RESEARCH

Microarray profile of B cells from Graves’ disease patients reveals biomarkers of proliferation

Xuechao Jiang²,³,* Yonghui Wang¹,** Xiaoying Li¹,** Leqi He⁴, Qian Yang¹, Wei Wang¹, Jun Liu¹ and Bingbing Zha¹

¹Department of Endocrinology, Fifth People’s Hospital of Shanghai Fudan University, Shanghai, China
²Scientific Research Center, Xinhua Hospital, Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China
³Department of Pediatric Cardiology, Xinhua Hospital, Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China
⁴Department of Clinical Laboratory Medicine, Fifth People’s Hospital of Shanghai Fudan University, Shanghai, China

Correspondence should be addressed to B Zha: bingbingzha@fudan.edu.cn
*(X Jiang, Y Wang and X Li contributed equally to this work)

Abstract

B lymphocytes are the source of autoantibodies against the thyroid-stimulating hormone receptor (TSHR) in Graves’ disease (GD). Characterization of autoimmune B-cell expression profiles might enable a better understanding of GD pathogenesis. To reveal this, the expression levels of long noncoding RNAs (lncRNAs) and mRNAs (genes) in purified B cells from patients with newly diagnosed GD and healthy individuals were compared using microarrays, which elucidated 604 differentially expressed lncRNAs (DE-lncRNAs) and 410 differentially expressed genes (DEGs). GO and pathway analyses revealed that the DEGs are mainly involved in immune response. A protein–protein interaction network presented experimentally validated interactions among the DEGs. Two independent algorithms were used to identify the DE-lncRNAs that regulate the DEGs. Functional annotation of the deregulated IncRNA–mRNA pairs identified 14 pairs with mRNAs involved in cell proliferation. The IncRNAs TCONS_00022357-XLOC_010919 and n335641 were predicted to regulate TCL1 family AKT coactivator A (TCL1A), and the IncRNA n337845 was predicted to regulate SH2 domain containing 1A (SH2D1A). TCL1A and SH2D1A are highly involved in B-cell proliferation. The differential expression of both genes was validated by qRT-PCR. In conclusion, IncRNA and mRNA expression profiles of B cells from patients with GD indicated that the IncRNA–mRNA pairs n335641–TCL1A, TCONS_00022357-XLOC_010919–TCL1A, and n337845–SH2D1A may participate in GD pathogenesis by modulating B-cell proliferation and survival. Therefore, the identified IncRNA and mRNA may represent novel biomarkers and therapeutic targets for GD.

Introduction

Graves’ disease (GD) is the most common cause of hyperthyroidism and diffuse goiter, affecting 1.0–1.5% of the population (1). GD is characterized by lymphocytic infiltration and the production of autoantibodies directed against components of the thyroid, including thyroglobulin (TG), thyroid peroxidase (TPO), and thyroid-stimulating hormone receptor (TSHR) (2, 3). TSHR autoantibodies mimic the function of TSH and lead to uncontrolled thyroid hormone production and hyperthyroidism (4, 5). B cells are the source of pathognomonic activating autoantibodies (TRAb) against TSHR in GD (6, 7). B cells also present antigens to T cells
Although B cells are essential for GD pathogenesis, little is known regarding the expression profiles of autoimmune B cells during GD development.

Long noncoding RNAs (lncRNAs), defined as non-protein-coding transcripts longer than 200 nucleotides, have been primarily studied for their effects on development, cancer, and cell differentiation (9, 10). There is emerging evidence that lncRNAs play crucial roles in immune regulation and autoimmunity. T lymphocytes, B lymphocytes, macrophages, dendritic cells, and NK cells all express lncRNAs, which have been shown to be involved in the differentiation and activation of immune cells (11). Accumulating evidence suggests that lncRNA dysregulation is involved in the pathogenesis of autoimmune diseases, such as systemic lupus erythematosus, rheumatoid arthritis, Crohn’s disease, multiple sclerosis, and autoimmune thyroid disease (AITD) (11, 12, 13, 14, 15). AITD, which includes GD and Hashimoto’s thyroiditis, is caused by an immune response to self-thyroid antigens (16). Single-nucleotide polymorphisms (SNPs) in the promoter of SAS-ZFAT, an antisense transcript of the ZFAT gene, were reported to determine susceptibility to AITD. SAS-ZFAT is exclusively expressed in CD19+ B cells in peripheral blood, indicating that SAS-ZFAT might control susceptibility to AITD by modulating B-cell function (12, 16). Recent studies have illustrated that lncRNAs can regulate the expression of protein-coding genes through cis-acting mechanisms on neighboring genes or through trans-acting mechanisms involving protein complexes such as transcription factors, splicing factors, or RNA-decay machinery (17, 18). In this study, we aimed to identify lncRNA–mRNA pairs that play a functional role in GD.

We found that patients with newly diagnosed GD had significantly higher numbers of B cells among their peripheral blood mononuclear cells (PBMCs) than healthy individuals. To reveal the mechanism of the increase in peripheral B cells, we studied the lncRNA and mRNA profiles of peripheral B cells from patients with GD by microarray analysis. We focused on lncRNA–mRNA pairs that were abnormally regulated in GD, especially those involving protein-coding genes that were related to cell proliferation. Among the identified genes, TCL1 family AKT coactivator A (TCL1A), and SH2 domain containing 1A (SH2D1A) have been reported to be associated with B-cell proliferation and survival. Using computational algorithms, we predicted that the lncRNAs TCONS_00022357-XLOC_010919–TCL1A, and n337845–SH2D1A might be useful biomarkers related to the proliferation of B lymphocytes. Furthermore, drugs targeting these genes might represent a novel therapy for GD.

Materials and methods

GD patients and healthy individuals

Peripheral blood was obtained from patients with newly diagnosed GD without thyroid associated ophthalmopathy (n = 34) and age-matched healthy individuals (n = 34). The patients with GD were clinically diagnosed with recent-onset GD in the Fifth People’s Hospital of Shanghai medical ward from 2014 to 2016 on the basis of the presence of overt hyperthyroidism, diffuse goiter or normal thyroid volume on ultrasonography, and a high rate thyroid iodine uptake. The characteristics of the patients are summarized in Table 1. Healthy controls normalized for sex, age, smoking status, thyroid function, and thyroid autoantibodies levels were recruited from the physical checkup center of the Fifth People’s Hospital of Shanghai, Shanghai, China. The controls were all in good health and had no family or personal history of thyroid disease. The exclusion criteria were a history of other autoimmune diseases, pregnancy during the study period, or a history of immunosuppressive medication. All procedures were in accordance with the ethical standards of the institutional and national research committees and the Declaration of Helsinki. The study protocols were approved by the Medical Ethics Committee of the Fifth People’s Hospital of Shanghai, Fudan University, Shanghai, China (NO. 002). All participants provided written informed consent.

Thyroid function and antithyroid antibodies detection

The levels of free triiodothyronine (FT3) (reference range, 3.1–6.8 pmol/L), free thyroxine (FT4) (reference range, 12.0–22.0 pmol/L), thyroid stimulating hormone (TSH) (reference range, 0.270–4.20 mIU/L), thyroglobulin antibody (TGAb) (cut-off level, 115 IU/mL), thyroid peroxidase antibody (TPOAb) (cut-off level, 34 IU/mL), and TSH receptor antibody (TRAb) (cut-off level, 1.75 IU/L) were measured automatically by Cobas c8000 machine via utilizing a commercial electro-chemiluminescence
immunoassay kit (Roche Diagnostics, Cobas 8000 analyzer e602 module). The intra-assay CVs of FT3, FT4, TSH, TGAb, TPOAb, and TRAb were 1.4–6.5%, 2.6–4.3%, 1.1–3.0%, 1.3–4.9%, 2.7–6.3%, and 0.9–7.6%, respectively, and the inter-assay CVs of coefficients of these indexes were 1.9–7.2%, 5.1–8.4%, 3.3–7.2%, 3.4–6.3%, 4.2–9.5%, and 1.9–11.4%, respectively.

Isolation of peripheral blood mononuclear cells (PBMCs)

Discontinuous Lymphoprep (Axis-Shield PoC As, Oslo, Norway) gradient centrifugation was used to collect PBMCs, which were resuspended in PBS containing 2% fetal bovine serum (FBS).

Flow cytometry analysis

Isolated PBMCs were blocked by Fc receptors before staining with a specific antibody. The PE-conjugated anti-CD19 antibody (Biolegend, San Diego, CA, USA) was used for surface staining following the manufacturer’s recommended protocol. Flow cytometry was conducted using a CyAn™ ADP Analyzer (Beckman, Coulter, Inc., Carlsbad, CA, USA) and data were analyzed using the FlowJo software (Tree Star, Ashland, Or, USA).

B-cell purification

Human CD19+ B cells were isolated by positive selection using PE-conjugated anti-CD19 (Biolegend) and anti-PE microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany)
according to the manufacturer’s instructions, and the purity was confirmed to be over 90.0% as judged through flow cytometry (Supplementary Fig. 1, see section on supplementary materials given at the end of this article).

**RNA extraction and purification**

RNA was extracted from the purified CD19+ B cells from 21 healthy individuals and 24 GD patients (patients 11–34, Table 1) using the TRIzol Reagent (Life technologies) as per the manufacturer’s protocol. The purity and concentrations of RNA samples were measured through NanoDrop ND-2000 spectrophotometry (Thermo Scientific). For Microarray experiments, total RNA was further purified using the RNeasy micro kit (QIAGEN) and RNase-Free DNase Set (QIAGEN).

**IncRNA and mRNA expression profiling**

CD19+ B cells from the peripheral blood of three GD patients and three healthy individuals were analyzed using microarrays. The RNA from the three GD patients (patients 11–13, Table 1) and the three healthy individuals were pooled separately. IncRNA and mRNA expression profiling was performed using Affymetrix GeneChip Human Transcriptome Array 2.0 microarrays (Affymetrix). Total RNA was amplified, labeled, and purified using the Affymetrix WT PLUS Reagent Kit (Affymetrix) to obtain biotin-labeled complimentary DNA (cDNA) according to the manufacturer's instructions. The labeled cDNA was hybridized and washed using a GeneChip Hybridization, Wash, and Stain Kit in a Hybridization Oven 645 and a Fluidics Station (all from Affymetrix) according to the manufacturer's instructions.

**Data acquisition**

Arrays were scanned in an Affymetrix GeneChip Scanner 3000 (Affymetrix). The Command Console Software (Affymetrix) was used to control the scanner and summarize probe cell intensity data (CEL file generation) with default settings. Then, raw data were normalized by the Expression Console.

**Enrichment analysis of DEGs**

To identify differential gene expression between healthy individuals and GD patients, a \(|\text{Fold Change (linear)}| \geq 2\) was considered as differently expressed. Gene Ontology (GO, http://www.geneontology.org) and Kyoto Encyclopedia of Genes and Genomes (KEGG, http://www.genome.jp/kegg) analyses were conducted on these differentially expressed genes (DEGs) to categorize the biological functions and pathways that they are associated with.

**IncRNA target prediction**

To identify the cis- or trans-regulatory targets of IncRNAs, two independent algorithms were used. The first algorithm searched for cis-regulated target genes with the help of gene annotations from UCSC (http://genome.ucsc.edu/). Genes located within 10 kb upstream or downstream of the IncRNAs were considered cis-regulated target genes. The second algorithm was based on mRNA sequence complementarity and prediction of RNA duplex energy. After a first round of screening using BLAST software, the RNAplex software was used to identify trans-regulated target genes (19). The RNAplex parameters were set as in previous studies (19).

**Construction of a protein–protein interaction (PPI) network**

To construct a PPI immune network involving differentially expressed genes (DEGs), all possible combinations of gene pairs were defined using InnateDB (http://www.innatedb.ca/), a publicly available database of the genes, proteins, experimentally verified interactions, and signaling pathways involved in the innate immune response (20). A final interaction network was constructed using the force-directed layout in the Cytoscape software. Differently expressed IncRNAs and their associated protein-coding genes were identified on the basis of chromosomal localization and BLAST sequence alignment.

**Quantitative real-time polymerase chain reaction (qRT-PCR) validation**

Total RNA from each sample was reverse transcribed into cDNA (Takara Biotechnology Co.). The cDNA was used in qRT-PCR assays to validate the presence of differentially expressed mRNAs and IncRNAs in B cells from GD patients. Real-time PCR amplifications were performed on an ABI 7500 thermocycler in 20 μL reaction volumes containing cDNA, primers, and SYBR Green Real-time PCR Master Mix (Toyobo Co., Ltd., Osaka, Japan). The primer sequences are listed in Table 2. Thermocycling was performed as follows: initial denaturation at 95°C for 30 s; 45 cycles of denaturation at 95°C for 5 s, and elongation at 60°C for 15 s. The specificity of the PCR reaction was controlled by assessing the melting curves. The ΔΔCt
method was used to compute the target gene expression relative to that of ACTB. All primers were synthesized and purified by Shanghai Sangon Biotechnology Co. Ltd. (Shanghai, China).

Statistical analysis

All statistical analyses were performed using SPSS 17.0 (SPSS) and PRISM 6.0 (GraphPad Software Inc.). Results were presented as the mean ± s.e.m. For comparison between two groups, Student’s unpaired t-test or unpaired t-test with Welch’s correction was used. $P$ values <0.05 were considered statistically significant.

Results

Increased peripheral B lymphocyte levels in GD patients

Compared with the normal reference values, the levels of free triiodothyronine (FT3), free thyroxine (FT4), TRAb, TPOAb, and TGAb in the blood of GD patients were remarkably high, whereas that of TSH was strikingly low. The basic information about the patients is shown in Table 1. The average proportion of CD19$^+$ B cells in the PBMCs of GD patients was 10.7%, which was significantly higher than that in the PBMCs of the healthy controls (Fig. 1).

Table 2 Primer information.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′–3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIRC3-F</td>
<td>TTGCTGTGATGGTGGACCTAGGTGTTG</td>
</tr>
<tr>
<td>BIRC3-R</td>
<td>GGTAACTGGCTTGAATCTGGAATGGTGATGAGCT</td>
</tr>
<tr>
<td>EAF2-F</td>
<td>GCCCTCCGACTTGGCCTATGACTTC</td>
</tr>
<tr>
<td>EAF2-R</td>
<td>GTCACCCTTTGACCTCACCACAACTTGAAGG</td>
</tr>
<tr>
<td>IL7R-F</td>
<td>AAGGCTTCTGAGGTGAATGGTGATGCAAGGATCC</td>
</tr>
<tr>
<td>IL7R-R</td>
<td>GCCAAGATGCAAACACAGACGACAGAGA</td>
</tr>
<tr>
<td>FYN-F</td>
<td>CATTGGCCTTGGGAGCTGAGTGGAGTT</td>
</tr>
<tr>
<td>FYN-R</td>
<td>AAATCAGCGAGCGAAGCCGAGAAG</td>
</tr>
<tr>
<td>TCL1A-F</td>
<td>AAGGGAGACACAGGTGCTGCAAGGACATC</td>
</tr>
<tr>
<td>TCL1A-R</td>
<td>AAATGACAGCGGGAACCGGAGAAG</td>
</tr>
<tr>
<td>SH2D1A-F</td>
<td>ACCGTGATACAGCACAATAGGCAATCATA</td>
</tr>
<tr>
<td>SH2D1A-R</td>
<td>AAAGGAAATGTGGTGGTCAGAAG</td>
</tr>
<tr>
<td>n335641-F</td>
<td>TCCCTCTGATGCTGGTCATGGG</td>
</tr>
<tr>
<td>n335641-R</td>
<td>GCACCTCATCTCCTCTCTCAG</td>
</tr>
<tr>
<td>TCONS_00022357-XLOC_010919-F</td>
<td>GGACGGAGAGACAGGAGGAGG</td>
</tr>
<tr>
<td>TCONS_00022357-XLOC_010919-R</td>
<td>ACTGACCGAGCAGTGAGGACAG</td>
</tr>
<tr>
<td>n337845-F</td>
<td>CAGGTAGCTGAGCTATGAGGAG</td>
</tr>
<tr>
<td>n337845-R</td>
<td>TCTTACTGATGCTGGTGCAGAAG</td>
</tr>
<tr>
<td>ACTB-F</td>
<td>AAAGGGAAATGTGGTGGTCAGAAG</td>
</tr>
<tr>
<td>ACTB-R</td>
<td>TCCCTCTGATGCTGGTCATGGG</td>
</tr>
</tbody>
</table>

F, forward; R, reverse.

Figure 1 The ratio of CD19$^+$ B cells to PBMCs increased in GD patients. (A) Representative percentage of CD19$^+$ B cells in healthy individuals and new-onset GD patients as determined by flow cytometry. (B) CD19$^+$ B-cell frequencies in total PBMCs were measured in 34 healthy individuals and 34 new-onset GD patients, with each point representing a single person. $***P < 0.001$. The data shown are the mean ± s.e.m.
Differential mRNA and lncRNA expression profiles in B lymphocytes from GD patients

For revealing the mechanism of the abnormality of B cells in GD, the expression profiles of lncRNA and mRNA were investigated. In total, 604 lncRNAs were differentially expressed (DE-lncRNAs) between the purified B cells of healthy individuals and those of patients with GD. Of those, 303 were upregulated in GD patients relative to the healthy controls and the other 301 were downregulated in GD patients compared with the healthy controls (Fig. 2A). Furthermore, a total of 410 differentially expressed genes (DEGs) were also identified, 331 (80.7%) of which were downregulated in GD patients compared with the healthy controls and 79 (19.3%) of which were upregulated in GD patients compared with the healthy individuals (Fig. 2B). Our data showed that molecular events in peripheral blood B cells, such as lncRNAs and mRNAs, have changed during GD development.

Functional enrichment analysis

Gene ontology (GO) and KEGG pathway enrichment analyses were performed to identify the functions of the DEGs. GO project used to classify the significantly regulated genes according to biological process (BP), cellular component (CC), and molecular function (MF). The DEGs were significantly enriched with GO terms related to the immune response. The most common biological processes associated with the DEGs were immune response, T cell receptor signaling pathway, and regulation of immune response (Fig. 3A). Consistent with the enriched GO categories, the KEGG pathway analysis showed that the major signal-transduction pathways associated with the DEGs were related to immune signaling, such as T cell receptor signaling pathway, natural killer cell mediated cytotoxicity, and antigen processing and presentation (Fig. 3B).

Confirmation of the microarray results by qRT-PCR

To confirm the results of the microarray analysis, two mRNAs that were upregulated (BIRC3 and EAF2) and two mRNAs that were downregulated (IL7R and FYN) in the B lymphocytes of the patients with GD were randomly selected and analyzed by qRT-PCR in 21 healthy individuals and 24 GD patients. As expected, the expression levels of BIRC3 and EAF2 were increased in GD patients relative to those in the healthy individuals, whereas the expression levels of IL7R and FYN were decreased in GD patients relative to those in the healthy individuals, consistent with the microarray results (Fig. 4).

Analysis of differentially expressed lncRNA and mRNA associations

Because the roles of lncRNAs are mainly executed through effects on protein-coding target genes, the cis-regulatory targets of DE-lncRNAs were identified on the basis of the distance between the lncRNAs and their neighboring protein-coding genes (adjacent upstream or downstream) and the trans-regulatory targets of DE-lncRNAs were analyzed by RNAplex software. The results showed that 84 DE-lncRNAs were significantly associated with 81 differentially expressed protein-coding genes to generate...
85 lncRNA–mRNA pairs. Of the 84 lncRNA candidates, 61 were identified as sense lncRNAs, 5 as antisense lncRNAs, 3 as bidirectional lncRNAs, 2 as intergenic lncRNAs, and 13 as other lncRNAs.

To identify lncRNA–mRNA associations and evaluate the impact of the DE-lncRNAs on gene expression, the predicted target genes of the DE-lncRNAs were identified and combined with the mRNA transcriptomes of the study participants. Because the patients with GD had higher percentages of peripheral B lymphocytes than the healthy controls, we hypothesized that the regulation of B-cell proliferation was abnormal in the patients with GD. Fourteen pairs of DE-lncRNAs and adjacent protein-coding genes associated with cell proliferation were identified (Table 3). Among the identified genes, those encoding TCL1A (21, 22) and SH2D1A (23, 24, 25) have been reported to be involved in B-cell proliferation and survival. Thus, n335641–TCL1A, TCONS_00022357–XLOC_010919–TCL1A, and n337845–SH2D1A were identified as candidate lncRNA–mRNA pairs that were differentially regulated between GD patients and healthy individuals.

To validate the candidate lncRNA–mRNA pairs in independent patients with GD, the expression of the pairs was measured in an additional 21 healthy individuals and 24 GD patients by qRT-PCR. As expected, the lncRNAs n335641 and TCONS_00022357–XLOC_010919 were upregulated, whereas IL7R and FYN were downregulated significantly in purified B cells from GD patients as compared with healthy controls. The data shown are the means ± s.e.m. Three replicates showed similar results. ***P < 0.001.
roles and mechanisms of these lncRNA–mRNA pairs in dysfunction of B cells in patients with GD were not clear and need to be studied in depth.

**Correlation analysis for DEGs and DE-lncRNAs**

To reveal the PPI networks of the DEGs, immunologically focused network analyses of 126 DEGs (13 upregulated and 113 downregulated in the patients with GD relative to the controls) and 193 potential interactions were performed using the InnateDB platform (Fig. 6). The downregulated and upregulated DEGs with high degrees are illustrated in Table 4. Accordingly, the downregulated DEG *SH2D1A* had a connective degree of 7, potentially interacting with FYN (26, 27, 28, 29, 30), LCK (28), CD247 (31), CD5 (32) and so on, and the upregulated DEG *TCL1A* had a connective degree of 2, potentially interacting with JUNB (33) and FOS (33). To further uncover the correlations between the DEGs and the DE-lncRNAs, the potential functions of the DE-lncRNA–protein pairs were identified. In total, there were 28 regulatory relationships between DEGs and DE-lncRNAs, including 27 DEGs and 28 DE-lncRNAs (3 upregulated in GD and 25 downregulated in GD). For instance, *n335641* and *XLOC_010919* were predicted to regulate *TCL1A*, whereas *n337845* was predicted to regulate *SH2D1A* (Fig. 6).

**Table 3** Deregulated lncRNAs and their corresponding protein-coding genes (associated with cell proliferation).

<table>
<thead>
<tr>
<th>lncRNA</th>
<th>Log2(FC)-lncRNA</th>
<th>Type</th>
<th>mRNA</th>
<th>Log2(FC)-mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>n326505</td>
<td>−1.21</td>
<td>antisense</td>
<td>CDC14A</td>
<td>−1.39</td>
</tr>
<tr>
<td>n332938</td>
<td>−1.23</td>
<td>sense</td>
<td>EIF2AK2</td>
<td>+1.18</td>
</tr>
<tr>
<td>n407800</td>
<td>−2.38</td>
<td>sense</td>
<td>HPGD</td>
<td>−1.06</td>
</tr>
<tr>
<td>n409270</td>
<td>−2.76</td>
<td>sense</td>
<td>TCF7</td>
<td>−1.07</td>
</tr>
<tr>
<td>TCONS_00010118-XLOC_004584</td>
<td>−1.09</td>
<td>other</td>
<td>NDFIP1</td>
<td>−1.90</td>
</tr>
<tr>
<td>n384644</td>
<td>−2.58</td>
<td>sense</td>
<td>IL6ST</td>
<td>−3.09</td>
</tr>
<tr>
<td>n335470</td>
<td>−3.84</td>
<td>sense</td>
<td>TNFAIP3</td>
<td>−3.72</td>
</tr>
<tr>
<td>n378743</td>
<td>−2.72</td>
<td>bidirectional</td>
<td>PRKCQ</td>
<td>−2.13</td>
</tr>
<tr>
<td>n335567</td>
<td>−1.58</td>
<td>bidirectional</td>
<td>IFITM1</td>
<td>−1.16</td>
</tr>
<tr>
<td>TCONS_00022357-XLOC_010919</td>
<td>+2.31</td>
<td>sense</td>
<td>TCLI1A</td>
<td>+1.05</td>
</tr>
<tr>
<td>n335641</td>
<td>+1.97</td>
<td>sense</td>
<td>TCLI1A</td>
<td>+1.05</td>
</tr>
<tr>
<td>n338211</td>
<td>−1.21</td>
<td>other</td>
<td>ERN1</td>
<td>−1.20</td>
</tr>
<tr>
<td>n409231</td>
<td>−2.30</td>
<td>sense</td>
<td>TYROBP</td>
<td>−2.66</td>
</tr>
<tr>
<td>n337845</td>
<td>−1.50</td>
<td>intergenic</td>
<td>SH2D1A</td>
<td>−1.39</td>
</tr>
</tbody>
</table>

FC, fold change.

Figure 5

Validation of mRNA and IncRNA microarray results by qRT-PCR. (A, B, C, D, and E) The RNA of purified CD19⁺ B cells from another 21 healthy individuals and 24 GD patients were used to verify the significantly regulated IncRNA and their related target mRNA. (A and B) IncRNA n335641 and TCONS_00022357-XLOC_010919, which potentially regulate TCL1A were upregulated. (C) IncRNA n337845, which potentially regulates SH2D1A was downregulated. (D and E) TCL1A mRNA was upregulated, whereas SH2D1A was downregulated significantly in purified B cells from GD patients as compared with healthy controls. The data shown are the means ± S.E.M. Three replicates showed similar results. ***P < 0.001.
In GD, B cells are the source of TRAb against TSHR (2, 6, 34). It is widely accepted that the quantity and function of B cells are anomalous in GD patients (7, 8, 35). We previously reported that B cells infiltrate the thyroid tissues of patients with GD (36). In the present study, we found that the percentage of peripheral B cells was higher in GD patients than in healthy individuals; however, the molecular determinants of that difference remain unknown. Microarray analysis identified 604 lncRNAs and 410 mRNAs with differential expression between the B cells of patients with GD and those of healthy individuals. The differential expression of a subset of those lncRNAs and mRNAs was confirmed by qRT-PCR. GO and KEGG pathway analyses of the DEGs showed enrichment of biological functions and signaling pathways involved in the immune response, such as T-cell receptor signaling pathway, regulation of immune response, and antigen processing and presentation.

**Table 4** The downregulated genes with the connectivity degree ≥6 and upregulated genes with the connectivity degree ≥2 in the protein–protein interaction network.

<table>
<thead>
<tr>
<th>Downregulated DEGs</th>
<th>Connective degree</th>
<th>Downregulated DEGs</th>
<th>Connective degree</th>
<th>Upregulated DEGs</th>
<th>Connective degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>FYN</td>
<td>22</td>
<td>LGALS1</td>
<td>8</td>
<td>BIRC3</td>
<td>4</td>
</tr>
<tr>
<td>ZAP70</td>
<td>18</td>
<td>GAPDH</td>
<td>7</td>
<td>NSF</td>
<td>3</td>
</tr>
<tr>
<td>IL7R</td>
<td>18</td>
<td>HSPAS</td>
<td>7</td>
<td>IFIT3</td>
<td>3</td>
</tr>
<tr>
<td>LCK</td>
<td>17</td>
<td>SH2D1A</td>
<td>7</td>
<td>IFIT1</td>
<td>3</td>
</tr>
<tr>
<td>ARRB1</td>
<td>13</td>
<td>CD5</td>
<td>7</td>
<td>STX7</td>
<td>3</td>
</tr>
<tr>
<td>LCP2</td>
<td>12</td>
<td>TNFAIP3</td>
<td>6</td>
<td>TCL1A</td>
<td>2</td>
</tr>
<tr>
<td>CD247</td>
<td>10</td>
<td>CD2</td>
<td>6</td>
<td>XAF1</td>
<td>2</td>
</tr>
<tr>
<td>HSPA8</td>
<td>10</td>
<td>CD3E</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4</td>
<td>9</td>
<td>TUBB</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATXN1</td>
<td>8</td>
<td>ITK</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The PPI network constructed by InnateDB (experimentally validated interactors) showed that the mRNA expression of most of the DEGs was decreased in GD compared with that in healthy individuals. To our knowledge, no studies have examined lncRNA and mRNA expression profiles of peripheral B cells in GD.

The functions of the DE-lncRNAs might provide hints about the mechanisms of B cells dysfunction and pathogenesis in GD. Previous research reported that lncRNAs can regulate the expression of neighboring protein-coding genes via cis-acting mechanisms (37), and we predicted that some of the DE-lncRNAs in GD would function based on a cis-acting mechanism. Because the functions of most of the DE-lncRNAs were not known, we predicted them on the basis of the positional relationships between the lncRNAs and nearby genes. lncRNAs can also regulate the expression of protein-coding genes via trans-acting mechanisms including interactions with chromatin modifiers, transcription factors, splicing factors, or RNA-decay machinery (38, 39, 40). Owing to this, 85 lncRNA–mRNA pairs were identified between patients with GD and healthy controls.

In our study, we found the percentage of peripheral B lymphocytes was increased in GD patients. Previous research also reported that the percentage and absolute count of peripheral B lymphocytes were markedly increased in GD patients than in normal controls, but the proportions and absolute counts of B lymphocytes were reduced to normal in euthyroid patients with GD after antithyroid drug therapy (41). Thus, we hypothesized that abnormal regulation of proliferation was the cause of abnormal B-cell function in GD. Therefore, we focused on protein-coding genes linked to cell proliferation. After combining the lncRNA and mRNA expression data, we identified 14 pairs of DE-lncRNAs and corresponding protein-coding genes associated with cell proliferation (Table 3). Among those protein-coding genes, TCL1A acts as co-activator of AKT kinase and thus promotes AKT-induced cell survival and proliferation (21), which are part of the normal development of early B cells and T cells (42). TCL1A has also been reported to regulate cell survival and cell death pathways in human naive and IgM memory B cells (22). The PPI network shows that TCL1A is related to the pathogenesis of B-cell chronic lymphocytic leukemia (B-CLL) via physical interactions with JUNB and FOS (33), which play important roles in the regulation of thyroid cell function (43, 44, 45). That suggests that TCL1A might be involved in thyroid autoimmune disorders through interactions with JUNB and FOS. In addition, X-linked lymphoproliferative disease 1 (XLP1), manifested as lymphomas resulting from uncontrolled B-cell proliferation upon infection with Epstein-Barr virus, is a primary immunodeficiency caused by mutations or deletions of SH2D1A gene, which encodes the SLAM-associated protein (SAP) (46, 47). B-lymphoblastoid cell lines (B-LCLs) from XLP1 patients have an intrinsic defect that affects cell activation, apoptosis, and proliferation (23), and SAP-deficient mice also have a marked defect in B-cell proliferation (24, 48). SAP could form a ternary complex with the Lyn kinase and the inhibitory IgG Fc receptor FCGR2B to regulate B-cell proliferation and survival, leading to a decrease in B-cell proliferation and an increase in activation-induced cell death (25). Among the proteins that interacted with SH2D1A in the PPI network, LCK and CD247 were previously shown to be hypermethylated in CD4+ T cells of GD patients (49). Furthermore, CD247 is a susceptibility gene inAITD (50). CDS1 peripheral B lymphocytes were increased in GD and were associated with disease activity (51, 52, 53), indicating that SH2D1A may also play roles in GD pathogenesis via its interactors. The lncRNAs TCONS_00022357-XLOC_010919 and n335641 were predicted to target TCL1A, and the lncRNA n337845 was predicted to target SH2D1A. The differential expression of both target genes in GD was confirmed by qRT-PCR. The results indicate that the n335641–TCL1A, TCONS_00022357-XLOC_010919–TCL1A, and n337845–SH2D1A lncRNA–mRNA pairs might be novel biomarkers associated with the proliferation of B lymphocytes in GD. Further study should be conducted to verify whether those are biomarkers for the diagnosis, treatment, and prognosis of GD.

In summary, we analyzed lncRNA and mRNA expression profiles in B cells to identify molecules that influence GD pathogenesis by regulating B-cell function, proliferation, and survival. Our study is preliminary and therefore has some limitations. First, we only isolated B lymphocytes from peripheral blood and not from the thyroid or local neck lymph nodes surrounding the thyroid. B lymphocytes of different origins might contain different genetic information. Second, the lncRNA database is updated continuously, so the datasets used in our study might be limited. Therefore, the lncRNA signatures might need further identification. Furthermore, the relationships between lncRNAs and mRNAs predicted by the two algorithms were not verified by experiments in vivo and in vitro. Third, the mechanisms by which the lncRNA–mRNA pairs influence B-cell proliferation have not been uncovered. Further study should be carried out to verify whether the lncRNA–mRNA pairs can be regarded as biomarkers for the diagnosis, treatment, and prognosis of GD. Except for
three pairs of IncRNA–mRNA pairs mentioned previously, other differently expressed IncRNA–mRNA pairs may also be related to B-cell proliferation or other abnormal B-cell functions such as B-cell activity or suppression ability. We believe, however, that the IncRNA–mRNA expression pattern provides a basis for understanding the pathogenesis of GD and will lead to diagnostic biomarkers or therapeutic targets in future.

**Conclusions**

Our study provides an integrated analysis of IncRNA and mRNA expression profiles in peripheral B cells of GD patients. The results suggest that the IncRNA–mRNA pairs n335641–TCLIA, TCONS_00022357–XLOC_010919–TCLIA, and n337845–SH2D1A participate in GD pathogenesis by modulating B-cell proliferation and survival. The identified IncRNA–mRNA pairs should be investigated further as potential biomarkers and therapeutic targets in GD, which might provide a better understanding of the molecular basis of GD.

**Supplementary materials**

This is linked to the online version of the paper at https://doi.org/10.1530/EC-20-0045.

**Acknowledgements**

The authors appreciate the time and effort given by all participants. The authors also thank Lingfang Zhang for her help in bioinformatics analysis.

**References**

transcribed long noncoding RNA. PNAS 2019 116 16410–16419. (https://doi.org/10.1073/pnas.1904180116)
24 Schwartzberg PL, Mueller KL, Qi H & Cannons JL. SLAM receptors and SAP influence lymphocyte interactions, development and function. Nature Reviews Immunology 2009 9 39–46. (https://doi.org/10.1038/ml2456)
30 Li C, Josef C, Jia CY, Han VK & Li SS. Dual functional roles for the X-linked lymphoproliferative syndrome gene product SAP/SH2D1A in signaling through the signaling lymphocyte activation molecule (SLAM) family of immune receptors. Journal of Biological Chemistry 2003 278 3852–3859. (https://doi.org/10.1074/jbc.M206649200)


Received in final form 14 April 2020
Accepted 21 April 2020
Accepted Manuscript published online 21 April 2020