HHS Public Access

Author manuscript

*Arthritis Rheumatol*. Author manuscript; available in PMC 2017 November 01.

Published in final edited form as:

*Arthritis Rheumatol*. 2016 November ; 68(11): 2637–2645. doi:10.1002/art.39746.

# Limb Bud and Heart Development Gene Transcription is Regulated by the Interplay of an Enhancer Risk Allele and DNA Methylation in Rheumatoid Arthritis

**Deepa Hammaker, PhD**[1], **John W. Whitaker, PhD**[2,#a], **Keisuke Maeshima, MD, PhD**[1], **David L. Boyle**[1], **Anna-Karin H. Ekwall, MD, PhD**[3], **Wei Wang, PhD**[2], and **Gary S. Firestein, MD**[1]

[1] School of Medicine, Department of Rheumatology, Allergy and Immunology, University of California San Diego, La Jolla, California, USA

[2] Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California, USA

[3] Department of Rheumatology and Inflammation Research, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

## Abstract

**Objective—**To identify non-obvious therapeutic targets for rheumatoid arthritis, we performed an integrative analysis incorporating multiple ‘omics data and the ENCODE database for potential regulatory regions. This analysis identified the Limb Bud and Heart development (LBH) gene, which has risk alleles associated with RA/celiac disease and lupus, and can regulate cell proliferation in RA. We identified a novel LBH transcriptional enhancer with an RA-risk allele (rs906868 G (risk, Ref) /T) 6kb upstream of the LBH gene with a differentially methylated locus. The confluence of three regulatory elements, rs906868, an RA differentially methylated locus and a putative enhancer, led us to investigate their effect on LBH regulation in fibroblast-like synoviocytes (FLS).

**Methods—**We cloned the 1.4kb putative enhancer with either the rs906868 Ref allele or SNP variant into reporter constructs. The constructs were methylated in vitro and transfected into cultured FLS by nucleofection.

**Results—**We found that both variants increased transcription, thereby confirming the region’s enhancer function. Unexpectedly, the transcriptional activity of the Ref risk allele was significantly lower than the SNP variant and is consistent with low LBH levels as a risk factor for aggressive FLS behavior. Using RA FLS lines with homozygous Ref or SNP allele, we confirmed that homozygous Ref lines expressed lower LBH mRNA levels than the SNP lines. Methylation significantly reduced enhancer activity for both alleles, indicating that enhancer function is dependent on its methylation status.

Correspondence and reprint requests to: Dr. Gary S. Firestein, MD, UC San Diego School of Medicine, 9500 Gilman Dr. #0656, La Jolla, CA 92093-0656, transmed@ucsd.edu, Phone: (858) 534-2359, Fax: (858) 534-2606.

[#a]Current Address: Discovery Sciences, Janssen R&D of Johnson & Johnson, San Diego, California, USA

Research supported by Arthritis Foundation, Rheumatology Research Foundation (Disease Targeted Research), NIH (NIAMS) R01 AR065466 (GSF) and The Swedish Society of Medicine (AKE)

The authors have no financial conflicts of interest to disclose

**Conclusion—**This study shows how the interplay between genetics and epigenetics can affect expression of LBH in rheumatoid arthritis.

Rheumatoid arthritis (RA) is a chronic, immune-mediated disease characterized by synovial inflammation and joint damage(1-3). Fibroblast-like synoviocytes (FLS) are key players in rheumatoid synovial pathology and cartilage destruction through the elaboration of cytokines, small molecules, and proteases(4). These cells form the synovial intimal lining and assume a unique tumor-like phenotype, including the ability to invade into cartilage explants and to migrate between sites in a SCID mouse model(5, 6). However, therapies that target FLS are not currently available. With the advent of new unbiased genome-wide platforms, RA FLS are now amenable to genomic and epigenomic analyses that can potentially identify non-obvious therapeutic targets.

To understand the aggressive behavior of RA FLS, we previously analyzed DNA methylation patterns and identified an RA-associated epigenetic signature that implicated genes involved with cell migration, cell recruitment and matrix regulation(7). We then performed an integrative analysis of three genome-wide datasets, namely, RA risk-associated genetic variants (GWAS), changes in DNA methylation at gene promoters, and changes in gene expression in FLS(8). This analysis identified a limited subset of genes, where an overlap was observed in the three datasets(9). One of these genes, *LBH* (*Limb Bud and Heart Development* gene), is a transcription cofactor that regulates cell differentiation during development and is linked to cardiac and skeletal abnormalities in the partial trisomy 2p syndrome(10). We subsequently showed that *LBH* regulates synoviocyte proliferation by arresting cell growth at the G1 checkpoint(11, 12). Reduced LBH expression has been associated with inflammation and autoimmunity(12, 13). We hypothesized that low *LBH* expression could, therefore, be an RA risk factor by permitting the G1-S phase transition. Several *LBH* polymorphisms have been associated with autoimmune diseases including RA, celiac disease and systemic lupus erythematosus (SLE), which suggests a key role in regulating immune function(14, 15).

In this study, we expanded our DNA methylation analysis beyond promoters to consider distal regulatory loci by including genome-wide DNaseI hypersensitive sites (DHS). These open chromatin regions often function as regulatory elements such as transcriptional enhancers that interact with promoters through DNA looping(16). We identified a new putative *LBH* enhancer 6kb upstream of the transcription start site that contains a differentially methylated locus (DML) in RA FLS, but unexpectedly, also contains an adjacent RA-associated variant. This striking confluence of an RA risk-associated variant, an RA DML and an open DNA regulatory element suggests that *LBH* and this regulatory region might be important in RA pathogenesis. Here, we show that the open chromatin region is, in fact, an enhancer and that the RA-associated SNP decreases *LBH* gene transcription. The data show how a functional SNP and DNA methylation interact to alter expression of a gene that regulates the aggressive behavior of pathogenic cells in RA.

## Materials and Methods

### Human fibroblast-like synoviocytes and culture conditions

This study was approved by the Institutional Review Board of University of California, San Diego School of Medicine, and informed consent was obtained from all participants. Synovial tissue was obtained from patients with osteoarthritis (OA) and RA at the time of total joint replacement or synovectomy, as previously described. The diagnosis of RA conformed to American College of Rheumatology 1987 revised criteria(17). The synovium was minced and incubated with 0.5mg/ml collagenase type VIII (Sigma-Aldrich, St. Louis, MO) in serum-free RPMI 1640 (Life Technologies, Grand Island, NY) for 1h at 37°C, filtered, extensively washed, and cultured in DMEM (Life Technologies, Grand Island, NY) supplemented with 10% heat-inactivated fetal bovine serum (Gemini Bio Products, Calabasas, CA), and supplements (penicillin, streptomycin, gentamicin, and glutamine) in a humidified 5% CO2 atmosphere. Cells were allowed to adhere overnight, non-adherent cells were removed. Adherent FLS were split at 1:3 when 70–80% confluent and used from passages 4 through 7(18).

### Quantitative Real time PCR

RA FLS were serum starved for 24h in DMEM containing 0.1% FCS and supplements. Total RNA was isolated from FLS using RNASTAT-60 and reverse transcribed (Applied Biosystems, Foster City, CA). The cDNA served as template for amplification by qPCR using LBH TaqMan Gene Expression assays (Hs00368853_m1; StepOnePlus Instruments, Applied Biosystems, Foster City, CA). The Ct values were normalized to GAPDH (Hs02758991_g1)(19).

### Processing omics datasets and pathway analysis

Data processing for integrative analysis was performed as described previously(9). The putative enhancer regions were based on the ENCODE database of 1,281,988 DNaseI-seq peaks from 125 cell-types/condition combinations, including fibroblast cell lines. These were filtered to remove any peaks that overlap promoters (−2500bps to +500bps from a gene transcription start site (TSS)), resulting in a set of 1,055,570 putative enhancers. Promoter regions were identified from GENCODE gene models and represent 177,999 unique TSS. Putative enhancers containing DMLs were identified and then linked to potential target regulatory genes by identifying the closest TSS. KEGG pathway enrichment analysis of the double evidence gene sets was performed as described in our previous studies(8).

### Chromatin Immunoprecipitation (ChIP)-qPCR

ChIP assay was performed using Zymo-Spin ChIP kit according to the manufacturer's protocol (Zymo Research, Irvine, CA). Briefly, FLS were serum-starved for 24h with DMEM/0.1%FCS with penicillin, streptomycin, gentamicin and glutamine). The cells were fixed in 1% formaldehyde (ThermoScientific, Waltham, MA) for 8 minutes at RT. After sonication, the chromatin was immunoprecipitated with specific antibodies overnight at 4°C. The antibodies used were anti-H3K4me1 (ab8895, Abcam, Cambridge, MA), anti-RNA polymerase II (ab26721) or rabbit IgG (2729, Cell Signaling Technology, Danvers, MA).

The complexes were then immunoprecipitated with protein A magnetic beads for 1h at 4°C, and reverse-crosslinked at 65°C overnight. The eluted DNA was purified and used as template in the qPCR reaction. A 10% input sample was used as control. The enhancer primers (EpiTect ChIP qPCR primer GPH1007377(-)07A) were purchased from Qiagen (Valencia, CA). Control GAPDH promoter primers were purchased from Clontech (Mountain View, CA).

### Identifying RA FLS lines with homozygous Ref and SNP alleles

Purified genomic DNA from RA FLS was subjected to PCR and agarose gel analysis. Tetra-arms PCR primers were designed using BatchPrimer3 program(20). The outer primers were forward 5'-CACTTCAACAGCCAGAAAAGAGAGGTTA-3' and reverse 5'-GCAAGTGCTAGCTCAGGTTACAGTTAGG-3'. The inner primers were forward 5'-ACGTCTCCCACTATTTTACCCACGAC-3' and reverse 5'-ATGATGGAGTTTCCCTGGTCCAAAT-3'. Pyrosequencing of Ref and SNP FLS at chr2:30448529 (cg 20495819) was performed by EpigenDx (Hopkinton, MA).

### Constructs, transfections and luciferase reporter assays

A 1311bp enhancer fragment was amplified from homozygous Ref or SNP genomic DNA with specific primers (Forward – 5'-CTGAGCTCGCTAGCCTCGAGCAGGCTGATCTCGAACTC-3' and Reverse - 5'-ATACCCTCTAGTGTCTAAGCTTAGGCCTCAGGGTCTAGTAAG-3'). The enhancer fragment was cloned into the pGL4.23 vector with a minimal promoter and the empty vector was used as control. The constructs were validated by sequencing. Control, ref and SNP plasmids were methylated by incubating 50ug of plasmid with 20U/ul of CpG methyltransferase (M.SssI) and 32 mM S-adenosyl methionine, overnight at 37°C, according to manufacturer's instructions (New England Biolabs, MA). Complete methylation was verified by plasmid bisulfite modification and pyrosequencing (EpigenDx, Hopkinton, MA). For luciferase assays, 2ug of control, homozygous Ref or SNP plasmids were co-transfected with 50ng of renilla plasmid into $5\times10^5$ FLS by nucleofection (Lonza, Walkersville, MD). The cells were serum starved for 24h prior to harvesting. The cells were lysed and assayed for luciferase activity using the dual luciferase reporter assay kit (Promega, Madison, WI). The firefly luciferase activity was normalized to renilla and expressed relative to empty control vector.

## Results

### Expanding RA FLS integrative analysis to include putative transcription enhancers

Our previous integrative analysis of RA FLS focused on promoter DMLs and identified several candidate genes of interest in RA therapy, namely *ACOXL*, *AIRE, CASP8, CSF2, ELMO1, ETS1, HLA-DQA1* and *LBH*(9). To increase the coverage, we expanded the number of DMLs to include non-promoter regulatory regions. Putative enhancers containing DMLs were identified as described in Material and Methods and then linked to potential target regulatory genes by identifying the closest transcription start site. Integrative analysis of these newly identified differentially methylated genes (DMG), differentially expressed genes (DEGs), and RA GWAS variants, yielded 672 genes that were significantly different

between RA and non-RA in at least 2 of the 3 datasets. The 672 genes represent 614, 28, and 18 double evidence genes and 12 triple evidence genes (Fig 1A, Table S1). The 12 genes included the 8 genes found previously and the new genes, *IKZF3, IL2RB, PRKCH, PTPN11*, and *LBH*. Interestingly, *LBH* was identified twice, once for the promoter DML and the second, for a putative enhancer DML.

## Pathway analysis of the expanded multi-evidence gene set

The number of double evidence genes (i.e. identified in at least two of the three databases) and triple evidence genes increased from 357 genes found in our previous study(9) to 672 genes when the enhancer-related genes were included (multi-evidence genes (MEGs); Fig 1B; Table S2). KEGG pathway analysis was performed, and the 'Immune' and 'Signaling' groups were enriched with at least four pathways relevant to RA. The 'Cell adhesion molecules' pathway was enriched 3.9-fold ($P$=2.71E-07, $q$=3.75E-05) with 19 out of 129 genes in the MEGs. The 'Cytokine-cytokine receptor interaction' pathway was enriched 2.88-fold ($P$=4.8E-07, $q$=3.75E-05) with 27 out of 248 MEGs and the 'JAK-STAT signaling pathway' was enriched 2.85-fold ($P$=2.1E0-4, $q$=5E-03) with 15 out of 139 MEGs. The KEGG 'Antigen processing and presentation' pathway was enriched 3.55-fold ($P$= 8.4E-04, $q$=1.1E-02) with 9 out of 67 genes. Interestingly, the fold enrichment of these four pathways with 672 MEGs was similar to our previous data using 357 MEGs. Therefore, the expanded integrative analysis confirms and extends the association of immune-related pathways and abnormally methylated genes in RA FLS.

## Identification of a novel *LBH* regulatory region in RA FLS

We were particularly struck by the fact that the analysis of regulatory regions identified *LBH*, which was also discovered in our original analysis due to promoter DMLs. The *LBH* DML in the present analysis was found in an intergenic regulatory region 6kb upstream of the *LBH* transcription start site (Fig 2A). It contains a non-CpG RA/celiac-associated risk variant (rs906868, G [RA-risk allele, Ref] → T [protective, SNP]), and a DML that we identified, located 185bp downstream of the variant (Fig S1). We first confirmed hypomethylation of this DML by pyrosequencing 6 OA and 9 RA FLS lines (Fig 2B). We then performed chromatin immunoprecipitation on RA FLS nuclear extracts with anti-H3K4me1 antibody and qPCR with primers specific to the region of interest (hg19-chr2:30447672-30448982). The region is significantly enriched in H3K4me1 mark in FLS (1.35±0.3 percent of input; $t$-test, n = 4 RA FLS lines, $p$=0.005, Fig 2C), which supports its potential role as an enhancer(21). Myt1 primer, which binds repressed chromatin histone mark H3K27me3, was used as a negative control and was enriched 0.17±0.03 percent of input in the enhancer region.

## *LBH* enhancer region regulates gene transcription in FLS

The *LBH* enhancer region containing either the Ref (i.e., risk) or SNP (i.e., protective) allele was cloned into pGL4.23 luciferase reporter construct, transfected into RA FLS and analyzed for luciferase activity. As shown in Fig 3, both alleles increased luciferase activity compared with control, which confirms the enhancer's regulatory function. Of interest, SNP allele-driven transcription was significantly higher than the Ref allele (SNP: 2.9±0.6; Ref: 1.6±0.3, n=12 RA lines; paired $t$-test, $p$<0.0001). These data confirm that the region is an

enhancer and that SNP and Ref alleles differentially regulate *LBH* enhancer function, with lower transcriptional activity associated with the Ref (risk) allele.

### RA risk allele regulates *LBH* expression in FLS.

The differential effects of the Ref and SNP on transcription was then confirmed by determining if the risk allele for RA (G/G) is associated with lower *LBH* gene transcription in FLS. To identify RA FLS lines with homozygous Ref (G) or SNP (T) alleles, we genotyped 58 cell lines. Seven RA lines (12%) were homozygous for the Ref allele, 19 (32%) were homozygous for the SNP allele and 32 (55%) lines were heterozygous. The allele frequencies of G and T were 34.5% and 65.5%, respectively. *LBH* expression in the two homozygous genotypes was then determined. RA FLS lines with SNP alleles expressed significantly higher levels of *LBH* than the lines with Ref alleles, consistent with the enhancer functional assay (Fig 4; Ref: 0.65 ± 0.08, n=7; SNP 1.2 ± 0.13, n=9; *t*-test, $p$=0.01). These data indicate that the Ref allele decreases *LBH* enhancer activity, contributing to low *LBH* gene expression in RA FLS.

### Effect of DNA methylation on *LBH* enhancer function

Because a DML is located 185bp from the SNP, we determined whether the presence of the Ref or SNP allele is responsible for the DML methylation, such as by recruiting DNA methyltransferases to specific loci nearby. Pyrosequencing of 7 Ref homozygous and 14 SNP homozygous RA lines showed that the methylation at the RA DML was similar in both groups, indicating that the Ref and SNP alleles are likely not responsible for hypomethylation at the DML (median % methylation: Ref = 9.4, SNP = 9.0; Mann-Whitney test, p=0.62, Fig 5A). We then evaluated whether enhancer methylation alters its transcriptional activity by methylating all CpG loci in the Ref, SNP or control plasmids in vitro (Fig S2). We transfected methylated Ref, SNP or control plasmids and compared enhancer activity to their respective unmethylated plasmids. Figure 5B shows that enhancer methylation decreased transcription in both the SNP and Ref alleles (Inhibition: Ref 36±5% and SNP 30±9%, n=7 different RA lines/group, paired *t-test*; Fig 5B). These data suggest that the enhancer function is determined by a combination of its sequence (SNP vs. Ref) and its methylation state.

### Transcription factor binding to the enhancer region in Ref and SNP RA FLS

To determine if altered transcription factor binding to enhancer with Ref or SNP allele might contribute to differences in *LBH* expression, we screened the variant sequences using HaploReg(22), ENCODE ChIP-seq database (Genome Browser), TOMTOM(23) and LASAGNA –Search2.0(24). LASAGNA-Search2.0 and TOMTOM indicated a sequence specific binding of the transcriptional repressor RP58 and GATA6 while the HaploReg and ENCODE ChIP-seq database indicated that RNA polymerase II (pol II) and CTCF bind the enhancer region. RP58 and GATA6 mRNA expression was confirmed in RA FLS (data not shown). However, ChIP-qPCR analysis did not detect RP58 or GATA6 binding to the enhancer. In contrast, RNA pol II binding was detected in the enhancer but was unaffected by the presence of the variant (Fig 6). CTCF binding was very low or undetectable in both groups (data not shown). Together, these data show that RNA pol II binds the enhancer region to increase *LBH* expression in both SNP and Ref lines.

## Discussion

Over 100 non-HLA GWAS SNPs have been identified as RA-risk loci, at least 19 of which are categorized as missense variants(25). The best characterized are functional variants in PTPN22 and IL-6 receptor genes, which alter protein function and therefore, increase susceptibility to multiple autoimmune diseases(26-28). However, the majority of GWAS RA-risk alleles are intra- or intergenic non-coding variants without obvious phenotype or effect on protein function(29, 30). Non-coding GWAS variants are frequently located in DNaseI hypersensitivity sites and can potentially regulate gene expression in a cell- or stimulus-specific way(29). Identifying how these non-coding variants regulate gene expression, function, and disease susceptibility can increase our understanding of disease pathogenesis.

We previously identified candidate genes by a systematic integrative analysis of three genome-wide FLS data sets, namely promoter DMLs, mRNA expression and GWAS SNPs(9). Of the original candidate genes, we focused on *LBH*, a highly conserved transcriptional coactivator, since very little was known about its function in RA. Subsequent functional studies showed that *LBH* expression in FLS is regulated by PDGF, which is a major growth factor that contributes to synovial lining hyperplasia in RA(11). Low *LBH* expression, which has been previously documented for RA FLS(31), allows cells to progress from G1 to S phase and can contribute to hyperproliferative state of RA FLS. Understanding why low *LBH* expression could participate in RA pathogenesis can provide insights into disease mechanisms and possibly novel therapeutic targets.

In the present study, we expanded our original promoter DML analysis to include DMLs in DNaseI hypersensitive sites, which are likely transcriptional enhancers with RA specific activity. The full integrative analysis identified a set of 12 genes, in which *LBH* appeared twice: once because of a DML in its promoter and second, in a novel regulatory region. The intriguing co-localization of an RA-associated variant and an RA DML in an enhancer led us to dissect how this enhancer might regulate a pathogenic gene in RA.

After confirming that the enhancer regulates transcription, we demonstrated that the RA-risk (Ref) allele is associated with reduced enhancer activity compared with the protective SNP allele (rs906868). Furthermore, *LBH* expression was lower in RA FLS that were homozygous for the Ref allele compared with the SNP allele. This observation suggested that the polymorphism could contribute to increased FLS proliferation in RA. *LBH* gene expression is also reduced in SLE peripheral blood mononuclear cells (PBMC) compared with control, suggesting that *LBH* could play a role in multiple autoimmune diseases and that various *LBH* risk alleles might participate(15). *LBH* expression was modestly reduced in Ref compared with SNP PBMCs in SLE, although the mechanisms were not defined(13). In other studies, *LBH* over-expression increased phosphorylation of NF-κB inhibitor, IκB, while loss of *LBH* enhanced NF- κB activity, indicating that lower *LBH* expression promotes inflammation(12).

Although the focus of our study was rs906868, several *LBH* reference variants outside the enhancer region have been identified and are associated with RA(14). These variants have

high linkage disequilibrium (D'>0.95) with rs906868 (Fig S3). Recent studies showed four *LBH* SNPs (rs7579944, rs1355208, rs1396838, rs906868) are modestly protective in RA but confer risk in celiac disease; while the role of *LBH* vary depending on the disease, the data suggest that the gene plays a key role in many inflammatory states(14). These results are consistent with a meta-analysis study on Asian and European RA patients demonstrating that rs7579944 confers significant risk(32). The individual and cumulative effects of these variants on enhancer function can ultimately be evaluated using recently developed high-throughput strategies(33, 34).

We were particularly interested in the interactions between epigenetic marks, RA-associated SNPs and enhancer function. The fact that a differentially methylated locus was identified in such close proximity to a functional SNP led us to explore how methylation might affect the enhancer function. The complex relationship between non-CpG GWAS variants, the methylation status of neighboring CpG loci and their effect on gene expression is not well understood. In response to temporal or environmental cues, methylation can be associated with increased or decreased gene expression in enhancers, introns, and gene bodies, but the mechanisms are unclear(35). Recent genome-wide studies have shown that disrupting transcription factor binding motifs causes changes in DNA methylation(36, 37) and that SNPs can regulate (i) gene expression through methylation, (ii) methylation through expression or (iii) methylation and gene expression independently(38).

Our data show that altering *LBH* enhancer methylation decreased its activity, independent of the Ref or SNP variant. The CpG locus near the RA-associated variant is hypomethylated and is associated with lower LBH expression might appear counter-intuitive. However, we cannot assume that low methylation levels at this particular site is a primary event, as it could also be a specific regulatory mechanism to increase LBH expression and suppress FLS proliferation. The mechanism of differential methylation *in vivo* is an area of active investigation to address these questions. For example, non-coding RNAs that direct DNA methyltransferases(39) or steric hindrance from transcription factors could play a role, but are not yet defined in *LBH*. A direct effect of the SNP itself is unlikely because the SNP and Ref did not influence the methylation status of the enhancer. Future studies to evaluate the contribution of other CpGs in the enhancer and the mechanisms responsible for altered DNA methylation are needed to define the interplay between epigenetics and LBH expression in RA.

We also explored whether differential binding of transcription factors could account for lower transcription in Ref. Several candidate factors that could potentially bind at or near the SNP site were identified. However, ChIP-qPCR did not demonstrate a consistent difference between the two alleles, including RNA pol II. The SNP might affect enhancer activity by disrupting the DNA-binding site of an uncharacterized epigenomic regulator in FLS. Exploring this possibility would require the creation of genome-wide chromatin maps for RA FLS, which is an area of active investigation(30). Another possibility is that the SNP alters the sequence of enhancer RNAs that are transcribed from active enhancers(40) and regulate gene expression(41, 42) by forming complex with transcription factors and stabilizing their interaction with DNA(43).

In conclusion, an expanded integrative analysis of regulatory regions in RA FLS identified a novel *LBH* enhancer with an RA-associated polymorphism that affects enhancer function. The polymorphism is closely associated with a differentially methylated CpG locus, and methylation status of the region can dramatically affect enhancer function and gene expression. In RA, the combination of genetic and epigenetic marks can lead to the dysregulation of a pathogenic gene and subsequently, might contribute to synovial hyperplasia by allowing cells to progress through the G1 checkpoint and initiating the cell cycle.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## References

1. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat Biotechnol. 2013; 31:142–7. [PubMed: 23334450]
2. Viatte S, Plant D, Raychaudhuri S. Genetics and epigenetics of rheumatoid arthritis. Nat Rev Rheumatol. 2013; 9:141–53. [PubMed: 23381558]
3. Catrina AI, Joshua V, Klareskog L, Malmstrom V. Mechanisms involved in triggering rheumatoid arthritis. Immunol Rev. 2016; 269:162–74. [PubMed: 26683152]
4. Bottini N, Firestein GS. Duality of fibroblast-like synoviocytes in RA: passive responders and imprinted aggressors. Nat Rev Rheumatol. 2013; 9:24–33. [PubMed: 23147896]
5. Lefevre S, Knedla A, Tennie C, Kampmann A, Wunrau C, Dinser R, et al. Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. Nat Med. 2009; 15:1414–20. [PubMed: 19898488]
6. Steenvoorden MM, Tolboom TC, van der Pluijm G, Lowik C, Visser CP, DeGroot J, et al. Transition of healthy to diseased synovial tissue in rheumatoid arthritis is associated with gain of mesenchymal/fibrotic characteristics. Arthritis Res Ther. 2006; 8:R165. [PubMed: 17076892]
7. Nakano K, Whitaker JW, Boyle DL, Wang W, Firestein GS. DNA methylome signature in rheumatoid arthritis. Ann Rheum Dis. 2013; 72:110–7. [PubMed: 22736089]
8. Whitaker JW, Shoemaker R, Boyle DL, Hillman J, Anderson D, Wang W, et al. An imprinted rheumatoid arthritis methylome signature reflects pathogenic phenotype. Genome Med. 2013; 5:40. [PubMed: 23631487]
9. Whitaker JW, Boyle DL, Bartok B, Ball ST, Gay S, Wang W, et al. Integrative omics analysis of rheumatoid arthritis identifies non-obvious therapeutic targets. PLoS One. 2015; 10:e0124254. [PubMed: 25901943]
10. Briegel KJ, Baldwin HS, Epstein JA, Joyner AL. Congenital heart disease reminiscent of partial trisomy 2p syndrome in mice transgenic for the transcription factor Lbh. Development. 2005; 132:3305–16. [PubMed: 15958514]
11. Ekwall AK, Whitaker JW, Hammaker D, Bugbee WD, Wang W, Firestein GS. The Rheumatoid Arthritis Risk Gene LBH Regulates Growth in Fibroblast-like Synoviocytes. Arthritis Rheumatol. 2015; 67:1193–202. [PubMed: 25707478]
12. Liu Q, Guan X, Lv J, Li X, Wang Y, Li L. Limb-bud and Heart (LBH) functions as a tumor suppressor of nasopharyngeal carcinoma by inducing G1/S cell cycle arrest. Sci Rep. 2015; 5:7626. [PubMed: 25557837]
13. Chang Y, Sheng Y, Cheng Y, Lin Y, Zhu Z, Wen L, et al. Downregulated expression of LBH mRNA in peripheral blood mononuclear cells from patients with systemic lupus erythematosus. J Dermatol. 2015 DOI 10.1111/1346-8138.13006.
14. Zhernakova A, Stahl EA, Trynka G, Raychaudhuri S, Festen EA, Franke L, et al. Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. PLoS Genet. 2011; 7:e1002004. [PubMed: 21383967]

15. Yu ZY, Lu WS, Zuo XB, Hu J, Yao S, Li Y, et al. One novel susceptibility locus associate with systemic lupus erythematosus in Chinese Han population. Rheumatol Int. 2013; 33:2079–83. [PubMed: 23408151]
16. Whitaker JW, Nguyen TT, Zhu Y, Wildberg A, Wang W. Computational schemes for the prediction and annotation of enhancers from epigenomic assays. Methods. 2015; 72:86–94. [PubMed: 25461775]
17. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The american rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis & Rheumatism. 1988; 31:315–24. [PubMed: 3358796]
18. Rosengren S, Boyle DL, Firestein GS. Acquisition, culture, and phenotyping of synovial fibroblasts. Methods Mol Med. 2007; 135:365–75. [PubMed: 17951672]
19. Boyle DL, Rosengren S, Bugbee W, Kavanaugh A, Firestein GS. Quantitative biomarker analysis of synovial gene expression by real-time PCR. Arthritis Res Ther. 2003; 5:R352–60. [PubMed: 14680510]
20. You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, et al. BatchPrimer3: a high throughput web application for PCR and sequencing primer design. BMC Bioinformatics. 2008; 9:253. [PubMed: 18510760]
21. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet. 2007; 39:311–8. [PubMed: 17277777]
22. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. Nucleic Acids Research. 2016; 44:D877–D81. [PubMed: 26657631]
23. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. Genome Biol. 2007; 8:R24. [PubMed: 17324271]
24. Lee C, Huang CH. LASAGNA-Search 2.0: integrated transcription factor binding site search and visualization in a browser. Bioinformatics. 2014; 30:1923–5. [PubMed: 24578403]
25. Li MJ, Liu Z, Wang P, Wong MP, Nelson MR, Kocher JP, et al. GWASdb v2: an update database for human genetic variants identified by genome-wide association studies. Nucleic Acids Res. 2016; 44:D869–76. [PubMed: 26615194]
26. Stanford SM, Bottini N. PTPN22: the archetypal non-HLA autoimmunity gene. Nat Rev Rheumatol. 2014; 10:602–11. [PubMed: 25003765]
27. Ferreira RC, Freitag DF, Cutler AJ, Howson JM, Rainbow DB, Smyth DJ, et al. Functional IL6R 358Ala allele impairs classical IL-6 receptor signaling and influences risk of diverse inflammatory diseases. PLoS Genet. 2013; 9:e1003444. [PubMed: 23593036]
28. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. 2014; 506:376–81. [PubMed: 24390342]
29. Farh KK-H, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015; 518:337–43. [PubMed: 25363779]
30. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015; 518:317–30. [PubMed: 25693563]
31. Del Rey MJ, Usategui A, Izquierdo E, Canete JD, Blanco FJ, Criado G, et al. Transcriptome analysis reveals specific changes in osteoarthritis synovial fibroblasts. Ann Rheum Dis. 2012; 71:275–80. [PubMed: 22021863]
32. Kim K, Bang SY, Lee HS, Cho SK, Choi CB, Sung YK, et al. High-density genotyping of immune loci in Koreans and Europeans identifies eight new rheumatoid arthritis risk loci. Ann Rheum Dis. 2015; 74:e13. [PubMed: 24532676]
33. Arnold CD, Gerlach D, Stelzer C, Boryn LM, Rath M, Stark A. Genome-wide quantitative enhancer activity maps identified by STARR-seq. Science. 2013; 339:1074–7. [PubMed: 23328393]
34. Dailey L. High throughput technologies for the functional discovery of mammalian enhancers: new approaches for understanding transcriptional regulatory network dynamics. Genomics. 2015; 106:151–8. [PubMed: 26072436]

35. Plank JL, Dean A. Enhancer function: mechanistic and genome-wide insights come together. Mol Cell. 2014; 55:5–14. [PubMed: 24996062]
36. Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. Nature. 2015; 523:212–6. [PubMed: 26030523]
37. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, et al. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature. 2011; 480:490–5. [PubMed: 22170606]
38. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet. 2011; 12:529–41. [PubMed: 21747404]
39. Lai F, Shiekhattar R. Where long noncoding RNAs meet DNA methylation. Cell Res. 2014; 24:263–64. [PubMed: 24481530]
40. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010; 465:182–7. [PubMed: 20393465]
41. Lam MT, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, et al. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. Nature. 2013; 498:511–5. [PubMed: 23728303]
42. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. Nat Rev Mol Cell Biol. 2015; 16:144–54. [PubMed: 25650801]
43. Sigova AA, Abraham BJ, Ji X, Molinie B, Hannett NM, Guo YE, et al. Transcription factor trapping by RNA in gene regulatory elements. Science. 2015; 350:978–81. [PubMed: 26516199]

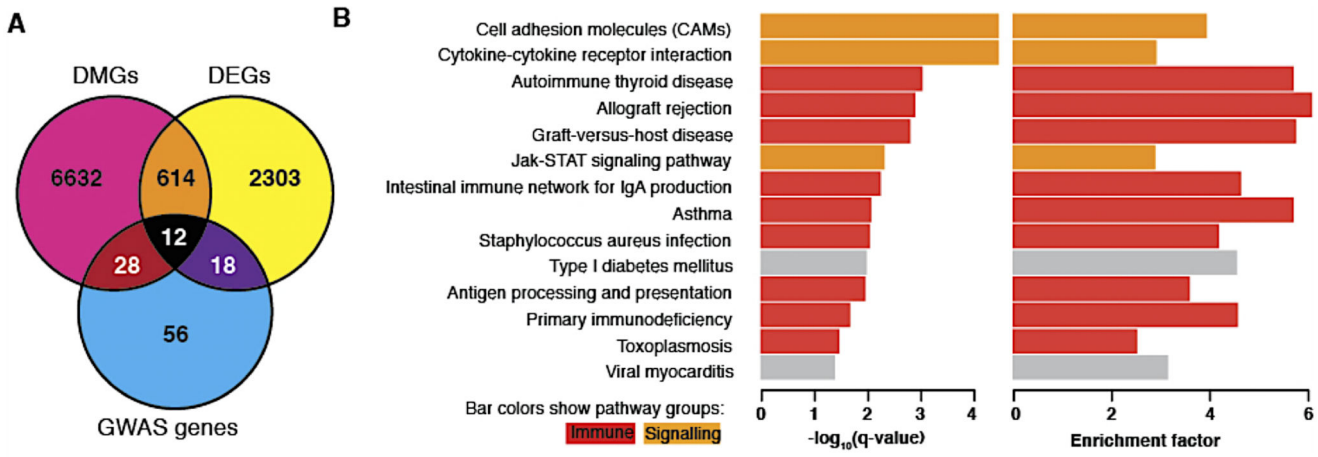


**Fig 1. Expanded integrative analysis of omics data and pathway analysis**

(A) A Venn diagram showing the overlap of three datasets: differentially methylated genes (DMGs) associated with differentially methylated loci at promoters and putative enhancers, differentially expressed genes (DEGs) and genome-wide association studies data (GWAS). (B) KEGG pathway enrichment analysis was performed using genes that were present in 2 or 3 of the datasets.

$$Enrichment \quad Factor = (numberoflocifromgeneAbeneathcut - off / totallocifromgeneA / (totalnumberoflocibeneathcut - off / totalnumberofloci)$$

*P*-values were calculated using the hypergeometric distribution and were corrected for multiple testing to produce *q*-values. Pathways with *q*-values <0.05 are shown.

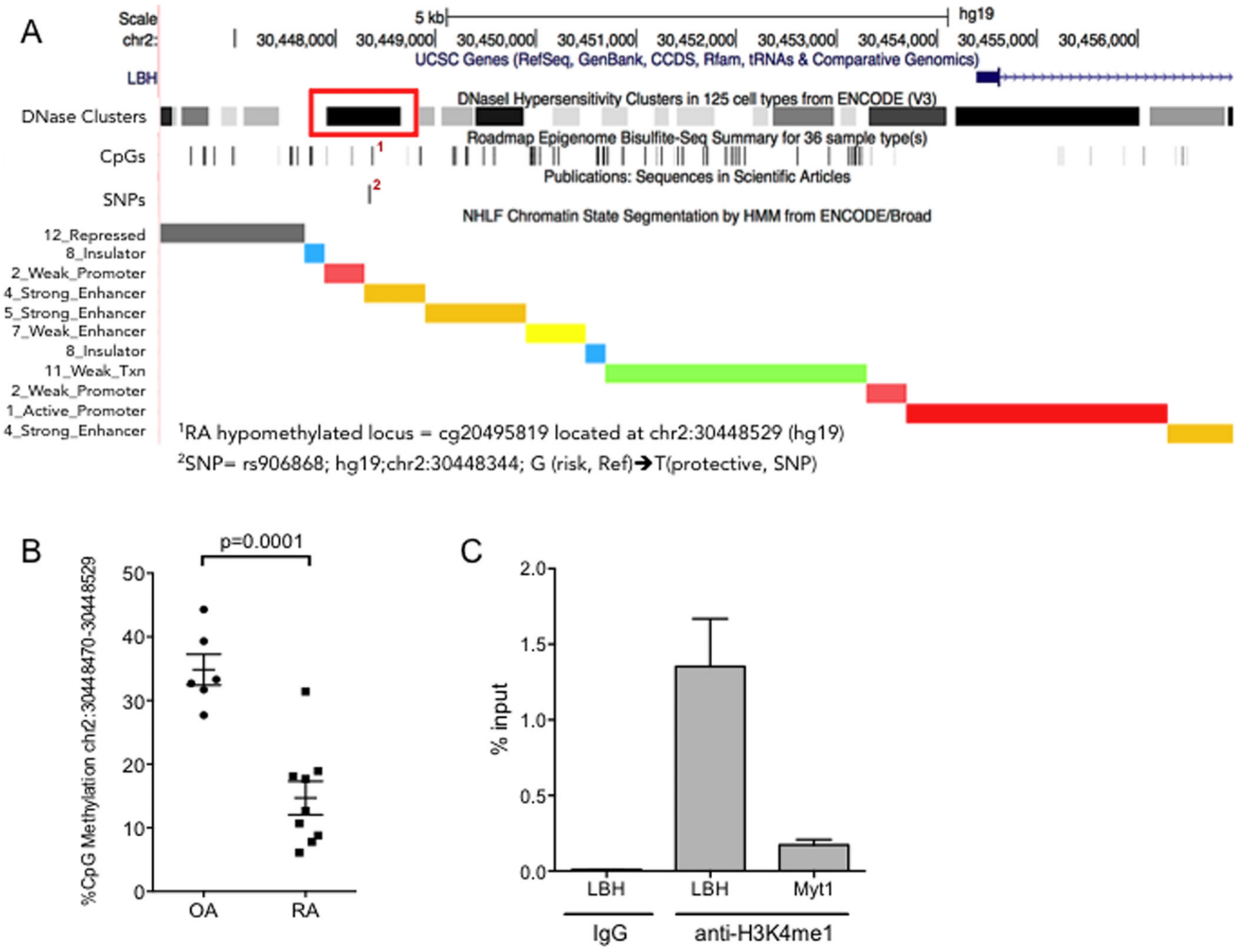


**Fig 2. Identification of a novel *LBH* regulatory region in RA FLS**

(A) The putative enhancer is in a DNaseI hypersensitive region located approximately 6kb upstream of the TSS. This region contains a non-CpG risk-variant (rs906868) and an RA hypomethylated locus (cg20495819) 185bp apart. (B) Pyrosequencing analysis confirmed significant hypomethylation of cg20495819 in RA compared with OA FLS. (C) FLS lysates were immunoprecipitated with anti-H3K4me1 antibody and the DNA was subjected to qPCR with enhancer-specific or negative control Myt1 primer. The enhancer region was significantly enriched for the H3K4me1 mark.

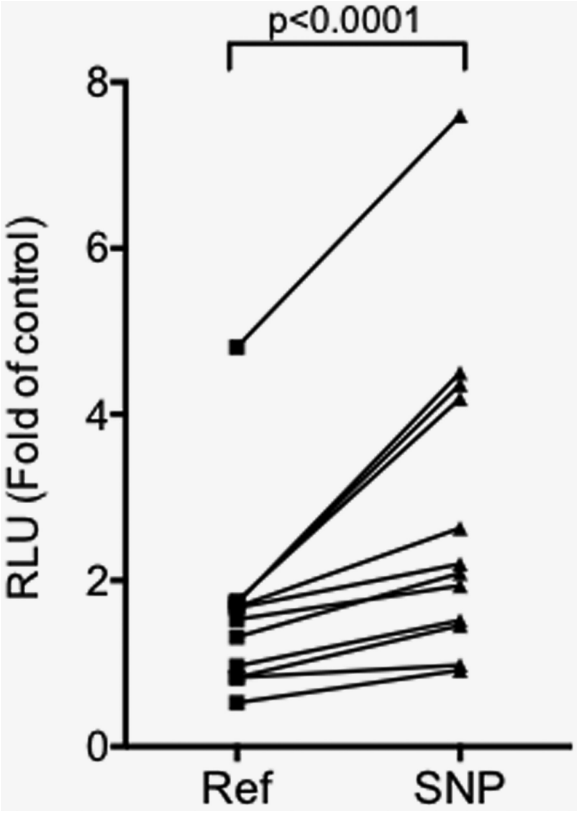


**Fig 3. *LBH* enhancer region regulates gene transcription in FLS**

The enhancer sequence with Ref or SNP allele was cloned into luciferase constructs, which were then transfected into RA FLS. Luciferase assays show significantly higher activity in SNP plasmids compared with Ref.

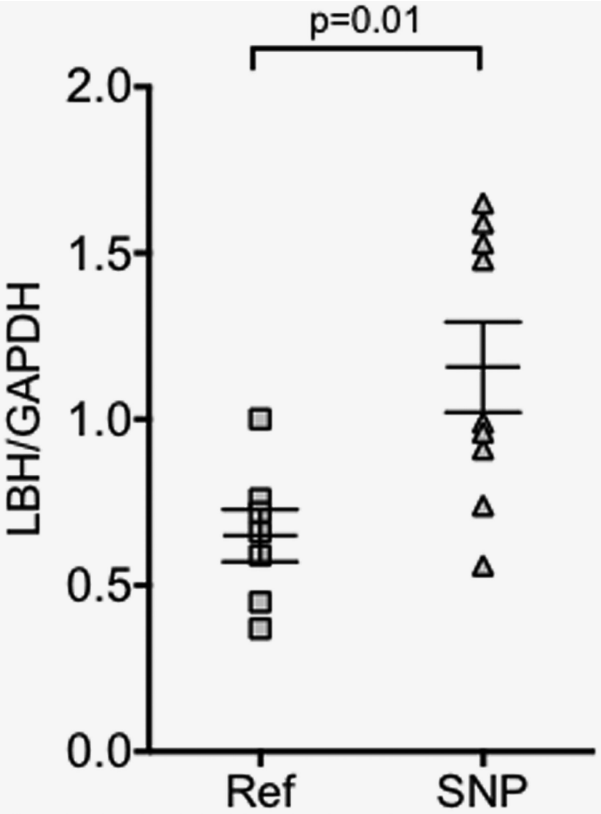


**Fig 4. RA risk allele regulates *LBH* expression in FLS**

RA FLS lines with Ref- or SNP-homozygous alleles were identified and *LBH* gene expression was measured by qPCR. *LBH* expression was higher in SNP- than the Ref-homozygous FLS lines.

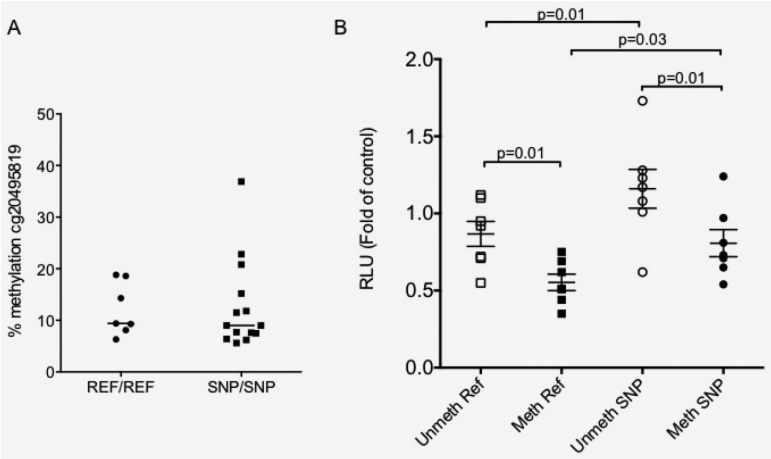


**Fig 5. Regulation and function of DNA methylation in *LBH* enhancer**

(A) Pyrosequencing was performed in RA FLS lines identified as Ref- or SNP-homozygous. Methylation of the DML was similar in both groups, indicating that the presence of the variant did not alter methylation status. (B) Fully methylated or unmethylated enhancer plasmids with Ref or SNP alleles were transfected into RA FLS and luciferase activity was measured. Methylation significantly decreased luciferase activity compared with unmethylated plasmids.

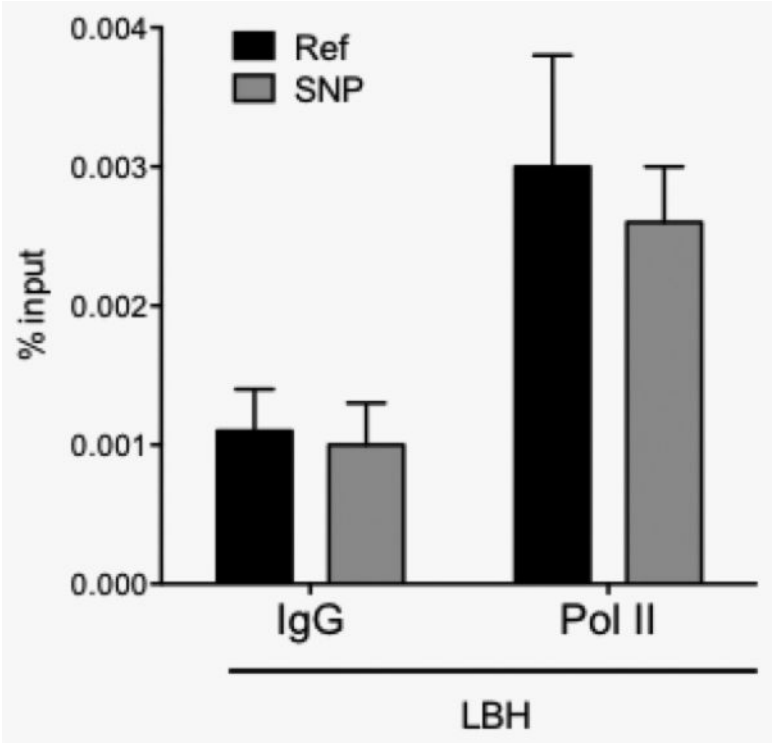


**Fig 6. Detection of transcription factor binding to the enhancer region**

Chromatin immunoprecipitation was performed using anti-RNA pol II antibody in Ref- or SNP-homozygous RA FLS lines and qPCR was performed with enhancer-specific primers. RNA pol II binding was detectable in the enhancer region with the SNP or Ref alleles.