1), sharing 2 interactions.

Ankyrin repeat domain 30A (ANKRD30A) encodes a transcription factor that is exclusively expressed in the breast and the testis. Deregulated expression levels have been correlated with breast cancer progression [59].

The expression of the locus containing the IncRNA RP11-20F24.2 has been already described in the breast, being altered in triple-negative breast cancer [60], while the IncRNA CTD-3032H12.1 has a role of a proliferation controller in in vitro experiment on breast cancer cell lines [61].

4.7. Cervix Uteri. TS-co-expression gene regulatory network in the cervix uteri involves 3 mRNAs (HOXA13, MAS1, and TG6M), 3 IncRNAs (RP11-246A10.1, XX-C2158C12.2 and RP11-893F2.15), and 1 miRNA (miR-4269), sharing 5 interactions.

Homeobox A13 (HOXA13) is highly expressed in the uterus, and its variation can alter development of the female reproductive tract [62].

For MAS1 proto-oncogene, G protein-coupled receptor (MAS1), transglutaminase 6 (TG6M), RP11-246A10.1, XX-C2158C12.2, RP11-893F2.15, and miR-4269 has not yet been reported a specific role in the cervix uteri.

4.8. Colon. TS-co-expression gene regulatory network in the colon involves 6 mRNAs (CHRM2, HAND1, MAB21L2, PHOX2A, PHOX2B, and PIR1) and 3 IncRNAs (CTB-158E9.1, RP11-227F19.1 and RP11-1336O20.2), sharing 6 interactions. Cholinergic receptor muscarinic 2 (CHRM2), a muscarinic cholinergic receptors belonging to a larger family of G protein-coupled receptors, and mab-21 like 2 (MAB21L2) have not yet reported in colon tissue.

Paired-like homeobox 2A and B (PHOX2A and PHOX2B) are involved in the intestinal neuronal differentiation playing a direct role in the pathogenesis of enteric disorders [63]. The heart and neural crest derivatives expressed 1 (HAND1), belonging to the basic helix-loop-helix family of transcription factors, is showed to be a tumor suppressor gene in colorectal cancer [64]. Phosphoinositide interacting regulator of transient receptor potential channels (PIRT) was already described in various regions of the gastrointestinal tract of the adult mouse [65]. Regarding the IncRNAs, RP11-1336O20.2 is the only IncRNA that has been described in colon cancer tissue [66].

4.9. Fallopian Tube. TS-co-expression gene regulatory network in the fallopian tube involves 25 mRNAs, 39 IncRNAs, and 7 miRNAs (miR-23b, miR-369, miR-1225, miR-3125, miR-3613, miR-3622a, and miR-3671), sharing 189 interactions.

According to the Reactome pathway, among 25 mRNAs, several are involved in GPCR downstream signaling such as bombesin receptor subtype 3 (BRS3), phosphoinositide-3-kinase regulatory subunit 2 (pIKR2), and olfactory receptor family 51 subfamily I member 2 (OR51I2), but there are not any mRNAs with a known role in the fallopian tube to date.

Previous studies have shown the role of miR-23b in cervical stem cells and ovarian cancer but no study has reported a role in the fallopian tube for miR-369, miR-1225, miR-3125, miR-3613, miR-3622a, and miR-3671, as well as for the 39 IncRNAs [67].


Creatine kinase, mitochondrial 2 (CKMT2) and dehydrogenase/reductase 7C (DHR57C) were found highly expressed in the heart [68, 69]. MYH6, MYBPC3, MYL3, MYL4, MYL7, MYOM2, and MYOZ2 encode proteins involved in cardiac function and they may be linked to modulation of cardiac contraction [70, 71].
Peroxidase-like (PXDNL) was found as a potential peroxidase homologue in the human heart [72]. SH3 domain-binding kinase family member 2 (SBK2) was found expressed in the atrium [37]. Some evidences have shown a role of small muscle protein, X-linked, SPHK1 interactor, AKAP domain containing (SPHKAP), T-box 5 (TBX5), and troponin I3, cardiac type (TNN13) in the heart [73–76].

Retinal degeneration 3-like (RD3L) and single-pass membrane protein with coiled-coil domains 1 (SMCO1) have not yet been reported to have a role in the heart. No publication is present on the role of the 11 lncRNAs in heart tissue.

4.11. Kidney. TS-co-expression gene regulatory network in the kidney involves 64 mRNAs, 68 lncRNAs, and 1 miRNA (mir-1227), sharing 385 interactions.

According to the Reactome pathway, 24 of 64 mRNAs are involved in transport of small molecules, 17/64 mRNAs in SLC-mediated transmembrane transport, and 9/64 mRNAs in transport of glucose, bile salts and organic acids.

Among mRNAs of the first group, there are solute carrier family, ATPase H + transporting (ATP6V0A4, ATP6V0D2, and ATP6V1G3), and FXYD domain-containing ion transport. No publication is present on the role of the 11 lncRNAs in heart tissue.

4.12. Liver. TS-co-expression gene regulatory network in the liver involves 133 mRNAs, 44 lncRNAs, and 1 miRNA (mir-135a1), sharing 476 interactions.

According to the Reactome pathway, 32 of 133 mRNAs are involved in biological oxidation, 77/133 mRNAs in metabolism, and 29/133 mRNAs in hemostasis.

mir-135a1 has an important role in proliferation and metastatic growth of hepatocellular cells and is associated with malignant thrombosis [77].

4.13. Lung. TS-co-expression gene regulatory network in the lung involves 11 mRNAs (ALPP, MS4A15, NAPSA, ROS1, SCGB1A1, SFTA2, SFTPA1, SFTPA2, SFTPB, SFTPC, and SFTPDP) and 7 lncRNAs (AC068134.8, CTD-2531D15.4, RP11-312J18.6, RP11-476D10.1, RP11-62I21.1, RP11-95I16.2, and U82670.4), sharing 13 interactions.

Napsin A aspartic peptidase (NAPSA) encodes a member of the peptidase A1 family of aspartic proteases, and it is involved in the processing of pulmonary surfactant protein B in the lung [78]. ROS proto-oncogene 1, receptor tyrosine kinase (ROS1), encodes an orphan receptor tyrosine kinase and is vulnerable to intrachromosomal or interchromosomal rearrangements in several tumor types including non-small-cell-lung cancer [79].

Secretoglobin family 1A member 1 (SCGB1A1) is involved in the regeneration of alveolar epithelia in response to severe pulmonary damage in mice [80]. Surfactant family genes (SFTA2, SFTPA1, SFTPA2, SFTPB, SFTPC, and SFTPDP) encode pulmonary surfactant-associated proteins [81]. Alkaline phosphatase, placental (ALPP) and membrane spanning 4-domains A15 (MS4A15) have not yet been reported with a specific role in the lung. None of the 7 lncRNAs has been described in lung tissue before.


Using the Reactome pathway, we found that 20/38 mRNAs (i.e., actinin alpha 2 (ACTN2), myosin heavy chain 13 (MYH13), troponin I1, slow skeletal type (tnn1), calcium voltage-gated channel subunit alpha1 S (CACNA1S), and calsequestrin 1 (CASQ1)) are involved in muscle contraction and 15/38 in striated muscle contraction (i.e., nebulin (NEB) and titin-cap (TCAP)).

mir-133a-2 along with mir-133a-1 is muscle-specific miRNAs that is modulated during muscle development by the serum response factor [82]. Although lncRNA RP11-612.4 has been recently associated to atherosclerosis [83], no publication is currently available for the identified lncRNAs in the muscle.

4.15. Nerve. TS-co-expression gene regulatory network in the nerves involves 4 mRNAs (HCN1, IL9, S100B, and SL25A48) and 3 lncRNAs (GS1-39E22.2, LINC00462, and PEX5L-AS1), sharing 5 interactions.

Hyperpolarization-activated cyclic nucleotide-gated potassium channel 1 (HCN1) has been found in the central and peripheral nervous system, where they are associated with synaptic integration, neuronal excitability, and formation of resting membrane potentials [84]. Interleukin 9 (IL9), S100 calcium-binding protein B (S100B), and solute carrier family 25 member 48 (SLC25A48) have not yet been reported to have a specific role in the nerves.

4.16. Ovary. TS-co-expression gene regulatory network in the ovary involves only 1 mRNA (WFIIKKN2) and 1 lncRNA (RP11-506D12.5), sharing only 1 interaction.

WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 2 (WFIIKKN2) has not yet been reported to have a specific role in the ovary.

RP11-506D12.5 along with other 12 lncRNAs was proposed as gene signature to identify breast cancer patients at high risk of tumor recurrence [85].

4.17. Pancreas. TS-co-expression gene regulatory network in the pancreas involves 3 mRNAs (IL22RA1, RNF186, and PRSS3) and 3 lncRNAs (RP11-133O22.6, RP11-91K11.2, and RP4-799P18.2), sharing 3 interactions. Interleukin 22 receptor subunit alpha 1 (IL22RA1) and protease serine 3 (PRSS3) were associated with chronic pancreatitis in mouse models [86] and with progression, metastasis, and prognosis of human pancreatic cancer [87].
Ring finger protein 186 (RNF186) has not yet been reported in the pancreas. No publication is found on the role of the 3 identified lncRNAs in the pancreas function.

4.18. Pituitary. TS-co-expression gene regulatory network in the pituitary involves 47 mRNAs, 34 lncRNAs, and 4 miRNAs (miR-7-3, miR-7-3hg, miR-137hg, and miR-770), sharing 137 interactions. According to the Reactome pathway, 4 of 47 mRNAs are involved in peptide hormone biosynthesis (i.e., glycoprotein hormones, alpha polypeptide (CGA), proopiomelanocortin (POMC), and luteinizing hormone beta polypeptide (LHB)), 3/47 glycoprotein hormones, and 4/47 in class A/1 (rhodopsin-like receptors) (i.e., arginine vasopressin receptor 1B (AVPR1B), and dopamine receptor D2 (Drd2)).

miR-7 was found to be enriched in the pituitary [88], miR-137hg was associated as schizophrenia-associated locus [89], and miR-770 have been linked to pituitary adenoma [90].


Acid phosphatase, prostate (ACPP) is regulated by androgen and is produced by the epithelial cells of the prostate gland [91]. Cholinergic receptors (CHRM1 and CHRNA2), homeobox B13 (HOXB13), matrix metalloproteinase 26 (MMP26), NK3 homeobox 1 (NKX3-1), prostate cancer susceptibility candidate (PRAC1 and PRAC2), SAM pointed domain containing ETS transcription factor (SPDEF), transglutaminase 4 (TGM4), and transient receptor potential cation channel subfamily M member 8 (TRPM8) are associated with prostate cancer [92–99].

Kallikrein related peptidase family (KLK2, KLK3, and KLK4) encode a serine proteases mainly expressed in prostatic tissue. They are highly expressed in prostate tumor cells and may be a prognostic maker for prostate cancer risk [100].

Lectin, mannos-binding 1 like (LMAN1L), and cell receptor gamma joining family (TRGJ1 and TRGJP) have not yet been associated with the prostate. No publication is currently available on the role of the 16 identified lncRNAs in prostate tissue.


Polymeric immunoglobulin receptor (PIGR), also known as the membrane secretary component, plays a central role in this network. It has a crucial function in intestinal immunity, but since the oral cavity is protected by the mucosal immune system, PIGR seems to have a role also in the salivary gland [101].

4.21. Skin. TS-co-expression gene regulatory network in the skin involves 121 mRNAs, 23 lncRNAs, and 1 miRNA (miR-205hg), sharing 711 interactions.

According to the Reactome pathway, 51 of 121 mRNAs are involved in keratinization (i.e., keratin family genes (KRT), desmoglein (DSG), and desmocollin (DSC)), and formation of the cornified envelope (i.e., late cornified envelope 6A and plakophilin).

miR-205hg was associated with cutaneous squamous cell carcinoma [102] and controls neonatal expansion of skin stem cells by regulating the PI(3)K pathway [103].

4.22. Small Intestine. TS-co-expression gene regulatory network in the small intestine involves 117 mRNAs, 30 lncRNAs, and 4 miRNAs (miR-147b, miR-192, miR-194-2, and miR-559), sharing 1197 interactions.

According to the Reactome pathway, 8/117 mRNAs are involved in antimicrobial peptides (i.e., intelectin 2 and regenerating family member 3 alpha and 4), 7/117 mRNAs in O-linked glycosylation of mucins (i.e., mucin 2, 3A, and 17; cell surface-associated and UDP-GlcNAc : betagal beta-1, 3-N-Acetylglucosaminyltransferase 6), 7/117 in MHC class II antigen presentation (i.e., tubulin alpha-like 3 and sucrase-isomaltase), and 5/117 in alpha-defensins (i.e., defensin alpha 5 and 6).

miR-147b was found as a biomarker for colorectal tumor, miR-192 seems to play a vital role in the differentiation and function of the intestinal epithelium, and miR-194 is regulated by HNF-1 alpha and seems to have important functions in intestinal epithelium maturation [104–106]. To date, miR-559 has not yet been stated to be associated with the intestine in the literature.

4.23. Spleen. TS-co-expression gene regulatory network in the spleen involves 70 mRNAs, 17 lncRNAs, and 1 miRNA (miR-1587), sharing 136 interactions.

According to the Reactome pathway, 42/70 mRNAs are involved in FCGR activation (i.e., immunoglobulin lambda constant family and Fc Receptor), 42/70 mRNAs in the role of phospholipids in phagocytosis, 42/70 in Fc gamma receptor (FCGR) dependent phagocytosis, and 47/70 in the innate immune system (i.e., integrin subunit alpha D).

miR-1587 has not yet been reported in the spleen. No publication is currently available on the role of the 17 identified lncRNAs in the spleen.

4.24. Stomach. TS-co-expression gene regulatory network in the stomach involves 22 mRNAs (ANXA10, ATP4A, ATP4B, BARX1, CA9, CCKBR, CHIA, CLDN18, CTSE, DPCR1, GIF, GKN2, KCNE2, LHX5, MUC5AC, NCOA7-AS1, PGA4, SLC26A9, SLC9A4, TFF1, TRIM50, and VSG1) and 10 lncRNAs (AP000320.6, CTD-2331C18.5, HCG21, HNF1A-AS1, RP11-231K24.2, RP11-567G11.1, RP11-708H21.4, RP11-
Annexin A10 (ANXA10) is a nuclear protein particularly expressed in the gastric mucosa, and it is a novel marker of gastric differentiation [107].

ATPase H+/K+ transporting subunit family (ATP4A and ATP4B) encodes a proton pump that catalyzes the hydrolysis of ATP. It is also responsible for gastric acid secretion [108].

BARX homeobox 1 (BARX1) regulates robust smooth muscle growth in the stomach by increasing proliferation of myogenic progenitors [109].

Carbonic anhydrase 9 (CA9) plays a role in the metabolism of carbon and has also been associated with hypoxia and proliferation in different cancers such as gastric cancer [110].

Cholecystokinin B receptor (CCKBR), chitinase acidic (CHA), claudin 18 (CLDN18), and cathepsin E (CTSE) are found mainly in the gastrointestinal tract with high levels in human stomach tissue [111–114].

Gastric intrinsic factor (GIF) and gastrokine 2 (GKN2) encode a glycoprotein secreted by the cells of the gastric mucosa [115, 116]. GIF is necessary for the correct absorption of vitamin B12 [115].

Potassium voltage-gated channel subfamily E regulatory subunit 2 (KCNE2), secreted mucin 5AC, oligomeric mucus/gel-forming (MUC5AC), pepsinogen 4, group 1 (Pepsinogen A), and trefoil factor 1 (TFF1) are essential for gastric acid secretion and for normal gastrointestinal function [117–120]. Some of them are the main components of the gastrointestinal mucose layer [117–120].

Tripartite motif containing 50 (TRIM50) checks vesicular trafficking for acid secretion in gastric cells [121]. V-set and immunoglobulin domain containing 1 (VSG1) is stomach-specific for the accurate differentiation of gastric epithelia [122].

Diffuse panbronchiolitis critical region 1 (DPCR1), LIM homeobox 5 (LHX5), and solute carrier family (SLC26A9 and SLC9A4) have not been reported with a specific role in the stomach. No publication is currently available on the role of the 10 identified lncRNAs in the stomach.

4.25. Testis. TS-co-expression gene regulatory network in the testis involves 849 mRNAs, 1318 lncRNAs, and 13 miRNAs (miR-let7dhg, miR-181c, miR-202, miR-371b, miR-519a2, miR-663a, miR-1302-11, miR-3150b, miR-3180-3, miR-3615, miR-3943, miR-4313, and miR-5188), sharing 18401 interactions.

According to the Reactome pathway, the major part of genes is involved in reproduction, fertilization, sperm motility, and meiosis.

miR-202 was found in Sertoli cells, and its expression varies between fertile and non-fertile men, suggesting it as a TS biomarker [123].

The other miRNAs have not yet been reported to have a comprehensive role in the testis.
project grant SysBioNet, Italian Roadmap Research Infrastructures 2012.

**Supplementary Materials**

Supplementary Table 1: the list of TS-lncRNAs. (Supplementary Materials)

**References**


[75] A. Khalil, R. Tanos, N. El-Hachem et al., “A HAND to TBX5 explains the link between thalidomide and cardiac diseases,” *Scientific Reports*, vol. 7, no. 1, p. 1416, 201


