AMERICAN SOCIETY of
GENE & CELL
THERAPY

Integrative Analysis of lncRNAs in Th17 Cell Lineage to Discover New Potential Biomarkers and Therapeutic Targets in Autoimmune Diseases

Shohreh Teimuri,[1,5] Aref Hosseini,[2,5] Ahmad Rezaenasab,[1] Kamran Ghaedi,[1,2] Elahe Ghoveud,[2] Masoud Etemadifar,[3] Mohammad Hossein Nasr-Esfahani,[2] and Timothy L. Megraw[4]

[1]Division of Cellular and Molecular Biology, Department of Biology, Faculty of Sciences, University of Isfahan, Isfahan, Iran; [2]Department of Cellular Biotechnology, Cell Science Research Center, Royan Institute for Biotechnology, ACECR, Isfahan, Iran; [3]Department of Neurology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran; [4]Department of Biomedical Sciences, Florida State University College of Medicine, Tallahassee, FL, USA

Th17 cells play a critical role in the pathogenesis of autoimmune diseases, including multiple sclerosis, rheumatoid arthritis, systemic lupus erythematosus, Sjogren's syndrome, and inflammatory bowel disease. Despite the extensive investigation into this T cell lineage, little is understood regarding the role of Th17 lineage-specific lncRNAs (long non-coding RNAs) > 200 nt. lncRNAs may influence disease through a variety of mechanisms; their expression could be regulated by SNPs. lncRNAs can also affect the expression of neighboring genes or complementary miRNAs, and their expression may have lineage-specific patterns. In the system biology study presented here, the effective lncRNAs from different criteria were predicted for each autoimmune disease, and we then evaluated their expression levels in 50 MS patients compared to 25 controls using qRT-PCR. We identified changes in the expression levels of AL450992.2, AC009948.5, and RP11-98D18.3 as potential peripheral blood mononuclear cell (PBMC) biomarkers for MS among our studied lncRNAs in which co-expression analysis of AL450992.2 had the most AUCs, and the relationship to _RORC_ was also assessed. We propose that the recurrently deregulated lncRNAs identified in this report could provide a valuable resource for studies aimed at delineating the relationship between functional lncRNAs and autoimmune disorders.

INTRODUCTION

Approximately 7% to 9% of people around the world suffer from autoimmune and immune-related diseases (AIDs) from a variety of heterogeneous disorders.[1] Th17 cell subsets with specific developmental requirements and functions play critical roles in the pathogenesis of autoimmune diseases.[2,3] Although transcriptional regulation of Th17 cell differentiation has been extensively studied, there is little understanding of the post-transcriptional regulation of Th17 cells.

Recently, reports suggest that long non-coding RNAs (lncRNAs), i.e., non-coding RNAs > 200 nt, have more stability compared to protein-coding genes in body fluids and tissues, have tissue-specific patterns

of expression, and are relatively easy to detect through various techniques in liquid biopsies. lncRNAs are attractive disease biomarkers and therapeutic targets, since they may modulate gene expression through several mechanisms.[4–7] In order to manifest the associated lncRNAs with the disease, several possibilities exist, including the possibility that they could be influenced by SNPs, affect the expression of neighboring genes or complementary microRNAs (miRNAs), or affect lineage-specific expression. In this paper, we address all of these influences.

The unique functions of each cell type are performed by its specialized gene expression. Therefore, regulating lineage-specific lncRNAs could affect cell fate and play a critical role in the pathology of disease.[8] Roughly 50% of protein-coding genes within 50 kb were co-expressed with the respective lncRNA gene, and this correlation will decrease by each kilobase farther away from the genomic position of the protein-coding gene.[9]

Integration of AID-associated SNPs with genome functional region data demonstrate that ~90% of SNPs are associated with regulatory DNA enhancer or promoter regions rather than with protein-coding regions.[10,11] Reported trait and/or disease-associated SNPs (TASs) were significantly overrepresented only in nonsynonymous sites

Figure 1. Association of Th17 Cell Differentiation Genes and lncRNAs

Th17 cell differentiation genes were deduced through data mining, and their proximities to selected lncRNAs were evaluated through Python programming language. Desired lncRNAs with the appropriate distance are indicated by red circles.

sequencing (RNA-seq) data. These lncRNAs included 54 transcripts (Table S1).

lncRNA Selection from Adjustment to Th17 Cell Differentiation Genes

From mining the literature, we identified 116 genes encoding proteins involved in Th17 cell differentiation and categorized them according to their signaling pathway involvement (Table S2). From this list, we identified 5 lncRNAs that neighbor the genes involved in Th17 cell differentiation with a distance less than 50 kb and for which 4 of them were also co-expressed with *IL21R*, *RORC*, and

BATF, which are also expressed in Th17 cell differentiation (Figure 1).

lncRNA Selection from Adjustment to Differentially Expressed Genes in AIDs

In the next phase, we analyzed array datasets of AID genes by comparing patient PBMC samples to controls and identified the most differentially expressed protein-coding genes (MS, 405; SLE, 102; SS, 55; RA, 23; and IBD, 238) (Table S3). We then examined the proximity of these genes to candidate lncRNA genes and recorded lncRNAs that were located less than 50 kb from protein-coding genes as associated or linked. We retrieved 1 lncRNA linked to MS and 3 linked to IBD genes with a potentially *cis*-regulatory role in neighboring genes, one of which was shared between these two groups (Figure 2).

lncRNA Selection from Adjustment to the AID-Associated SNPs

SNPs associated with each AID were retrieved from the Genome-Wide Association Studies (GWAS) catalog (MS, 210; SLE, 307; RA, 365; SS, 26; and IBD, 345) (Table S4). Accordingly, we predicted 3 lncRNAs related to AIDs via SNPs, including two associated with MS and one associated with SLE that was located less than 50 kb away from selected SNPs (Figure 3).

lncRNA-miRNA Interactions and Associated Diseases

Next, we inserted the sequence of every single exonic lncRNA transcript in FASTA format into the "LncRNAs Input" panel of LncDisease software, as shown in Table 1.

and 5-kb promoter regions compared to SNPs randomly selected from genotyping arrays.[12] Logistically, genomic regions containing SNPs associated with disease could include disease-associated genes like lncRNAs.

lncRNAs may have critical roles through targeting miRNA in a wide range of biological processes; accordingly, competing endogenous RNA theories could explain how lncRNAs regulate coding genes via competitively binding with their complementary miRNA.[13,14]

Identifying influential biomarkers could potentially shed light on the underlying pathogenic mechanisms in AIDs. From a practical standpoint, the most useful autoimmune biomarkers will be those measurable in peripheral blood mononuclear cells (PBMCs) and serum. The utility of most of these markers is limited by their restriction to relatively inaccessible anatomic sites.

One emerging view is that lncRNAs could be novel molecules for disease diagnosis and therapy.[15] In this system biology study, we aimed to deduce effective lncRNAs in several autoimmune diseases, including multiple sclerosis (MS), rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), Sjogren's syndrome (SS), and inflammatory bowel disease (IBD) through several criteria, and we show that the expression levels of some are significantly elevated in MS.

RESULTS
lncRNA Selection

Among the currently known lncRNAs, we selected 26 Th17 cell-lineage-specific lncRNAs that derived from previous RNA-
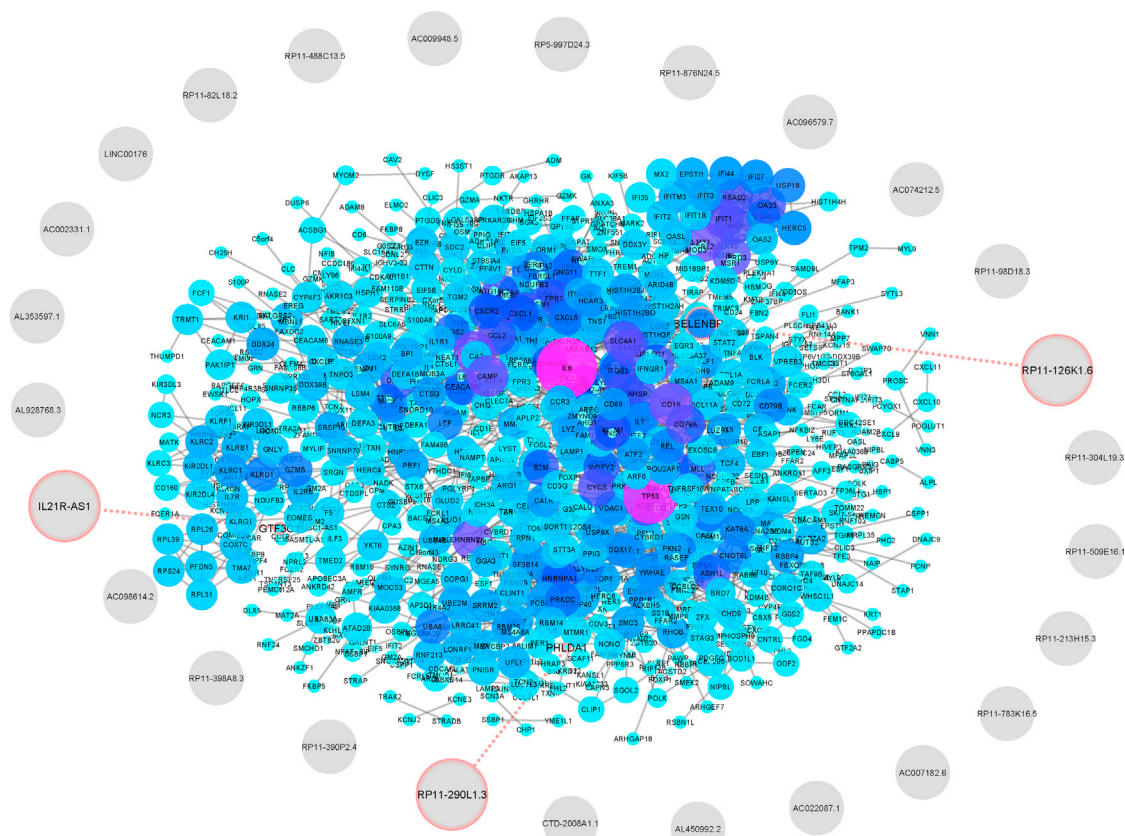
Figure 2. Association of lncRNAs and Array-Expressed Genes

Differentially expressed mRNAs in AID samples compared to controls were retrieved from the Gene Expression Omnibus (GEO) database, and their proximities to selected lncRNAs were analyzed by Python programming language, inputted to STRING-db, and visualized in Cytoscape. Appropriate lncRNAs in logical distance are indicated with red circles.

Putative lncRNAs Associated with MS

We selected 8 putative lncRNAs that, in addition to Th17 cell-lineage specificity, were also derived from at least one of the other criteria: 4 lncRNAs involved in Th17 cell-differentiation pathways, 1 lncRNA selected from an array expression criterion, 1 derived from LncDisease software, 1 lncRNA related to a SNP association criterion, and 1 derived from both SNP association and LncDisease software. AL450992.2 was associated with miRNA; therefore, we only evaluated this transcript for SNP association. This lncRNA was associated with MS (Table 2), so we explored its expression levels in a case-control study by real-time qPCR analysis.

Patient Characteristics

The clinical and pathological data including sex, age, illness duration, and score on the Expanded Disability Status Scale (EDSS) have been recorded. Statistical analyses showed no significant differences between all samples in each group (Table 3).

Significant Changes in Levels of Expression of Specific lncRNAs in MS Patients

We evaluated the expression levels of selected lncRNAs in the PBMCs of MS patients compared to the control group. Real-time qPCR anal-

ysis showed a significant elevation of RP11-126K1.6 expression in the PBMCs of MS patients compared to those of the healthy controls (p = 0.023). Furthermore, RP1198D18.3, AL450992.2, and AC007182.6 lncRNAs associated with the Th17 cell differentiation genes, and AC009948.5—an lncRNA associated with miRNA and SNP—were downregulated in patients compared to the healthy volunteers (p = 0.001, p < 0.001, p = 0.041, and p < 0.001, respectively). In contrast, expression levels of these lncRNAs showed no statistically significant difference between the MS samples and control group: AL928768.3 as a SNP-related lncRNA, RP11-290L1.3 as a miRNA-associated lncRNA, and IL21R-AS1 as associated with a differentially expressed protein-coding gene (Figure 4).

Significant Changes in Levels of Expression of Specific lncRNAs in Relapsing- and Remitting-Phase Patients

RP11-126K1.6, a lncRNA adjusted to array-expressed genes, had elevated expression levels in the remitting phase of disease compared to healthy volunteers (p = 0.039). RP1198D18.3 and AL450992.2, lncRNAs close to genes involved in Th17 cell differentiation, which were evaluated in relapsing-phase patients compared to controls using real-time qPCR, showed significantly decreased transcript levels (p = 0.030 and p < 0.001, respectively) and a decreased RP1198D18.3
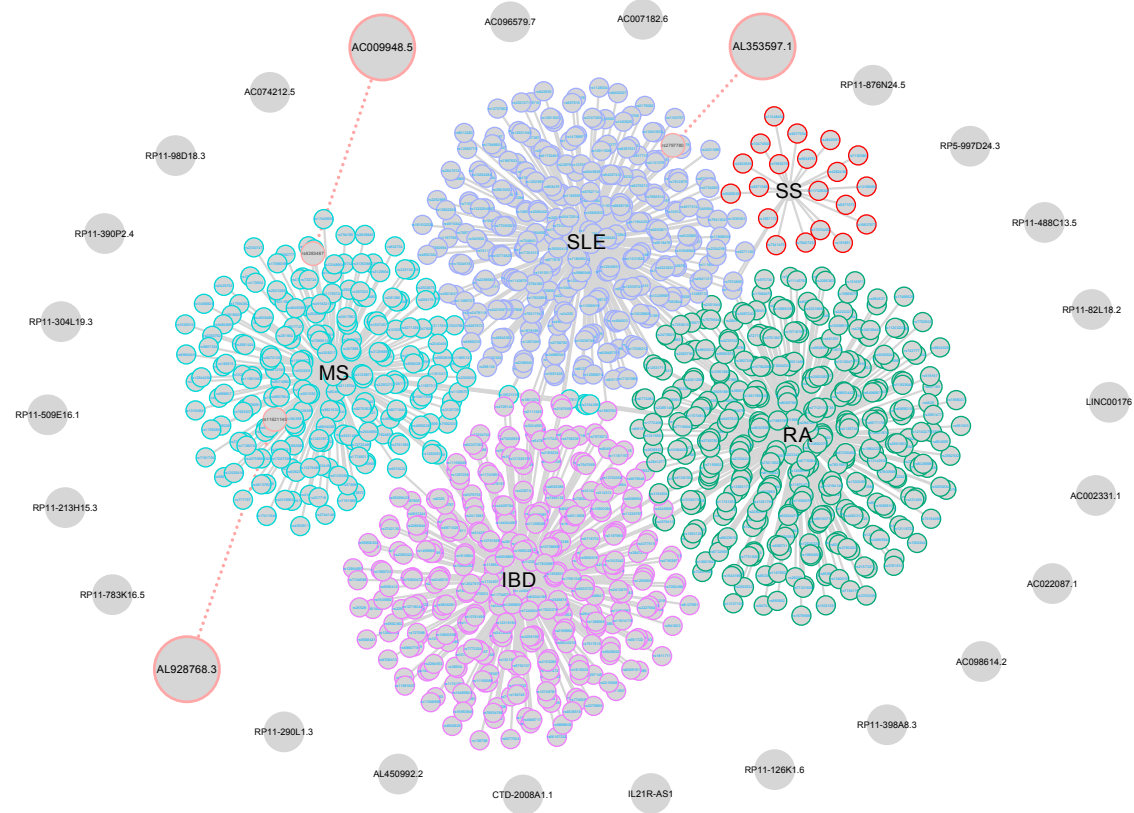
Figure 3. SNP-lncRNA Association

SNPs associated with selected AIDs were determined, and their distance from each lncRNA locus was assessed by Python programming language and visualized via Cytoscape. Selected lncRNAs in this analysis were connected to the associated SNP, as indicated with red circles.

expression level in relapsing-phase compared to remitting-phase patients (p = 0.016). Data indicated that AL450992.2 was downregulated in remitting-phase rather than healthy volunteers (p < 0.001). However, these values were not significant for IL21R-AS1 and AC007182.6 between the relapsing-phase, remitting-phase, and control groups.

Table 1. Deduced Potentially Associated lncRNAs according to miRNA Interaction Deduced from LncDisease Software

Disease	lncRNA
MS	AC009948.5-007
	RP11-290L1.3-001
IBD	LINC00877-001
RA	RP11-290L1.3-001
	RP11-488C13.5-002
SS	AC009948.5-001
SLE	RP11-213H15.3-003
	RP11-82L18.2-001
	AC098614.2-001

We assessed the expression levels of RP11-290L1.3 as a miRNA-associated lncRNA; however, the differences between the relapsing-phase, remitting-phase, and control groups were not as significant as for AL928768.5, a SNP-related lncRNA. However, there was a significant decrease in AC009948.5, a lncRNA associated with both miRNA and SNP, in the relapsing phase compared to the controls (p = 0.016) and in the relapsing phase compared to the remitting phase (p < 0.001) (Figure 5).

ROC Curve

We performed a comparative analysis of the lncRNAs generated, and then we ranked them as biomarker candidates according to decreasing area under the receiver operating characteristic (ROC) curve (AUC) and graphed the results. The first three lncRNAs ranked by AUC are AL450992.2, with an AUC of 0.866 (95% confidence interval [CI] = 0.770–0.962, p < 0.001); AC009948.5, with an AUC of 0.859 (95% CI = 0.761–0.958, p < 0.001); and RP11-98D18.3, with an AUC of 0.793 (95% CI = 0.682–0.904, p < 0.001) (Figure 6).

Correlation of lncRNAs Adjusted to *RORC*

We assessed *RORC* expression levels in the PBMCs of patients compared to those of controls, which were adjusted to RP11-98D18.3

Table 2. Putative lncRNAs in MS Pathology Derived from System Biology Analyses

Criterion	lncRNA
SNP association with AIDs	AC009948.5
	AL928768.3
miRNA interactions	AC009948.5-007
	RP11-290L1.3-001
Differentially expressed genes in AIDs	RP11-126K1.6
Th17 cell differentiation genes	IL21R-AS1
	RP11-98D18.3
	AL450992.2
	AC007182.6

and AL450992.2. AL450992.2 relative expression fold change in PBMCs and *RORC* in MS patients showed a significant positive correlation (Pearson's correlation = 0.306, $p < 0.05$) (Figure 7), while no correlation was observed between RP11-98D18.3 and *RORC* (Pearson's correlation = 0.210, $p = 0.11$); moreover, there was no significant difference in *RORC* expression levels in MS patients compared to those of the controls.

DISCUSSION

To identify disease-associated lncRNAs, we applied an integrated method considering several criteria. Investigation of the chromosomal locations of obtained results of criteria, including Th17 cell differentiation genes, differentially expressed genes in AIDs; and AIDs-associated SNPs with lineage-specific lncRNAs indicated that ~4.31% of Th17 cell differentiation genes were neighbors to the selected lncRNAs, while only ~0.04% of genes retrieved from array expression were in association with lncRNA and only ~0.02% of SNPs were in association with lncRNAs involved in disease suseptibility, which points to the fact that genes encoding Th17 cell differentiation are enriched in the genomic regions containing genes encoding lineage-specific lncRNAs. Hence, by applying these techniques to cell types involved in different diseases and protein-coding genes in specific cells, we could potentially identify more disease-associated lncRNAs.

Effector Th17 cell populations can cause CNS inflammation and demyelinating lesions.[16] Accordingly, as demonstrated earlier, pro-

tein-coding genes associated with lncRNAs[9] (RP11-98D18.3, AL450992.2, AC007182.6, and IL21R.AS1) were decreased in effector Th17 cells compared to primary cell cultures,[9] and our investigations showed consistency with these results, as three of these lncRNAs (RP11-98D18.3, AL450992.2, and AC007182.6) were downregulated in MS patients compared to the controls.

Real-time qPCR of RP11-126K1.6 showed upregulation in MS patients compared to the controls, which was in accordance with its differentially expressed gene in MS samples, as this study evaluated expressed genes in MS patients compared to the controls with no discriminatory group in patients, consequently, in order to verify that lncRNAs associated with this criterion-specific sampling of patients on differentially expressed genes could influence the results.

AC009948.5 expression results indicated that this putative lncRNA could, indeed, be associated with decreased level of MS pathogenesis, as it was identified from SNP (rs9283487) association and LncDisease criteria, while the reported risk SNP (rs11621145) from GWAS that is associated with the expression of AL928768.7 was not dysregulated, which suggests functional mechanisms underlying the findings from GWAS that are more complex than regulatory variants or expression levels of nearby lncRNA genes. Chromatin-looping models could explain how a contributing lncRNA is not necessarily the closest gene that is influenced by SNPs.[17]

The insignificant expression level of RP11-290L1.3 could be due to the limitations of LncDisease software for determining putative lncRNAs associated with miRNAs, as the exact mechanism by which miRNAs regulate lncRNAs is not clear, which results in the prediction that lncRNA targets of miRNAs will have high false positives and high false negatives.

The expression of lncRNAs is more disease and tissue specific than that of protein-coding genes and are more associated with their biological function.[18] lncRNAs as regulators of diverse biological processes by immune cells and the molecular mechanism of autoimmunity might be associated with numerous autoimmune diseases. Even recently, microvesicles found to contain lncRNAs involved in AIDs have been reported.[19–21] Previous studies demonstrated the therapeutic and biomarker roles of lncRNAs in various diseases such as RA, IBD, and SLE. These include the effect of large intergenic noncoding

Table 3. Patient Clinical and Demographic Data

Characteristics	RRMS Patients	Relapsing	Remitting	Control	p Value Comparison between Patients and Control/Relapsing, Remitting, and Control Samples
Sex: female/male	38/12	19/6	19/6	17/8	0.46/0.76
Age (and SD)	34 (8.30)	32.48 (7.07)	35.52 (9.26)	32.84 (6.06)	0.49/0.23
Duration of illness (and SD)	3.08 (4.41)	0	6.16 (4.47)	–	–
Disease severity (EDSS)	1.47	1.56	1.38	–	–
EDSS, expanded disability status scale.					
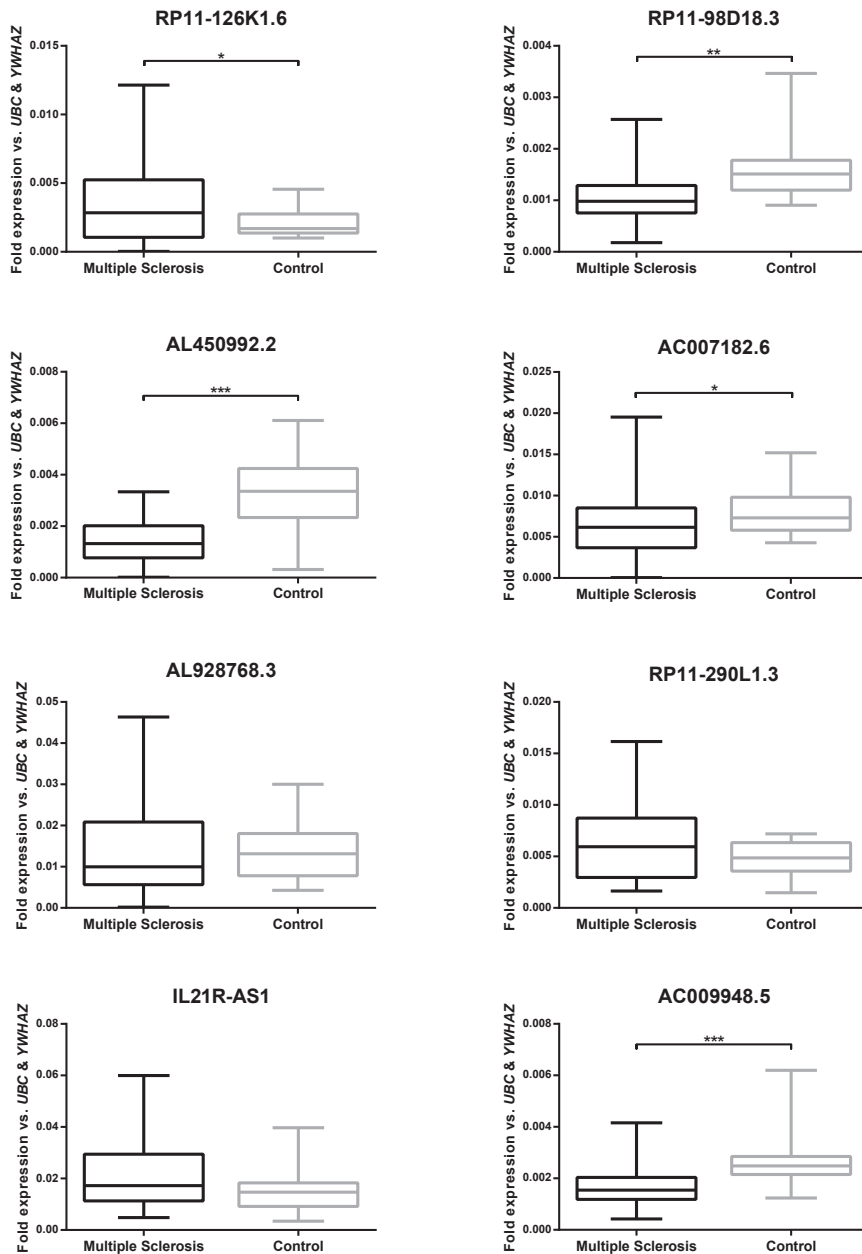
Figure 4. lncRNA Expression in MS compared to Controls

Boxplots of the expression levels of lncRNA genes in MS and control samples examined by real-time qPCR; values are given as mean normalized expression relative to *UBC* and *WYHAZ*; asterisks indicate significant differences to the controls (*p < 0.05; **p < 0.01; ***p < 0.001).

is important.[26] In the present study, we have identified that expression levels of AL450992.2, AC009948.5, and RP11-98D18.3 may potentially represent such PBMC biomarkers for MS among our studied lncRNAs. Moreover, the discriminatory power of these three lncRNAs was shown in the principal-component analysis (PCA) plot (Figure 8).

AL450992.2 and AC009948.5, as two potential MS biomarkers, are located adjacent to *RORC*, which is a key transcriptional regulator of Th17 lineage-specific lncRNAs and predominantly expressed in Th17 cells.[9,27] In accordance with the importance of *RORC* in the Th17 cell lineage, *RORC* correlation was assessed with the selected lncRNAs, revealing that AL450992.2 significantly correlated with *RORC* (Table S5). The possible role of AL450992.2 in different signaling pathways is determined through co-lncRNA and manifested through involvement with the cell cycle, T cell receptor signaling pathway, and many others (Figure 9).

MATERIALS AND METHODS

lncRNA Selection

As the first step to the preliminary selection of Th17 lncRNAs, the study by Spurlock et al.[9] on the whole-genome RNA-seq data of Th17 cell-lineage-specific lncRNAs was implemented. At the next step, the transcripts and chromosomal locations of each lncRNA were retrieved from Ensembl GRCh37 for further analyses to reach the defined list of potentially more effective lncRNAs based on the criteria and approach shown in the flowchart of the study in Figure 10, which is presented in detail.

lncRNAs Adjusted to Th17 Cell Differentiation Genes

Proteins and respective genes involved in Th17 cell differentiation were deduced through manual data mining,[28] in which keywords such as Th17 cell, differentiation, and signaling pathway were used to filter the outcomes. Systematic literature mining was performed on papers that were published until April 2017 from several databases, including PubMed, Science Direct, and Web of

RNA (lincRNA)-p21 on nuclear factor κB (NF-κB) activity in RA, PlncRNA1 mediated the function of the intestinal epithelial barrier in IBD, and NEAT1 as a regulator of inflammatory pathway in SLE.[22–24]

Arriving at more effective biomarker targets for AIDs seems a critical step, as Bernards declared that "a poor biomarker can be just as bad for the patient as a bad drug."[25] Improved clinical management of autoimmune diseases through biomarkers, which could be helpful for earlier and instant diagnosis, determining therapeutic strategies, and predicting outcome in this broad group of enervate disorders,
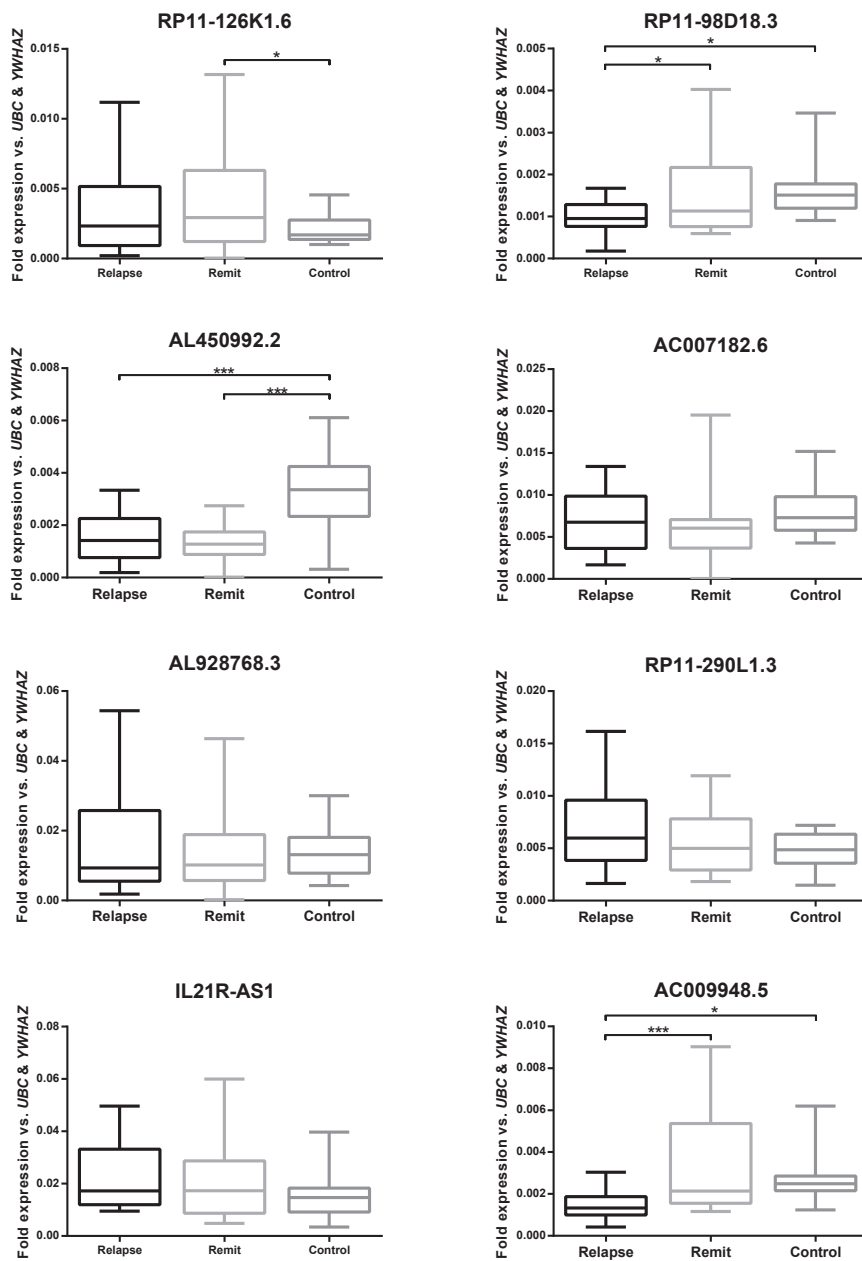
Figure 5. lncRNA Expression in Relapsing and Remitting Phases compared to the Controls

Boxplots of differential expression levels of lncRNA genes in relapsing-phase, remitting-phase, and control samples ($*p < 0.05$; $***p < 0.001$).

lncRNAs Adjusted to Differentially Expressed Genes in AIDs

To identify differentially expressed mRNAs in AID PBMCs, we used investigative analysis of array datasets that were retrieved from the Gene Expression Omnibus (GEO) database: for IBD, GEO: GSE3365 (Platform: GPL96; Affymetrix Human Genome U133A Array; 127 samples); for MS, GEO: GSE21942 (Platform: GPL570; Affymetrix Human Genome U133 Plus 2.0 Array; 29 samples); for RA, GEO: GSE15573 (Platform: GPL6102; Illumina Human-6 v2.0 expression beadchip; 33 samples), for SS, GEO: GSE84844 (Platform: GPL570; Affymetrix Human Genome U133 Plus 2.0 Array; 60 samples), for SLE, GEO: GSE10325 (Platform: GPL96; Affymetrix Human Genome U133A Array; 67 samples). These were utilized for identification of differentially expressed mRNAs with the GEO2R analyzer between patient samples and controls. We considered differentially expressed genes with adjusted p values < 0.05 and absolute log fold changes > 1. Next, each protein-coding gene's proximity to lncRNAs was checked via Python programming language in order to select adjusted lncRNAs (a similar code as that described in the immediately preceding section).

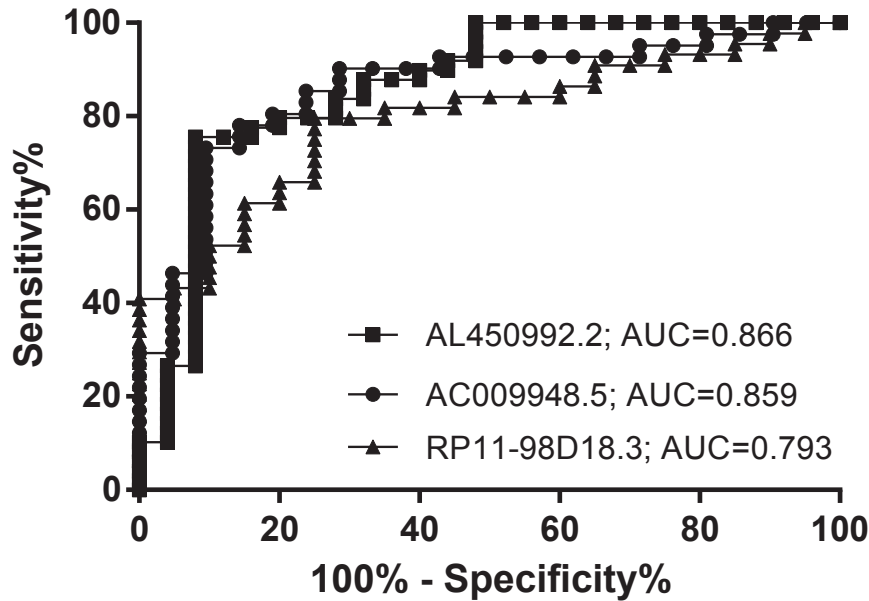### lncRNAs Adjusted to Autoimmune Diseases Associated with SNPs

SNPs ($p \leq 5 \times 10^{-8}$) associated with selected AIDs were extracted from the GWAS catalog[29] and were chromosomally localized in Ensembl GRCh37. The distance of each lncRNA locus from selected SNPs was assessed with Python programming language (v3.6.0) (Python Script S2). To do such analysis, SNPs relevant to each lncRNA categorized on the same chromosome and their positions were evaluated to reach the SNPs with less than 5-kb nt from selected lncRNAs.

lncRNA-Associated Diseases through miRNA Interactions

In order to predict lncRNA-associated diseases, a sequence-based bioinformatics tool, LncDisease, was used to identify potentially associated lncRNAs with autoimmune diseases.[14] TargetScan and

Science. Next, chromosomal locations of selected genes were determined in Ensembl GRCh37. Proximities of previously selected lncRNAs to these genes were analyzed using Python programming language (v3.6.0) (Python Script S1). In this step, Python programming language code categorizes the co-localized protein-coding genes and individual lncRNAs located in the same chromosome; then their distances were checked on each side of each lncRNA up to a distance of 50 kb away. Meanwhile, in order to consider more effective lncRNAs, we only selected lncRNAs that were co-expressed with protein-coding genes in Th17 cell differentiation.[9]

tial disease-associated lncRNAs through miRNAs interactions were listed.

**PBMC Isolation of Patients and Control**

In this step, we performed a case-control association study including 75 adults, divided into three groups. 10 mL peripheral blood was obtained from 25 patients recently diagnosed as having MS, according to the revised McDonald's criteria,[30] and 25 in the remitting phase, and both groups had no history of other diseases of the CNS, tumor(s) and systemic hematologic diseases, recent infection, and concomitant use of antineoplastic or immuno-

miRanda criteria were used to perform such analysis in this software. Each individual sequence of the lncRNA transcripts was input to LncDisease to predict the interacting miRNAs by using TargetScan and miRanda criteria. LncDisease then used the Human microRNA Disease Database (HMDD) for further analysis on the predicted miRNAs according to the TAM method. Finally, poten-

modulating therapies prior to blood sampling. Blood samples were gathered in EDTA-containing tubes from the Alzahra Hospital, Isfahan, Iran. Both groups were between 18 and 60 years old. Patients were also characterized according to the EDSS method of calculating and understanding disability in MS patients.[31] Moreover, 25 age- and gender-matched blood samples were collected from healthy
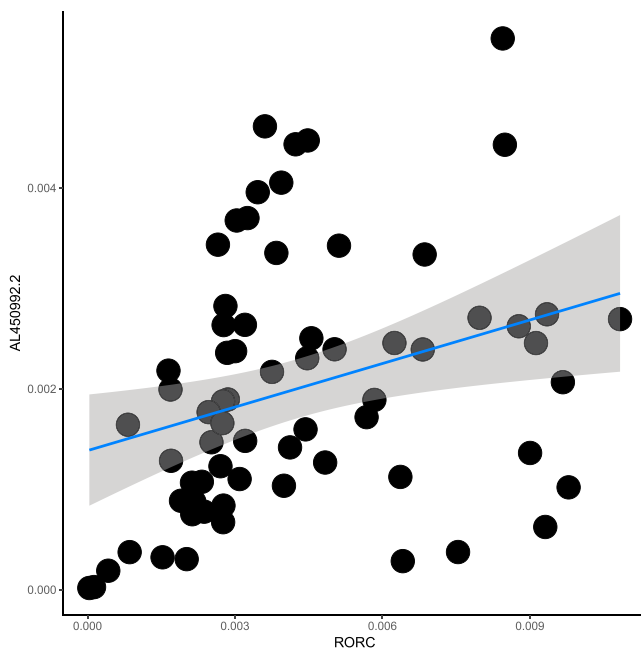

Figure 7. Correlation between AL450992.2 and *RORC* in MS Patients

Correlation between AL450992.2 relative expression fold change in PBMCs and *RORC* (Pearson's correlation = 0.31, p < 0.05).
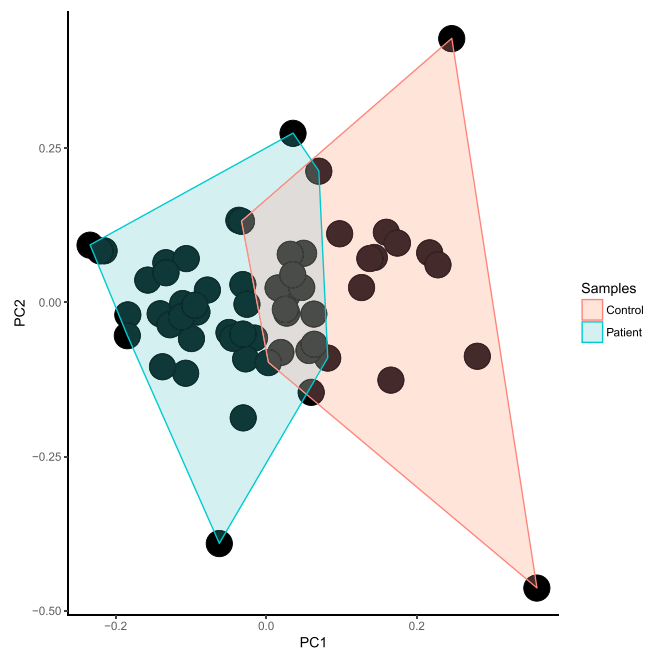
Figure 8. PCA

The scatterplot of principal-component analysis using biomarker lncRNAs in order to indicate a distinct overlap of patients and controls.

Significant Enriched Pathway Term (TOP 20)
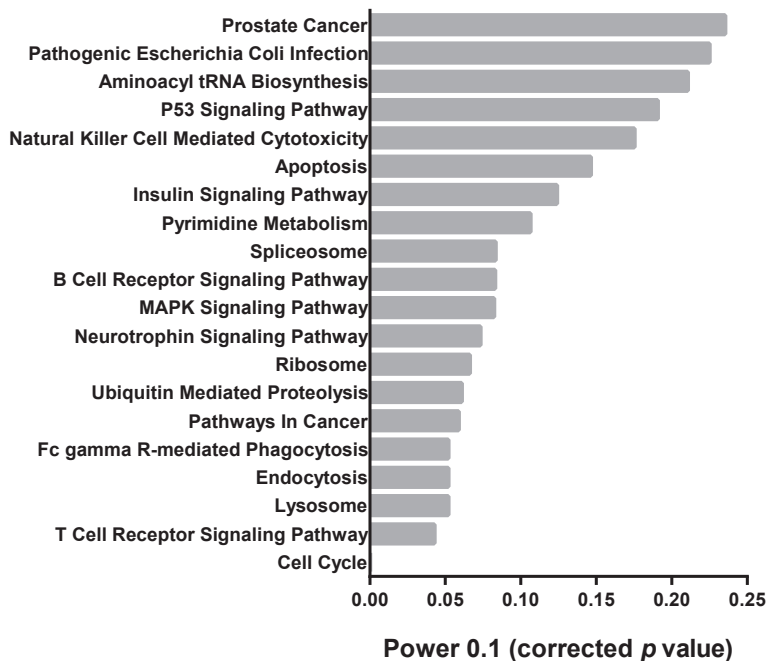
Figure 9. Enrichment Pathway Analysis

Top 20 pathways for the AL450992.2 lncRNA co-expressed genes based on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways analysis of co-lncRNA. The most significant column is shown in the bottom of the graph, with a lower p value.

PBMCs were isolated with the Lymphoprep density gradient medium (STEMCELL Technologies, Cambridge, MA, USA).

RNA Extraction and cDNA Synthesis

To isolate total RNA, TRIzol reagent (Invitrogen, Carlsbad, CA, USA) was used according to the manufacturer's supplied instructions. The RNA was quantified, and the quality was evaluated according to a ratio of absorbance of 260/280 nm using a Nanodrop spectrophotometer (Nanodrop 1000, Thermo Scientific, Waltham, MA, USA). To eliminate any potential contaminating DNA, total RNA samples were treated with RNase-free DNase (Thermo Scientific, Waltham, MA, USA) prior to real-time qPCR. cDNA was synthesized from the total RNA using a RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA, USA). Afterward, cDNA samples were stored at −80°C until further use.

volunteers, and written informed consent was obtained. The human subject protocol was approved by an institutional review board of Royan Institute (Project ID no. 91000573). All study protocols were carried out in accordance with the approved guidelines. Human
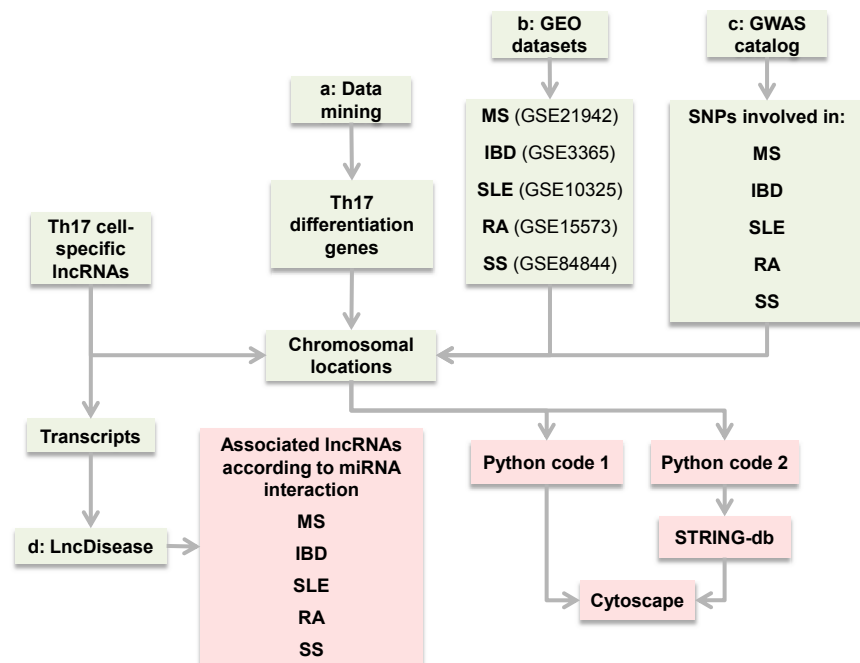
Figure 10. _In Silico_ Workflow of the Study

Lineage-specific lncRNAs were obtained from RNA-seq data. In step a, SNPs associated with each AID were retrieved from the GWAS catalog for analysis of their proximities to lncRNAs through Python programming language and were then visualized in Cytoscape. In step b, differentially expressed mRNAs in AID patients were obtained from corresponding GEO datasets. After chromosomal localization, their distances from selected lncRNAs were checked via Python programming language and, after inputting to STRING-db, visualized in Cytoscape. In step c, Th17 cell differentiation genes were selected through data mining. Their neighboring to the selected lncRNAs was determined and, after inputting to STRING-db, visualized in Cytoscape. In step d, lncRNAs' transcripts were deduced and inputted to the LncDisease software, which checked lncRNA-miRNA interaction and outputted the diseases related to the predicted miRNAs.

Table 4. Primer Pairs Used in This Study

Primer	Sequence (5′–3′)	GC%	3′ΔG (kcal/mol)	Amplicon Size (bp)	Ta (°C)
AC007182.6 F	GTAAGTCAGCATGTAGCAC	47.7	−7.5	86	50
AC007182.6 R	GCTCCATCAATAACATCTTCAT	36.4	−5.7		
AC009948.5.007 F	ATCAAGTTACAGAGCAGAG	4.1	−64	83	50
AC009948.5.007 R	AACATTACCGAGGACAAC	44.4	−6.5		
AL450992.2.001 F	TTAGACTCTCCTTGACCAT	42.1	−6.6	113	52
AL450992.2.001 R	TTCTCCTTCTGTGCTTTC	44.4	−5.7		
AL928768.3 F	ACAGGGAGGAAGTGTGGAG	57.9	−6.9	93	56
AL928768.3 R	GTGAGTAAGGGCGGGTC	64.7	−7.4		
IL21R-AS1 F	CTCCGACCACTCATTCAG	55.6	−6.1	119	56
IL21R-AS1 R	CTTATCACCTTGCCGTCTG	52.6	−6.6		
RP11.126K1.6 F	AACAGTCGCCTTCCAACAC	52.6	−6.5	118	62
RP11.126K1.6 R	TCCATCACTCCTACCCATCATT	45.5	−5.7		
RP11-98D18.3 F	AGGCTCAGTCACCTTTCC	55.6	−6.2	81	54
RP11-98D18.3 R	CCTTCTCTGTGACCGTCGA	57.9	−7.4		
RP11-290L1.3 F	GCGAGTGCGGCTCGTGATCTC	66.7	−5.8	168	64
RP11-290L1.3 R	CCGGTCAAGCTCAAGGAACTGC	59.1	−7.5		
RORC F	TGCCAGAATGACCAGATTGTGCTT	45.8	−7.1	132	60
RORC R	GAACAGCTCCATGCCACCGTA	57.1	−7.1		
UBC F	GGATTTGGGTCGCAGTTCTTG	52.4	−6.1	135	58
UBC R	TGCCTTGACATTCTCGATGG	50	−6.5		
YWHAZ F	ACTTTTGGTACATTGTGGCTTC	40.9	−6.9	94	62
YWHAZ R	CCGCCAGGACAAACCAGTA	57.9	−5.8		

F and R, forward and reverse primers, respectively; 3′ΔG stands for Gibbs energy for binding of the 3′ part of the primer with the template. Ta, annealing temperature of PCR reactions.

Real-Time qPCR

Transcription levels were measured in triplicate by real-time qPCR using SYBR Green Master Mix: SYBR Premix Ex TaqII (TaKaRa, Tokyo, Japan) and were carried out using specific primer pairs in the Step One Plus Real-Time PCR thermal cycler (Applied Biosystems, Foster City, CA, USA). *UBC* and *YWHAZ* were used as the most stably expressed reference genes, which were in accordance with gene expression analysis of PBMCs between healthy volunteers, relapsing-remitting MS (RRMS) and RRMS-interferon (IFN)-β patients.[32] Primer pairs used in these reactions are listed in Table 4.

Statistical Analysis

We performed analyses using a standard two-tailed Student's t test and a one-way ANOVA to assess differences, followed by pairwise comparisons using Tukey's correction. Pearson correlation was also performed for lncRNA-mRNA co-expression in this study. The discriminatory power of biomarker panels was assessed by ROC analysis between the controls and the MS patients. The aforementioned statistical analyses were performed using SPSS 20 software (SPSS, Chicago, IL, USA) and GraphPad Prism (v6; GraphPad software). For all analyses, $p < 0.05$ was considered statistically significant. Clustering of samples was performed using PCA with the package ggfortify (v0.4.1) in R software (v3.1.1).[33]

Network Construction

The protein-protein interactions of significant genes were assessed by STRING-db[34] and visualized by Cytoscape 3.6.0 software. Additionally, Cytoscape supplies a basic set of features for data integration, analysis, and visualization for complicated networks. Briefly, we transformed mined data into STRING-db to assess interactions and STRING-db results in Cytoscape. Next, to construct the network graph, we used methods available in Cytoscape to visualize and analyze the network by the number of direct edges. Co-LncRNA, a web-based computational tool that provides an overview of the relevant pathways of expressed protein-coding genes with inputted lncRNAs was also used in this study.[35]

Conclusions

In this study, most of the predicted lncRNAs of MS were derived from association with two criteria: Th17 cell differentiation genes and differentially expressed genes in AIDs. Most of the dysregulated lncRNAs were derived from these criteria. We propose that recurrently deregulated lncRNAs identified in this report could provide a valuable resource for studies aimed at delineating the relationship between functional lncRNAs and AIDs.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five tables and two Python scripts and can be found with this article online at https://doi.org/10.1016/j.omtn.2018.05.022.

AUTHOR CONTRIBUTIONS

S.T. and A.H. both participated in designing research studies, conducting experiments, acquiring data, analyzing data, providing reagents, and writing the manuscript; A.R. participated in designing research studies, conducting experiments, and acquiring data; K.G. participated in acquiring and analyzing data and writing the manuscript; E.G. carried out designing research studies, data interpretation, and manuscript writing; M.E. participated in acquiring data, data interpretation, manuscript writing, and giving final approval of the manuscript; M.H.N.-E. participated in designing research studies, data interpretation, manuscript writing, and giving final approval of the manuscript; and T.L.M. performed data interpretation, manuscript writing, and giving final approval of the manuscript.

CONFLICT OF INTERESTS

We declare no competing financial interests.

ACKNOWLEDGMENTS

REFERENCES

1. Cooper, G.S., Bynum, M.L., and Somers, E.C. (2009). Recent insights in the epidemiology of autoimmune diseases: improved prevalence estimates and understanding of clustering of diseases. J. Autoimmun. 33, 197–207.

2. Peters, A., Lee, Y., and Kuchroo, V.K. (2011). The many faces of Th17 cells. Curr. Opin. Immunol. 23, 702–706.

3. Dong, C. (2008). TH17 cells in development: an updated view of their molecular identity and genetic programming. Nat. Rev. Immunol. 8, 337–348.

4. Geisler, S., and Coller, J. (2013). RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat. Rev. Mol. Cell Biol. 14, 699–712.

5. Qi, P., and Du, X. (2013). The long non-coding RNAs, a new cancer diagnostic and therapeutic gold mine. Mod. Pathol. 26, 155–165.

6. Tong, Y.-K., and Lo, Y.M. (2006). Diagnostic developments involving cell-free (circulating) nucleic acids. Clin. Chim. Acta 363, 187–196.

7. Ayers, D. (2013). Long non-coding RNAs: novel emergent biomarkers for cancer diagnostics. J. Cancer Res. Treatment 1, 31–35.

8. Seim, I., Ma, S., and Gladyshev, V.N. (2016). Gene expression signatures of human cell and tissue longevity. NPJ Aging Mech. Dis. 2, 16014.

9. Spurlock, C.F., 3rd, Tossberg, J.T., Guo, Y., Collier, S.P., Crooke, P.S., 3rd, and Aune, T.M. (2015). Expression and functions of long noncoding RNAs during human T helper cell differentiation. Nat. Commun. 6, 6932.

10. ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74.

11. Ricaño-Ponce, I., and Wijmenga, C. (2013). Mapping of immune-mediated disease genes. Annu. Rev. Genomics Hum. Genet. 14, 325–353.

12. Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. USA 106, 9362–9367.

13. Chen, G., Wang, Z., Wang, D., Qiu, C., Liu, M., Chen, X., Zhang, Q., Yan, G., and Cui, Q. (2013). LncRNADisease: a database for long-non-coding RNA-associated diseases. Nucleic Acids Res. 41, D983–D986.

14. Wang, J., Ma, R., Ma, W., Chen, J., Yang, J., Xi, Y., and Cui, Q. (2016). LncDisease: a sequence based bioinformatics tool for predicting lncRNA-disease associations. Nucleic Acids Res. 44, e90.

15. Wapinski, O., and Chang, H.Y. (2011). Long noncoding RNAs and human disease. Trends Cell Biol. 21, 354–361.

16. Baron, J.L., Madri, J.A., Ruddle, N.H., Hashim, G., and Janeway, C.A., Jr. (1993). Surface expression of alpha 4 integrin by CD4 T cells is required for their entry into brain parenchyma. J. Exp. Med. 177, 57–68.

17. Ricaño-Ponce, I., Zhernakova, D.V., Deelen, P., Luo, O., Li, X., Isaacs, A., Karjalainen, J., Di Tommaso, J., Borek, Z.A., Zorro, M.M., et al.; BIOS Consortium; Lifelines Cohort Study (2016). Refined mapping of autoimmune disease associated genetic variants with gene expression suggests an important role for non-coding RNAs. J. Autoimmun. 68, 62–74.

18. Brunner, A.L., Beck, A.H., Edris, B., Sweeney, R.T., Zhu, S.X., Li, R., Montgomery, K., Varma, S., Gilks, T., Guo, X., et al. (2012). Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. Genome Biol. 13, R75.

19. Buzas, E.I., György, B., Nagy, G., Falus, A., and Gay, S. (2014). Emerging role of extracellular vesicles in inflammatory diseases. Nat. Rev. Rheumatol. 10, 356–364.

20. Sigdel, K.R., Cheng, A., Wang, Y., Duan, L., and Zhang, Y. (2015). The emerging functions of long noncoding RNA in immune cells: autoimmune diseases. J. Immunol. Res. 2015, 848790.

21. Wu, G.-C., Pan, H.-F., Leng, R.-X., Wang, D.-G., Li, X.-P., Li, X.-M., and Ye, D.Q. (2015). Emerging role of long noncoding RNAs in autoimmune diseases. Autoimmun. Rev. 14, 798–805.

22. Wang, J., Peng, H., Tian, J., Ma, J., Tang, X., Rui, K., Tian, X., Wang, Y., Chen, J., Lu, L., et al. (2016). Upregulation of long noncoding RNA TMEVPG1 enhances T helper type 1 cell response in patients with Sjögren syndrome. Immunol. Res. 64, 489–496.

23. Spurlock, C.F., 3rd, Tossberg, J.T., Matlock, B.K., Olsen, N.J., and Aune, T.M. (2014). Methotrexate inhibits NF-κB activity via long intergenic (noncoding) RNA-p21 induction. Arthritis Rheumatol. 66, 2947–2957.

24. Zhang, F., Wu, L., Qian, J., Qu, B., Xia, S., La, T., Wu, Y., Ma, J., Zeng, J., Guo, Q., et al. (2016). Identification of the long noncoding RNA NEAT1 as a novel inflammatory regulator acting through MAPK pathway in human lupus. J. Autoimmun. 75, 96–104.

25. Bernards, R. (2010). It's diagnostics, stupid. Cell 141, 13–17.

26. Gibson, D.S., Banha, J., Penque, D., Costa, L., Conrads, T.P., Cahill, D.J., O'Brien, J.K., and Rooney, M.E. (2010). Diagnostic and prognostic biomarker discovery strategies for autoimmune disorders. J. Proteomics 73, 1045–1060.

27. Hamzaoui, K., Borhani Haghighi, A., Ghorbel, I.B., and Houman, H. (2011). RORC and Foxp3 axis in cerebrospinal fluid of patients with neuro-Behçet's disease. J. Neuroimmunol. 233, 249–253.

28. Võsa, U., Vooder, T., Kolde, R., Vilo, J., Metspalu, A., and Annilo, T. (2013). Meta-analysis of microRNA expression in lung cancer. Int. J. Cancer 132, 2884–2893.

29. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L., and Parkinson, H. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 42, D1001–D1006.

30. Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J.A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., et al. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann. Neurol. 69, 292–302.

31. Kurtzke, J.F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology *33*, 1444–1452.

32. Oturai, D.B., Søndergaard, H.B., Börnsen, L., Sellebjerg, F., and Christensen, J.R. (2016). Identification of suitable reference genes for peripheral blood mononuclear cell subset studies in multiple sclerosis. Scand. J. Immunol. *83*, 72–80.

33. Tang, Y., Horikoshi, M., and Li, W. (2016). ggfortify: unified interface to visualize statistical results of popular R packages. R J. *8*, 478–489.

34. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. *43*, D447–D452.

35. Zhao, Z., Bai, J., Wu, A., Wang, Y., Zhang, J., Wang, Z., Li, Y., Xu, J., and Li, X. (2015). Co-lncRNA: investigating the lncRNA combinatorial effects in GO annotations and KEGG pathways based on human RNA-seq data. Database (Oxford) *2015*, bav082.

Supplemental Information

Integrative Analysis of lncRNAs in Th17 Cell

Lineage to Discover New Potential Biomarkers

and Therapeutic Targets in Autoimmune Diseases

Shohreh Teimuri, Aref Hosseini, Ahmad Rezaenasab, Kamran Ghaedi, Elahe Ghoveud, Masoud Etemadifar, Mohammad Hossein Nasr-Esfahani, and Timothy L. Megraw

Table S1: lncRNA transcripts

lncRNA	Number of transcripts	lncRNA	Number of transcripts	lncRNA	Number of transcripts
RP11-398A8.3	9	AL928768.3	1	RP11-304L19.3	1
AC002331.1	1	CTD-2008A1.1	1	RP11-390P2.4	4
AC007182.6	1	IL21R-AS1	1	RP11-488C13.5	2
AC009948.5	7	LINC00176	4	RP11-509E16.1	1
AC022087.1	1	RP11-126K1.6	1	RP11-783K16.5	2
AC074212.5	2	RP11-98D18.3	1	RP11-876N24.5	1
AC096579.7	2	RP11-290L1.3	1	RP11-213H15.3	3
AL450992.2	1	AC098614.2	2	AL353597.1	1
RP11-82L18.2	2	RP5-997D24.3	1		

Table S2: Th17 cell differentiation genes classification according to their pathway involvement

Signaling pathway involvements	Coding genes
Interleukin receptors	*TGFBR2, IL-23R, IL12RB1, IL6R, gp130, IL21R, IL1R1, IL1RAP, IL4R, IL12RB1, IL12RB2, IFNAR1, IFNGR2, IL27RA (WSX-1), IL2RA, IL2RB, IL2RG*
Chemokine receptors	*CCR2, CCR4, CCR5, CCR6, CCR7, CXCR3, CXCR6, ACKR2*
Activator transcription factors	*RORC, RORA, STAT3, SMAD2, SMAD3, SMAD4, IRF4, BATF, NFKBIZ, RUNX1, NFKB1, NFATC2, RELA, REL, HIF1A, IKZF3, AHR, NR4A1, TCF3, EP300, CHUK, CREM, PPARA, SOX5, PRDM1*
Inhibitory transcription factors	*FOXP3, TBX21, ETS1, GATA3, RARA, RARB, RARG, RXRA, PPARG, PPARD, NR1H3, NR1H2, VDR, IRF8, NR2F6, FOSL2, MAF, GFI1, SREBF1, KLF4, Tcf7, STAT1, STAT4, STAT6, STAT5A, STAT5B*
Inhibitory factors	*SOCS3, PIAS3, FOXO1, FOXO3, EOMES, TSC1, TSC2, AKT1S1, DEF6, SIGIRR, EGR2, DUSP2*
Activator factors	*HIF1A, ROCK2, SOCS1, SOCS2, AKT1, PDK1, RHEB, RPS6KB1, RPS6KB2, TRAF6, SGK1, IRAK1, RELA, CSNK2A1*
Th17 cell cytokine secretion profile	*IL17A, IL17F, IL22, IL26, IL21, IL6, IL9, IL10, CCL20, CCR9, TNF, IFNG, CSF2*

Table S3: Array expressed coding genes in AIDs

Disease	mRNA expression profile
IBD	*SERPINB2, TMEM158, AQP9, SNRNP70, HCAR3, IL1R1, CYP1B1, ABLIM3, CYLD, CXCL1, THBS1, MARK2, PHLDA1, SDC2, NPRL2, IL1R2, RUFY3, MYL9, XCL2///XCL1, CTTN, PF4V1, MYCBP2, ITGA2B, , KCNJ15, EWSR1, FOSL1, , RBBP6, FOSB, GK, ILF3, PTGS2, CXCL2, CAPN3, KLF6, TNPO3, , FCAR*, CLIC3*, THBD, ALAS2*, TMOD1, CXCL3, CCL2*, SLC6A8, HBQ1, CXCL8*, EIF5B, EPB42*, DUSP6, DDX3Y, DKC1, TREM1, SAT1, PLAUR, CXCL5, HIST1H3F, EGR3, S100A11, NRGN, PRKAR2B, BEX3, TTC3P1, KDM2A, ADGRE2, CCNL2, MFAP3L, KRI1, MYLIP, XK, ADGRG3, SAFB*, CELF1*, CELF2*, PF4, ZAP70, SUGP2, CCR1, GNG11, FPR1, ZFP36L2, MLH3, EGR2, SRSF10, SNRNP35, HMGXB4, GP1BB, FAXDC2, EPB41L3, MMP9*, KLRF1*, HIST1H2AH, VNN3, TGM2, IL1RN*, CD160*, CLU, SLC38A1, MPL, DDX18, SMOX, MRPL35, EGR1*, GZMK, TMEM259, TOP3B, FAM110B, CA2, KIF5B, HSP90AA1, FHL2, ACSBG1, THUMPD1, ZNF37BP*, FPR2*, FCF1, DDX39B*, AKR1C3*, PRRC2C*, UBXN7, SLC25A37*, ITGB3*, HIST1H2BH, CSNK2A1, EREG*, EIF5*, BRD3, ZMYND8, TTC38, CTDSPL, ZCCHC24, CRY2, GMPR, ANXA3*, NSUN5P1, ZNF12, ZNF428, ALPL, SART3, NSUN5P2, ZBTB20*, MED13L, DDX11, ZNF91, CD6, POGLUT1, UBE3A, ZNF551, KMT2A*, FFAR2*, TNFRSF25, GPRC5A, HIST1H2BO, ELMO2, GZMA, TFPI, ANKHD1, SNCA*, MPHOSPH9, SEC31B, PCTP, UBE4B*, CACNA2D3, HNRNPU, S100P*, TCF4*, CABP5, EZR, DLG5, SETD5, CCL7, TXNIP, KLRC2*, HSPH1, TNFRSF10C, PLEKHA1, CTSL, AHSP*, CXCR1, RPS6KB1, HNRNPH1, EMC1, SLC4A1*, NKTR, PI3, TRA2A, CYP4F3, TTF1, RNF144A, GUSBP9, GHRHR, CHMP6, ABCC3, AKAP13, PBXIP1, SMARCA4, NCR3, MME, TNS1, RBM10, CA1*, ZER1, LSM4, HERC4, IGHA2, TRMT1, HIST1H2BJ, SELENBP1*, CXCR2, ORM1*, DDX24, KIR2DS4, GTF3C1, ADRA1A, ANKZF1, PROSC, ARID4B, RNF24, DLX5, EMP1*, KIR2DS2, CAV2, VNN1, CH25H, MGAM, PCYOX1, PAK1IP1, HEY1, NMT1, CXorf57, MIA, ADAM8, MIS18BP1, TPM2, FKBP8, CFHR1, SLC6A6, MYOM2*, STEAP4, SGK3, NFIB, HS3ST1*, USP9Y, KDM5D, CHI3L1**
MS	*NEAT1, HINT3, CPA3, STT3A, MS4A3, STAT2, LTF*, HBA2///HBA1, ADAM9, HBB, SERPINB9, HSPA1B///HSPA1A, SNORD10, ALYREF, FBXO16///ZNF395, C21orf33, SLC25A37*, SNRNP40, G0S2*, MALAT1, CLEC7A, MDM4, FGD4, ARF6, TOMM22, POLK, DDX3X, MAT2A, TCN1*, ARPC4, PER1, STRBP, PLPP5, PRKAR1A, KCNE3, LILRA5, APLP2, CASP2, ACTB, HBM*, HBG2///HBG1, GATA2, FCAR*, HBD*, SLC8A1, GNAS, HIBADH, CD22, AZIN1, CLC*, SNX20, SLC4A1*, EIF2S3, MOB3A, CD200, FCRL1, SAMSN1, C1orf174, STAP1, HIPK3, SNORA28///EIF5*, CXCL8*, FAM129C, CD24*, IGHD, ICAM1, FAM30A, SELENBP1*, ALAS2*, BOD1L1, GM2A, TEX10, ATP6V1G2-DDX39B*, ODF3B, KLHL14, HLA-DOB, MS4A6A, TCL1A, VPREB3, IGHM, LAMP1, STRADB, RBM25, EBF1, EIF5A, FCRL2, HS3ST1*, NFKBIZ, ANKRD11, CA1*, NR4A2*, PTPRK, P2RX5, BANK1, KLF13, SMCHD1, DDX17, RSBN1L, STRAP., THRAP3, ATRX, TBL1XR1, FCRLA, PPP6R3, CEACAM8*, AHSP*, KIAA1033, CNTRL, AFF3, OSBPL10., DEFA1B*, QSER1, EGR1*, SNCA*, PRF1*, PTP4A1, SRRM2, UBA6, SCN3A, EPB42*, F5, SRGN, CD72, SF1, MS4A1, RNF213, RIF1, RIM58, CD79A, PCDH9, SYNE2, , TBL1X, BLK, FCRL5, , PAX5, NAMPT, CD19, LYZ, ASMTL-AS1, OSM, 2-Sep, CALR, FOXP1, 7-Sep, EPG5, GGNBP2, FAM126B, GPR155, CLINT1, ZBTB20*, RAPGEF6, KMT2A*, CHI3L1*, LONRF1, BCL11A, TGIF1, CCR3, RAB30, HDC*, ESF1, CAMP*, WHSC1L1, CELF1*, MLLT10, ZNF567, PNISR, , SWAP70, LUC7L3, RUSC1-AS1, SF3B6, KAT6A, IL7, CEP295, FBRSL1, GGA3, HNRNPA1, SGO2, PRRC2C*, KDM4B, USP9X, BRD7, TCF4*, ARG1*, NIPBL, SMC3, TSPAN13, RBM41, RASGRP3, SNORA21, ASH1L, B2M, MPHOSPH8, NPAT, PAWR, YTHDC1, CD3G, CCDC186, TNFRSF10A, CD69, PCBP2, SFXN1, TRAK2, NT5E, TMEM5, CSPP1, SCAF11, RORA, KRAS, PPP4R3B, C5orf24, ANKRD12, PPIG, CDCA7L, PLEKHA2, MYNN, AKAP12, EBLN3P, PALM2-AKAP2///AKAP2, ARHGEF7, RBM39, TOP1, ZFP36L1, MIRLET7DHG, NAA15, CR2, SKI, BACE2, MMP8*, TARDBP, GTF2A2, BTG1, SDCCAG3, YME1L1, GALNT1, EB1, CEP290, ABLIM1, SPPL2B, UFL1, SGSM2, ZNF37BP*, RNF103, CEP95, NFATC2IP, NEK1, BLNK, VPS13B, LUZP1, CHD9, CD79B, NCOA3, RASEF, TMCC1, ADAM28, GPATCH8, CYCS, STAG3, SPEN, PPP1R12A, SECISBP2L, FCER2, TCP11L2, AREG*, TSPAN5, KLHL6, RALGPS2, AP3D1, SSBP1, FMNL2, LOC102724611, PCNP, ASAP1, ATG16L2, LMO7, REL, KIAA0368, C19orf43, PHC2, CNTNAP2, SAMD9L, MEX3C, FLJ10038, PKIG, HEMGN, POU2AF1, CLIP1, SYTL3, KRT1, HIVEP3, ARHGAP18, LYST, IGHV3-23, NDUFS1, ZFAND5, SESN3, DEFA4*, HBP1, GYPA, ATAD2B, RPN1, COPG1, GLUD1, SLC16A1, PKN2, CAPZA2, DNAJC14, NONO, VDAC1, GLUD2, IDH3A, SAFB*, RBBP4, NDRG3, TP53, CYBB, ATF2, MPP7, G3BP1, PRKDC, CBFB, TMOD3, RTN3, DUSP3, LIX1L, SYNRG, , EXOSC6, DEDD, C5orf22, IFNGR1, PRIM2B///PRIM2, CRKL, MBNL1, YWHAZ, CDV3, CTSZ, CPM, STYX, GNS, ALKBH5, QKI,*

MRPS12, CRTAP, NFYC, SPTLC2, CORO1C, MTMR1, STX6, ZFX, SERTAD3, ST8SIA4, NADK, APH1A, ADAM17, MFAP3, WTH3DI///RAB6A, UBE2M, SS18, PFN1, ACTBP9, FPR2, UBE4B*, LRRC41, CELF2*, MAP3K2, YKT6, CDS2, WDFY2, FLI1, TSPAN14, TAPBP, SETSIP///SETP4///SET, FEM1C, LRRC25, RAB8B, LRRD1///CYP51A1, GSN, TSPAN4, TIRAP, FBN2, SORT1, MOCS3, FAM198B, IL1RN*, DCBLD2, C1orf228, NUTF2P4///NUTF2, PDGFC, YIPF4, , CEBPZOS, , ST6GAL1, CALU, HOMER3, FKBP5, MAX, NAIP, FAM49B, NUDT4, SMAD4, PLA2G4A, TMED2, DUSP10, SOWAHC, SIAH1, CHP1, ANKRD50, TAS2R45///TAS2R43, ANKRD10-IT1, UGGT1, CD1E, RBM14, CDKAL1, GALC, ODF2, ANKRD42, ARGLU1, KCNJ2, RHOB, , NRNPH1, MGEA5, TAF9B, CBX5, BMS1P5, CNOT6L, FOSL2, KANSL1, LPP, TYMP, CYBRD1, CDC42SE1, IL23A, IFI6*, ITGB3*, YWHAE, FPR3, CLIC3*, CD9, AMFR*

SLE	*FEZ1, RNASE2*, NCALD, GNLY, IFI27*, KLRF1*, PTGDR, ECRP, PRSS23, SYTL2, HDC8, SH2D1B, FCRL6, KLRB1, KLRG1, AKR1C3*, AUTS2, PPP2R2B, S1PR5, FGFBP2, FCER1A, KIR3DL1, KLRC3, HOPX, TCN2, OTOF, KIR2DL4, EOMES, PRF1*, KLRD1, IL7R, TGFBR3, IL18RAP, IFI44*, MATK, USP18*, CD160*, IL2RB*, IFI44L*, PLEKHF1, MCEMP1, RNASE1, EPSTI1*, PTGDS, ADGRG1, CLIC3*, ACSL1, HIST1H4H, KLRC1, IFITM3, GZMB, OAS1*, KLRC2*, MX1*, MMP9*, OLFM4, IFI6*, ADM, HP, OAS3*, HERC5*, CEACAM8*, RNASE3, RSAD2*, DYSF, CCL4L1, KIR2DL1, CEACAM1, ARG1*, CD24*, ELANE, ORM1*, IFIT1*, MPO, DEFA4*, CEACAM6, LCN2, DEFA1B*, PGLYRP1*, MMP8*, CLC*, CAMP*, S100P*, DEFA3, CTSG, ANXA3*, TACSTD2, DEFA1, BPI, GRN, AZU1, OLR1, LTF*, HBD*, CA1*, RETN, AHSP*, HBM*, TCN1*, IFIT1B, MYOM2*, ALAS2**
RA	*RNASE2*, S100A8, S100A9, TXN, ORM1*, IL2RB*, NDUFB3, CAMP*, LY96, ARG1*, PGLYRP1*, TMA7, COX7C, S100P*, MMP9*, PFDN5, RPL31, RPL39, COMMD6, RPS24, RPL26*
SS	*RSAD2*, EPSTI1*, IFI44L*, MX1*, IFIT1*, OAS2, HERC5*, IFI6*, IFIT3, OASL, OAS3*, MX2, SIGLEC1, PARP9, OAS1*, SPATS2L, EIF2AK2, USP18*, IFIT2, , IFI35, CMPK2, POBEC3A, IFI44*, LY6E, SERPING1, AREG*, IL1B, ANKRD22, FCGR1A, NR4A2*, IL8, TNFAIP6, CXCL9, HBEGF, MSR1, G0S2*, NOG, HIST1H3C, IGLV6-57, IFI27*, DDX60, LGALS3BP, LAMP3, PLSCR1, C2, HERC6, FFAR2*, CXCL11, CXCL10, EREG*, CCL2*, RGS1, MOP-1, SOCS3, NFIL3,*

Table S4: SNP associated with each autoimmune disease (AID)

Disease	SNP
IBD	Rs7547569, rs11209026, rs11581607, rs11742570, rs6880778, rs10748781, rs113653754, rs10781499, rs4409764, rs4077515, rs10761659, rs2188962, rs9836291, rs11236797, rs3024493*, rs56167332, rs2836883, rs2836878, rs75900472, rs3197999, rs10758669, rs35730213, rs4795397, rs56399423, rs6871626, rs17622378, rs3024505*, rs12994997, rs12946510, rs2066844, rs1801274*, rs2143178, rs11741861, rs10800309, rs7608910, rs6426833, rs2155219, rs7848647, rs6062496, rs2413583, rs7554511, rs6556412, rs4246905*, rs1505992, rs7134599, rs17800987, rs6927022, rs1297258, rs2823286, rs8127691, rs11564258, rs6752107, rs2847278, rs1893217*, rs12318183, rs13300483, rs3091316, rs7282490, rs9313808, rs11010067, rs34779708, rs6596, rs6062504, rs11743851, rs11614178, rs744166*, rs116855232, rs4946717, rs12942547, rs62037363, rs1388585, rs6935723, rs26528, rs6920220*, rs3749171, rs17293632, rs11554257, rs1819333, rs9889296, rs13407913, rs1420098, rs395157, rs917997, rs35164067, rs3091315, rs1250566, rs13300218, rs6568421, rs921720, rs72634258, rs17085007, rs2024092, rs11879191, rs12627970, rs1250546, rs4676410*, rs9868809, rs2179070, rs12568930, rs4845604, rs6017342, rs444210, rs55808324, rs2266959, rs11641016, rs80262450, rs529866, rs10956252, rs12720356, rs2488397, rs6545800, rs78487399, rs4380874, rs6586030, rs3764147, rs35675666, rs7097656, rs4246215, rs4768236, rs925255, rs11677953, rs2279990, rs1003342, rs148319899, rs1842076, rs1363907, rs17694108, rs6933404, rs1456896, rs4676408, rs2315008, rs4256159, rs2395022, rs1260326, rs11641184, rs941823, rs9557195, rs8005161, rs11230563, rs449454, rs2412970, rs3776414, rs12718244, rs6863411, rs1517352, rs9358372, rs6708413, rs1569723, rs3851228, rs516246, rs2328546, rs6740847, rs1734907, rs7657746, rs2266961, rs559928, rs9557207, rs12103, rs2488389, rs1292053, rs3742130, rs7746082, rs259964, rs3766606, rs78534766, rs28449958, rs4976646, rs1042058, rs2382817, rs1250550*, rs12722515, rs11229555, rs6740462, rs10495903, rs9297145, rs2153283, rs2050392, rs2297559, rs17119, rs2538470, rs17780256, rs16940202, rs12585310, rs913678, rs1049526, rs11548656, rs13001325, rs670523, rs11672983, rs4703855, rs11708026, rs6074022*, rs7495132, rs13126505, rs7240004, rs6716753, rs34920465, rs1847472,

rs423674, rs11187157, rs7556897, rs71559680, rs907611, rs2231884, rs194749, rs2688608, rs243505, rs5743289, rs10500264, rs4836519, rs1990760*, rs2816958, rs314313, rs13333062, rs2076756, rs13204742, rs6856616, rs6142618, rs2227564, rs10896794, rs3853824, rs254560, rs6651252, rs56116661, rs7404095, rs10521318*, rs4911259, rs1182188, rs3184504*, rs111456533, rs2412973, rs8049439, rs1991866, rs561722, rs17207042, rs11681525, rs3766920, rs653178*, rs7555082, rs1535, rs915286, rs1569328, rs2593855, rs11195128, rs4743820, rs727088, rs17061048, rs2284553, rs17656349, rs7773324, rs630923*, rs4656958, rs2274351, rs2790216, rs11168249, rs7517810, rs10051722, rs2395128, rs477515, rs11612508, rs2930047, rs38904, rs34856868, rs10142466, rs4692386, rs7236492, rs79206939, rs501916, rs59043219, rs144344067, rs4256018, rs6592362, rs2651244, rs2111485*, rs12654812, rs12199775, rs72810983, rs4802307, rs7011507, rs16953946, rs11221332, rs6478109, rs7911264, rs4899554, rs2472649, rs6025, rs4664304, rs6058869, rs2945412, rs7015630, rs76527535, rs503734, rs67289879, rs149169037, , rs113010081, rs11768365, rs5763767, rs3731257, rs11734570, rs17057051, rs11064881, rs3180018, rs483905, rs10065637, rs7758080, rs13277237, rs2073505, rs3798544, rs1479918, rs564349, rs212388, , rs10486483, rs10798069, rs2187668*, rs798502, rs1811711, rs2189234, rs7165170, rs16967103, rs3116494, rs4728142*, rs1077773, rs17229285, rs4821558, rs10797432, rs11150589, rs4243971, rs7438704, rs490608, rs13204048, rs6667605, rs10061469, rs1819658, rs11583043, rs11083840, rs200349593, rs727563, rs7973572, rs6677524, rs7954567, rs732072, rs2641348, rs11054935, rs1321366, rs17771967, rs1816854, rs80244186, rs9525625, rs11788518, rs4903214, , rs138788, rs4812833 |

| MS | rs3135388, rs3129889, rs9271366, rs11621145, rs3135338, rs2040406, rs9271640, rs7090512, rs2523393, rs6457617*, rs7200786*, rs3129871, rs1335532, rs4613763, rs3129934, rs10136766, rs669607, rs3957148, rs3828840, rs3129720, rs11810217, rs1738074, rs4648356, rs1800693, rs1077667, rs11154801, rs17066096, rs4902647, rs9260489, rs9282641, rs2546890, rs650258, rs2248359, rs3118470, rs703842, rs7595037, rs9275563, rs9657904, rs2119704, rs10201872, rs9891119, rs10492972, rs2300747, rs744166*, rs11581062, rs9272105, rs2425752, rs12466022, rs3817963, rs4939490, rs9292777, rs11129295, rs12368653, rs7522462, rs2293370, rs7238078, rs17445836, rs17824933, rs2303759, rs7923837, rs2283792, rs2019960, rs354033, rs802734, rs1250550*, rs533259, rs4410871, rs771767, rs10466829, rs13333054, rs17174870, rs874628, rs13192841, rs170934, rs2300603, rs6897932, rs140522, rs1323292, rs4285028, rs12722489, rs4961252, rs2150702, rs6718520, rs2293152, rs12212193, rs12048904, rs4308217, rs2744148, rs2104286*, rs180515, rs6074022*, rs882300, rs12025416, rs8070463, rs228614, rs6062314, rs11865121, rs6896969, rs2503875, rs4953911, rs10411936, rs2681424, rs7592330, rs1520333, rs630923*, rs9568281, rs733724, rs9283487, rs931555, rs1250542, rs806321, rs4075958, rs307896, rs2935183, rs794185, rs7191700, rs4409785*, rs756699, rs1790100, rs9596270, rs10866713, rs10936599, rs6659742, rs397020, rs17411949, rs9807334, rs290986, rs4271113, rs9320598, rs9523762, rs8049603, rs3780792, rs3745672, rs8112449, rs233100, rs7255066, rs13279485, rs7924357, rs2920001, rs1458175, rs1529316, rs1386330, rs12638253, rs1478091, rs6984045, rs8074980, rs2038256, rs281380, rs9321490, rs1062158, rs11026091, rs908821, rs1755289, rs17157903, rs2116078, rs6604026, rs3761959, rs4792814, rs11755724, rs793108*, rs17090640, rs6498169, rs1557351, rs299175, rs10259085, rs10518025, rs261902, rs12644284, rs12456021, rs6952809, rs17594362, rs386965, rs9821630, rs5978649, rs2842483, rs4149584, rs4669226, rs12513380, rs12047808, rs6941421, rs7191888, rs180358, rs10243024, rs2039485, rs4680534, rs17149161, rs7789940, rs651477, rs4704970, rs9480865, rs6917747, rs11666377, rs210428, rs13117816, rs2243123, rs1821625, rs4916321, rs10984447, rs7672826, rs1841770, rs1437898, rs1927457, rs716595, rs758944, rs7779014, rs11962089, rs17267338, rs17749211, rs1109670, rs337718, rs7253363, rs11957313, rs8007846, rs2602397 |

| RA | rs6910071, rs9268839, rs6457620, rs2476601*, rs3763309, rs660895, rs6457617*, rs13192471, rs12194148, rs2157337, rs6679677*, rs9296015, rs1571878, rs7752903, rs615672, rs3087243, rs77331626, rs12531711*, rs7765379, rs7731626, rs2051549*, rs71508903, rs12617656, rs2233424, rs5029924*, rs653178*, rs8026898, rs11889341*, rs2233434, rs3093024, rs2301888, rs9653442, rs8032939, rs4239702, rs10821944, rs8083786, rs10790268, rs874040, rs11933540, rs909685, rs12379034, rs7748270, rs34695944, rs5987194, rs34536443, rs9269234, rs11574914, rs1858037, rs41291794, rs706778, rs1854853, rs11676922, rs2841277, rs1980422, rs953387, rs3761847, rs7574865*, rs12525220, rs2736337, rs9603612*, rs2105325, rs6920220*, rs10892279, rs2867461, rs13330176, rs13017599, rs9603616, rs9275406, rs28411352, rs1893217*, rs1893592, rs2736340*, rs6859219, rs4409785*, rs864537, rs3093023, rs2561477, rs2283790*, rs1516971, rs2451258, rs1953126, rs10488631*, rs4452313, rs10985070, rs1950897, rs11580078, rs36001488, rs4625, rs7725052, rs10822050, rs17885785, rs72743477, rs117372389, rs2066363, rs7660520, rs7100025, rs1332099, rs773125, rs73013527, rs2395148, rs9378815, rs13277113*, rs755374, rs805297, rs2298428, rs657075, rs947474, rs17466626, rs951005, rs934734, rs3781913, rs9372120, rs624988, rs2469434, rs9826828, rs10499194, rs10175798, rs793108*, rs72634030, rs3783782, rs11089637, rs12140275, rs6715284, rs11741255, rs4810485, rs3824660, rs67250450, rs9373594, rs8133843, rs227163, rs2671692, rs13142500, rs2072438, rs72632736, rs2075876,

	rs2228145, rs1689510, rs12598357, rs2833522, rs4272, rs3125734, rs231735, rs77150043, rs12601925, rs1043099, rs678347, rs10774624, rs998731, rs6732565, rs3806624, rs726288, rs4780401, rs2596565, rs11203203, rs7579944*, rs11900673, rs4246905*, rs1250563, rs1502644, rs2807264, rs12519788, rs975730, rs1876518, rs2240335, rs12529514, rs2847297, rs968567, rs1877030, rs1943199, rs2812378, rs73194058, rs2664035, rs62324212, rs11145763, rs7258015*, rs284511, rs881375, rs12901682, rs26232, rs6479800, rs4305317, rs729302*, rs11984075, rs11121380, rs12565755, rs13315591, rs73081554, rs2582532, rs602662, rs2836882, rs9860428, rs9557321, rs12863738, rs1633360, rs12413578, rs1678542, rs2294369, rs2075184, rs4869313, rs36051895, rs3890745, rs4688011, rs7940423, rs12109285, rs16977065, rs7155603, rs2220327, rs331463, rs72717009, rs4921283, rs6568431*, rs7672495, rs7042370, rs6028945, rs6479891, rs1898036, rs4942242, rs17118552, rs1406428, rs11937061, rs800586, rs12570744, rs2317230, rs6689858, rs4676410*, rs114846446, rs6593803, rs10945919, rs1020388, rs12411988, rs6138892, rs1809529, rs1543922, rs1340317, rs9571178, rs9598783, rs1340319, rs3807306, rs12232497, rs1823549, rs1178121, rs11761231, rs1957895, rs16938910, rs12131057, rs7624766, rs6448432, rs34884278, rs62131887, rs7046653, rs1980493, rs7164176, rs72991, rs11148643, rs9271348, rs757278, rs10956445, rs6956740, rs7831697, rs12517545, rs6500395, rs6026990, rs10520789, rs12652364, rs12928404, rs75908454, rs79575701, rs12719740, rs9604529, rs1914816, rs13119723, rs10988542, rs284515, rs1772408, rs4937362, rs1329568, rs16906916, rs13031237, rs1679568, rs6774280, rs1273516, rs13137105, rs2872507, rs13385025, rs2738774, rs743777, rs12046117, rs6496667, rs11051970, rs4336372, rs2104286*, rs3790022, rs4395908, rs10128264, rs11225055, rs55705316, rs11839053, rs13015080, rs114940806, rs1279094, rs864089, rs481331, rs2837960, rs854555, rs2062583, rs3783637, rs2280381, rs11045392, rs7141276, rs6427528, rs840016, rs2230926*, rs10865035, rs4750316, rs17374222, rs518167, rs7839040, rs1079467, rs8097070, rs1527934, rs78507369, rs12445022, rs9299346, rs6138150, rs11870477, rs508970, rs10113213, rs61996546, rs112165031, rs6853094, rs73401585, rs9595973, rs9633402, rs62438583, rs13393173, rs437943, rs42041, rs10876993, rs7404928, rs1539909, rs3794271, rs45475795, rs11113818, rs2961663, rs41005, rs62359376, rs1501138, rs2741200, rs12422918, rs983332, rs1600249, rs66861122, rs6506122, rs7712113, rs1885747, rs111580313, rs2002842, rs12831974, rs885814, rs3184504*, rs3218251, rs60624478, rs7800668, rs1498103, rs6506569, rs2802369, rs255758, rs6448119, rs1624005, rs960902, rs923880, rs3816587, rs17679567, rs10084630, rs16973500
SS	rs117026326, rs9271573, rs9271588, rs4282438, rs10168266., rs3823536, rs11889341*, rs5029939*, rs7574865*, rs17074492, rs7341475, rs9277554, rs11048434, rs1957173, rs7999279, rs9938751, rs16837677, rs7192380, rs181851, rs79407237, rs4842091, rs10474500
SLE	rs1270942, rs11889341*, rs10488631*, rs1150757, rs17849501, rs34572943, rs35000415, rs35472514, rs3131379, rs3757387, rs1143679, rs10036748, rs7574865*, rs2476601*, rs2205960, rs6932056, rs12539741, rs13239597, rs1150754, rs2187668*, rs2431697, rs6590330, rs7812879, rs6889239, rs9270984, rs12531711*, rs2732549, rs4728142*, rs9888739, rs4917014, rs7444, rs2736332, rs2051549*, rs558702, rs3794060, rs2736340*, rs5029924*, rs12612769, rs704840, rs58721818, rs12537284, rs10912578, rs13332649, rs2431098, rs2230926*, rs11644034, rs1887428, rs10845606, rs10028805, rs6705628, rs9652601, rs6736175, rs4639966, rs131654, rs11185602, rs17321999, rs6804441, rs7726414, rs9275572, rs13385731, rs6679677*, rs1734787, rs6762714, rs2289583, rs11073328, rs9311676, rs6568431*, rs4852324, rs1059702, rs73135369, rs3768792, rs1059312, rs11085727, rs10911628, rs2254546, rs9603612*, rs387619, rs3747093, rs76413021, rs2304256, rs58688157, rs9273076, rs9271100, rs3734266, rs1801274*, rs6671847, rs74290525, rs139110493, rs2301271, rs34889541, rs5029939*, rs17603856, rs113478424, rs597325, rs548234, rs34330, rs10753074, rs1913517, rs1128334, rs7097397, rs4622329, rs11697848, rs2111485*, rs2297550, rs1385374, rs2283790*, rs7708392, rs11574637, rs7601754, rs2176082, rs4948496, rs77583790, rs12736195, rs11235667, rs3821236, rs2736345, rs849142, rs1061502, rs13277113*, rs2275247, rs11860650, rs7941765, rs268134, rs9462027, rs10048743, rs4963128, rs12711490, rs10516487, rs5986948, rs4388254, rs887369, rs4902562, rs597808, rs12802200, rs729302*, rs9782955, rs7579944*, rs10499197, rs564799, rs2286672, rs3129716, rs10774625, rs3024505*, rs2663052, rs1150753, rs143123127, rs494003, rs6445975, rs340630, rs10946940, rs877819, rs2941509, rs11059919, rs4916342, rs7329174, rs8023715, rs820077, rs4917385, rs2618476, rs1635852, rs1167796, rs960709, rs12822507, rs7200786*, rs1170426, rs7197475, rs7172677, rs7897633, rs7258015*, rs12531540, rs13306575, rs6697139, rs1990760*, rs2222631, rs3024493*, rs2267828, rs2532871, rs9267531, rs1267499, rs7090925, rs10798269, rs4684256, rs1049564, rs4921283, rs9398235, rs1878186, rs10498070, rs1133906, rs10466455, rs2785197, rs1028488, rs4660116, rs10995092, rs7186852, rs8012283, rs2647012, rs1535001, rs1874791, rs7582694, rs6914831, rs11185603, rs643955, rs3827644, rs1364989, rs274068, rs12599402, rs525410, rs912784, rs12753920, rs11057864, rs11655550, rs7773456, rs11231824, rs1048257, rs7692514, rs1108131, rs4956211, rs7834765, rs601162, rs2303745, rs4850410, rs10032909, rs2928402, rs958476, rs11150610, rs5754217, rs918959, rs956237, rs4544377, rs6946131, rs6084875, rs752010, rs6538678, rs9267972, rs4901847, rs967616, rs6850606, rs11908000, rs61732491, rs181502228, rs8009420, rs6696619, rs12589674, rs710987, rs11101442, rs11243676, rs3130320, rs10857712, rs10857712, rs10041935, rs2252996,

rs7411387, rs3934007, rs959260, rs79404002, rs190029011, rs7892586, rs7098187, rs979233, rs6695567, rs4963128, rs1355223, rs12993006, rs742108, rs2288012, rs10521318*, rs1534618, rs2669010, rs12565776, rs2786111, rs2478118, rs2236178, rs16860537, rs4522865, rs1429411, rs9402743, rs2797780, rs8105429, rs427221, rs10992211, rs17083844, rs10911390, rs10276619, rs10954650, rs12141391, rs10737562, rs7927370, rs2764208, rs9303277, rs11717455, rs12917712, rs9937837, rs3862260, rs4574684, rs330048, rs17807624, rs7031325, rs931127, rs10142203, rs4149228, rs2313132, rs12949531, rs1463525, rs6049839, rs2396545, rs1478897, rs10254284, rs4641121, rs17301013, rs17039212, rs9263871, rs12529935

Table S5: *RORC* and lncRNAs correlation

	RORC	
	Pearson Correlation	Sig. (2-tailed)
AC009948.5.007	0.139	0.265
RP11.290L1.3	0.153	0.219
AL928768.3	0.238	0.051
RP11.98D18.3	0.106	0.404
IL21R.AS1	0.028	0.824
AC007182.6	0.138	0.267
AL450992.2	0.306	0.010
RP11.126K1.6	0.239	0.051

File S1: Python's scripts

```python
gene=open("C:\\Users\\Aref\\Desktop\\newgenes.txt","r")

lnc=open("C:\\Users\\Aref\\Desktop\\lnc.txt","r")

distance=50000

res=open("C:\\Users\\Aref\\Desktop\\result.txt","w")

for g in range(1,117):

    gline=gene.readline()

    import string

    gcht1=str.find(gline,"\t") # first TAB character

    gcrn=int(gline[:gcht1]) #choromosomal number

    gcht2=str.find(gline,"\t",gcht1+1) # second TAB character

    gfn=int(gline[gcht1+1:gcht2])#first nucleotide in gline

    gln=int(gline[gcht2+1:])#last nucleotide in gline

    lnc=open("C:\\Users\\Aref\\Desktop\\lnc.txt","r")

    for l in range(1,27):

        lline=lnc.readline()

        import string

        lcht1=str.find(lline,"\t") # first TAB character

        lcrn=int(lline[:lcht1]) #choromosomal number

        lcht2=str.find(lline,"\t",lcht1+1) # second TAB character

        lfn=int(lline[lcht1+1:lcht2])#first nucleotide in lline

        lln=int(lline[lcht2+1:])#last nucleotide in lline

        if gcrn==lcrn:

            if abs(gfn-lln)<=distance or abs(lfn-gln)<=distance:

                res.write(("%-2d\t%-2d\n")%(g,l))

            if gfn>lfn and gfn<lln:

                res.write(("%-2d\t%-2d\n")%(g,l))

            if gln>lfn and gln<lln:

                res.write(("%-2d\t%-2d\n")%(g,l))

    res.write("\n")

res.close()
```

```
gene.close()

lnc.close()
```

File S2: Python's script

```
snp=open("C:\\Users\\Aref\\Desktop\\SLE SNP.txt","r")

lnc=open("C:\\Users\\Aref\\Desktop\\lnc.txt","r")

distance=5000

SNPres=open("C:\\Users\\Aref\\Desktop\\SNPresult.txt","w")

for s in range(1,313):

    sline=snp.readline()

    import string

    scht1=str.find(sline,"\t") # first TAB character

    scrn=int(sline[:scht1]) #choromosomal number

    sn=int(sline[scht1+1:])#SNP Nuclotide Position

    lnc=open("C:\\Users\\Aref\\Desktop\\lnc.txt","r")

    for l in range(1,27):

        lline=lnc.readline()

        import string

        lcht1=str.find(lline,"\t") # first TAB character

        lcrn=int(lline[:lcht1]) #choromosomal number

        lcht2=str.find(lline,"\t",lcht1+1) # second TAB character

        lfn=int(lline[lcht1+1:lcht2])#first nucleotide in lline

        lln=int(lline[lcht2+1:])#last nucleotide in lline

        if scrn==lcrn:

            if abs(sn-lln)<=distance or abs(sn-lfn)<=distance:

                SNPres.write(("%-2d\t%-2d\n")%(s,l))

            if sn>lfn and sn<lln:

                SNPres.write(("%-2d\t%-2d\n")%(s,l))

    SNPres.write("\n")

SNPres.close()

snp.close()

lnc.close()
```