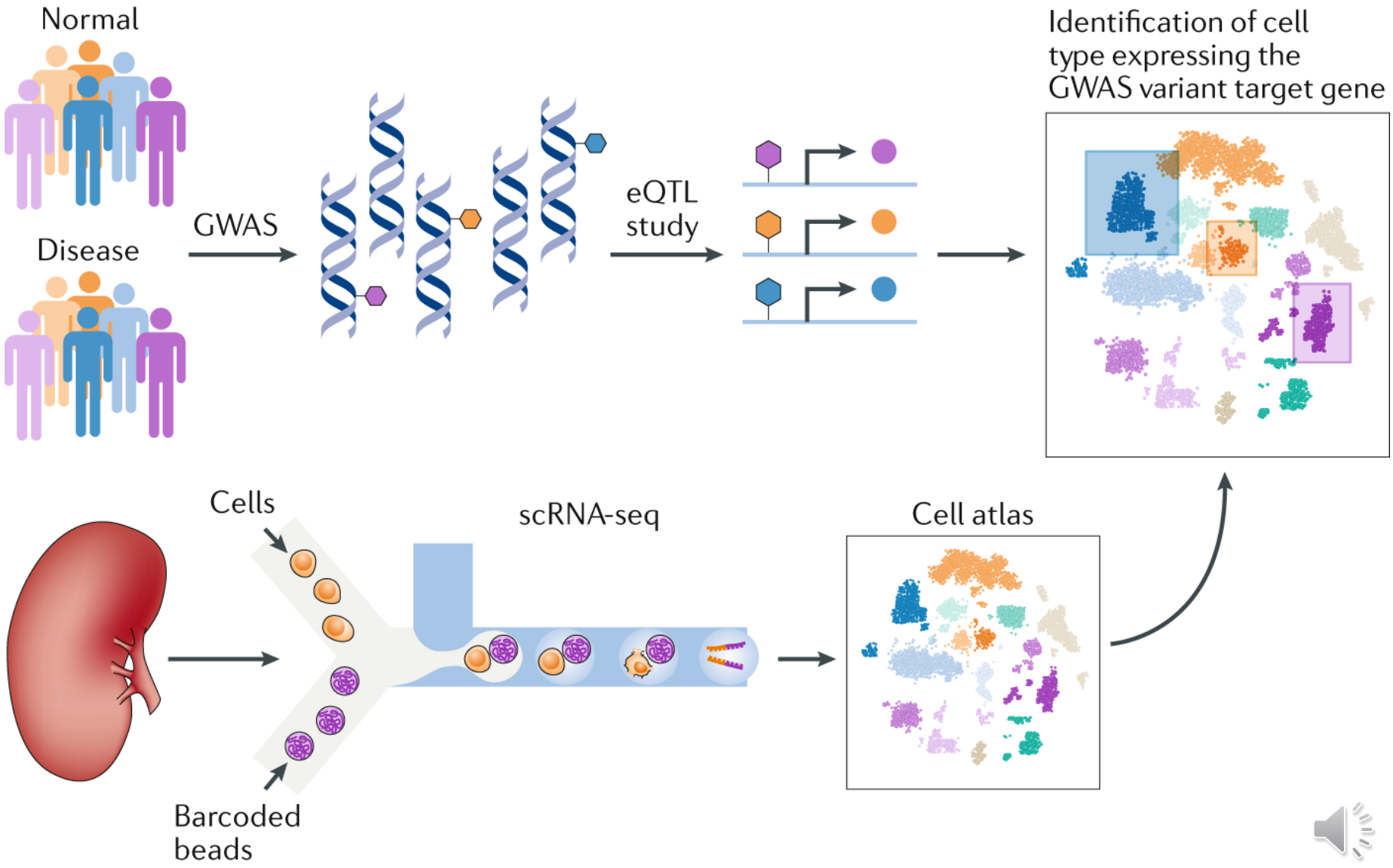
Single Cell Sequencing application to study single nucleotide variants



IMMUNE REGULATION

A long noncoding RNA associated with susceptibility to celiac disease

Ainara Castellanos-Rubio,[1] Nora Fernandez-Jimenez,[2] Radomir Kratchmarov,[1] Xiaobing Luo,[3] Govind Bhagat,[4,5] Peter H. R. Green,[5] Robert Schneider,[6] Megerditch Kiledjian,[3] Jose Ramon Bilbao,[2] Sankar Ghosh[1]*

Recent studies have implicated long noncoding RNAs (lncRNAs) as regulators of many important biological processes. Here we report on the identification and characterization of a lncRNA, lnc13, that harbors a celiac disease–associated haplotype block and represses expression of certain inflammatory genes under homeostatic conditions. Lnc13 regulates gene expression by binding to hnRNPD, a member of a family of ubiquitously expressed heterogeneous nuclear ribonucleoproteins (hnRNPs). Upon stimulation, lnc13 levels are reduced, thereby allowing increased expression of the repressed genes. Lnc13 levels are significantly decreased in small intestinal biopsy samples from patients with celiac disease, which suggests that down-regulation of lnc13 may contribute to the inflammation seen in this disease. Furthermore, the lnc13 disease-associated variant binds hnRNPD less efficiently than its wild-type counterpart, thus helping to explain how these single-nucleotide polymorphisms contribute to celiac disease.

Celiac disease (CeD)

is a chronic, immune-mediated intestinal disorder that is caused by intolerance to ingested gluten and develops in genetically susceptible individuals.

The NEW ENGLAND JOURNAL *of* MEDICINE

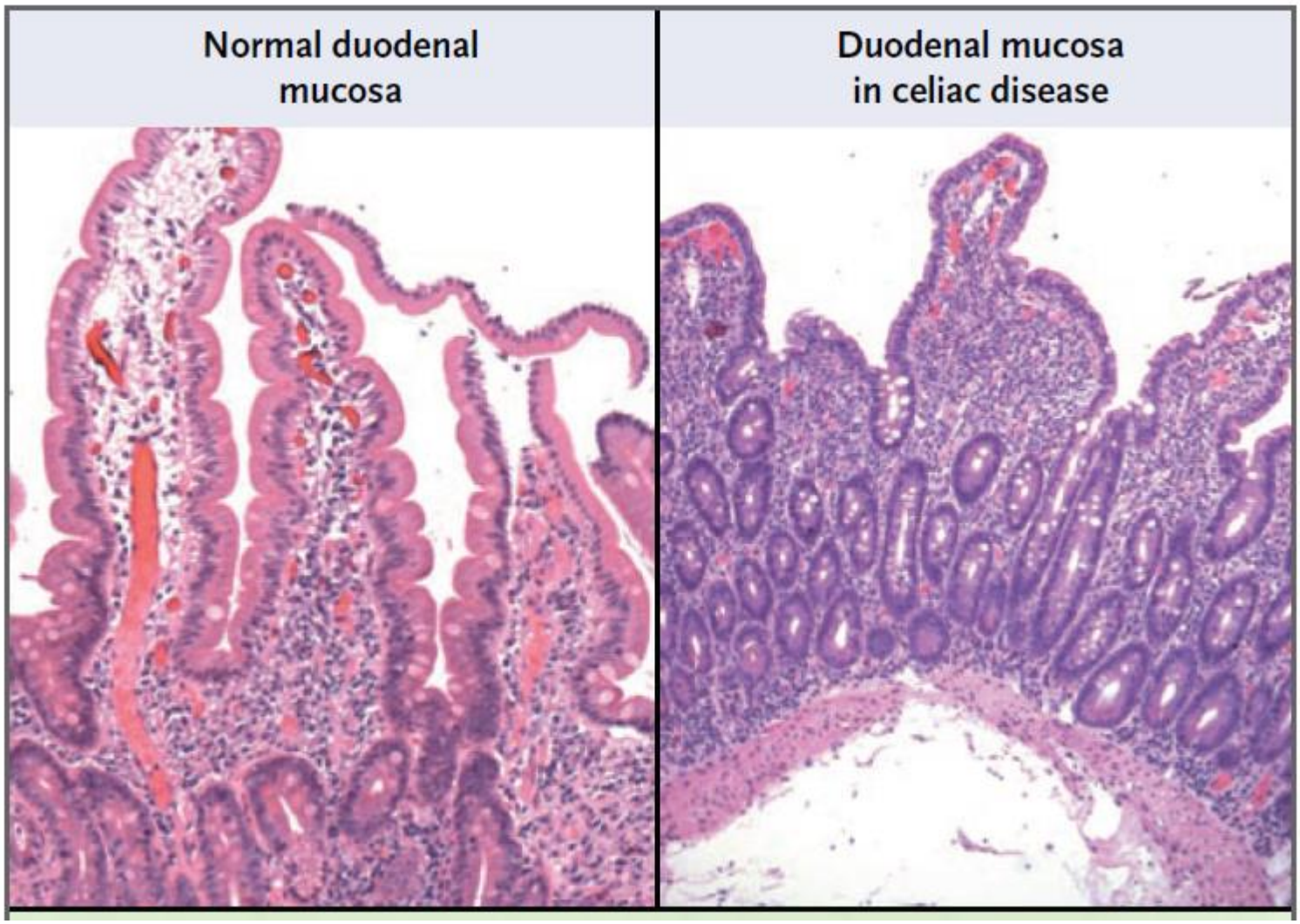
REVIEW ARTICLE

MEDICAL PROGRESS

Celiac Disease

Peter H.R. Green, M.D., and Christophe Cellier, M.D., Ph.D.

Immune system activation from gluten damage duodenal mucosa



Molecular Mechanisms that underpin CELIAC DISEASE

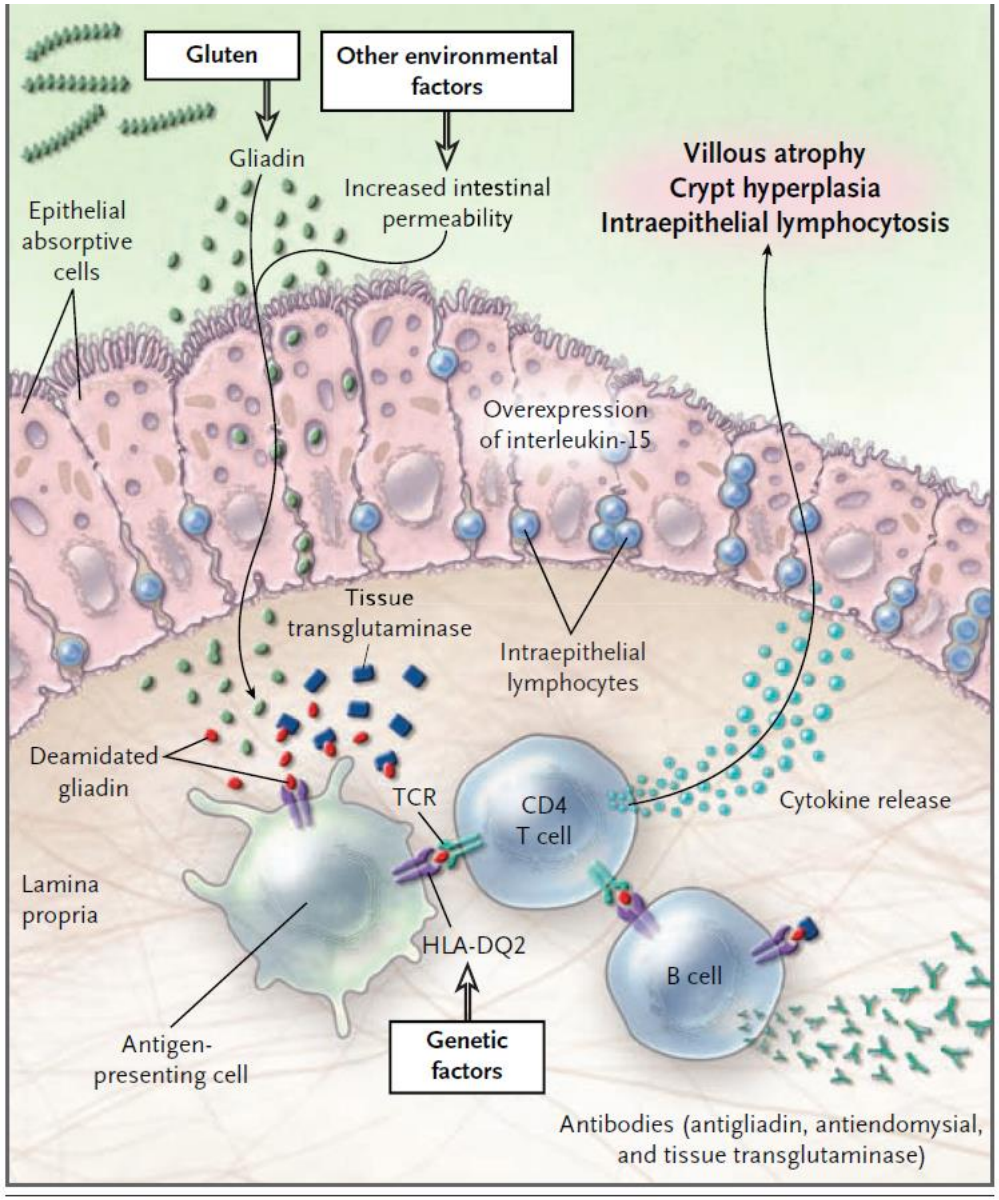


Figure 1. Interaction of Gluten with Environmental, Immune, and Genetic Factors in Celiac Disease.

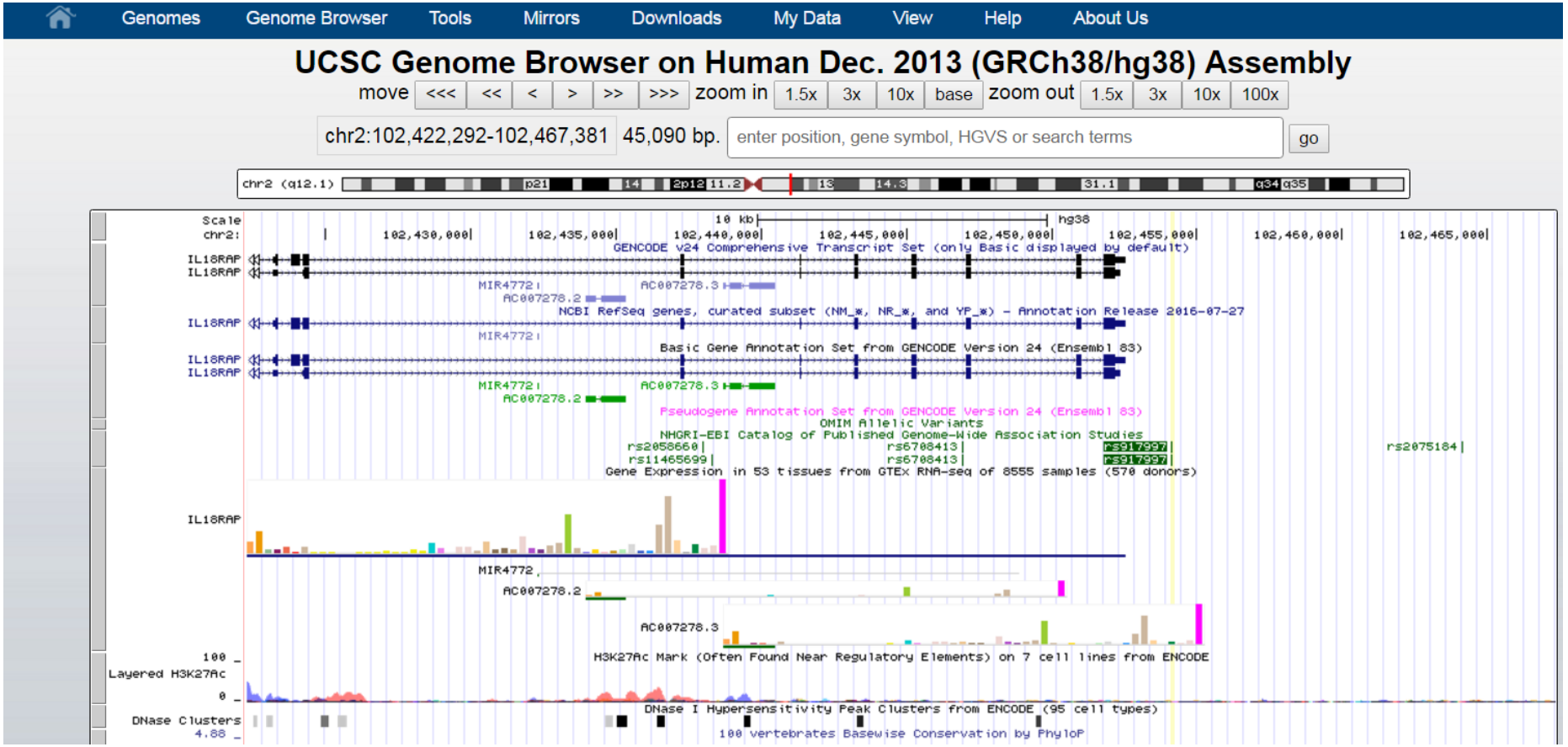
Gluten is digested by luminal and brush-border enzymes into amino acids and peptides. The gliadin peptides induce changes in the epithelium through the innate immune system and, in the lamina propria, through the adaptive immune system. In the epithelium, gliadin damages epithelial cells, resulting in increased expression of interleukin-15, which in turn activates intraepithelial lymphocytes. These lymphocytes become cytotoxic and kill enterocytes that express MIC-A (a stress protein) on their surface. During infections or as the result of permeability changes, gliadin enters the lamina propria, where it is deamidated by tissue transglutaminase, allowing interaction with HLA-DQ2 (or HLA-DQ8) on the surface of antigen-presenting cells. Gliadin is presented to gliadin-reactive CD4+ T cells through a T-cell receptor, resulting in the production of cytokines that cause tissue damage. This leads to villous atrophy and crypt hyperplasia, as well as the activation and expansion of B cells that produce antibodies. Images of mucosa courtesy of Govind Bhagat, M.D.

Steps for the identification of the role for SNP associated with Celiac Disease, rs917997

- **Annotation of SNP on the reference genome for the localization in the specific locus**

- **Searching in the Genome-wide database for epigenetic features that describe SNP-associated locus for cell type or tissues**

rs917997 is 1.5 kb apart from the 3′ end of the IL18RAP coding gene in chromosome 2



Human lnc13 genomic location is predicted based on homology with mouse lnc13 (TransMap alignment 66.5% identity). Conserved regions are represented as thicker lines on mouse lnc13 (AK161196)

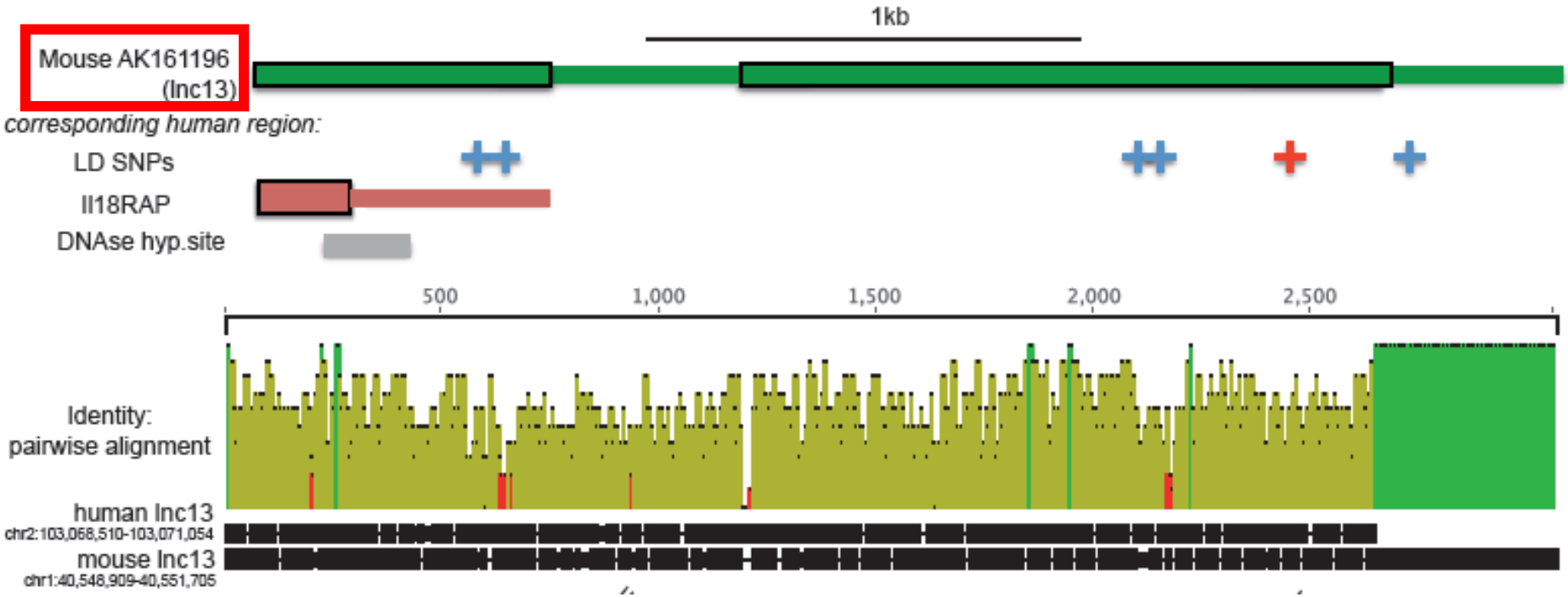
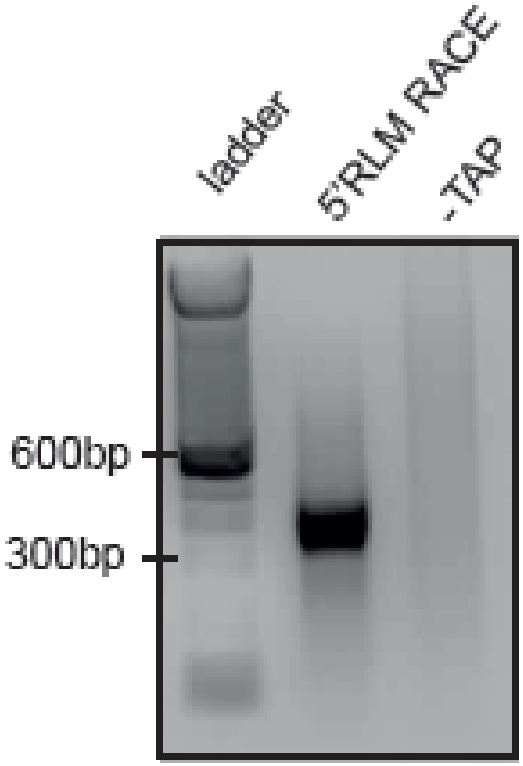


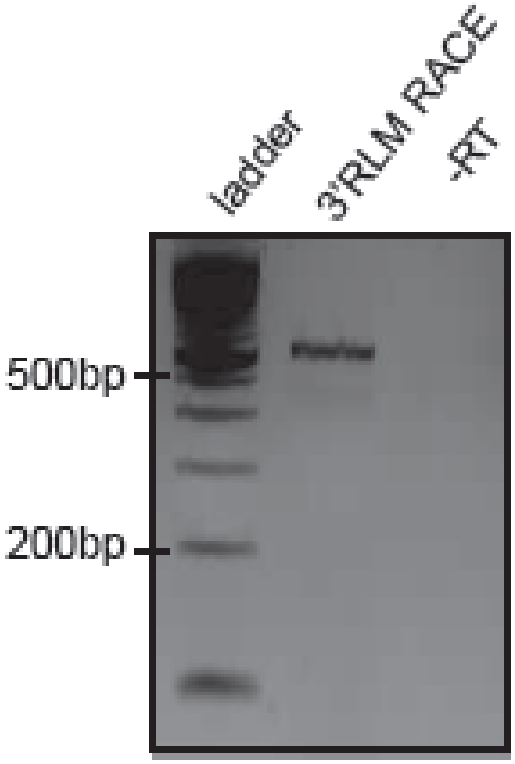
fig. S1A

Steps for the chacterization of Lnc13

- Cloning of Lnc13 using RACE-PCR
- Northen blot to identify Lnc13 expression
- Lnc 13 expression in cell lines and tissues using RNAscope and PCR

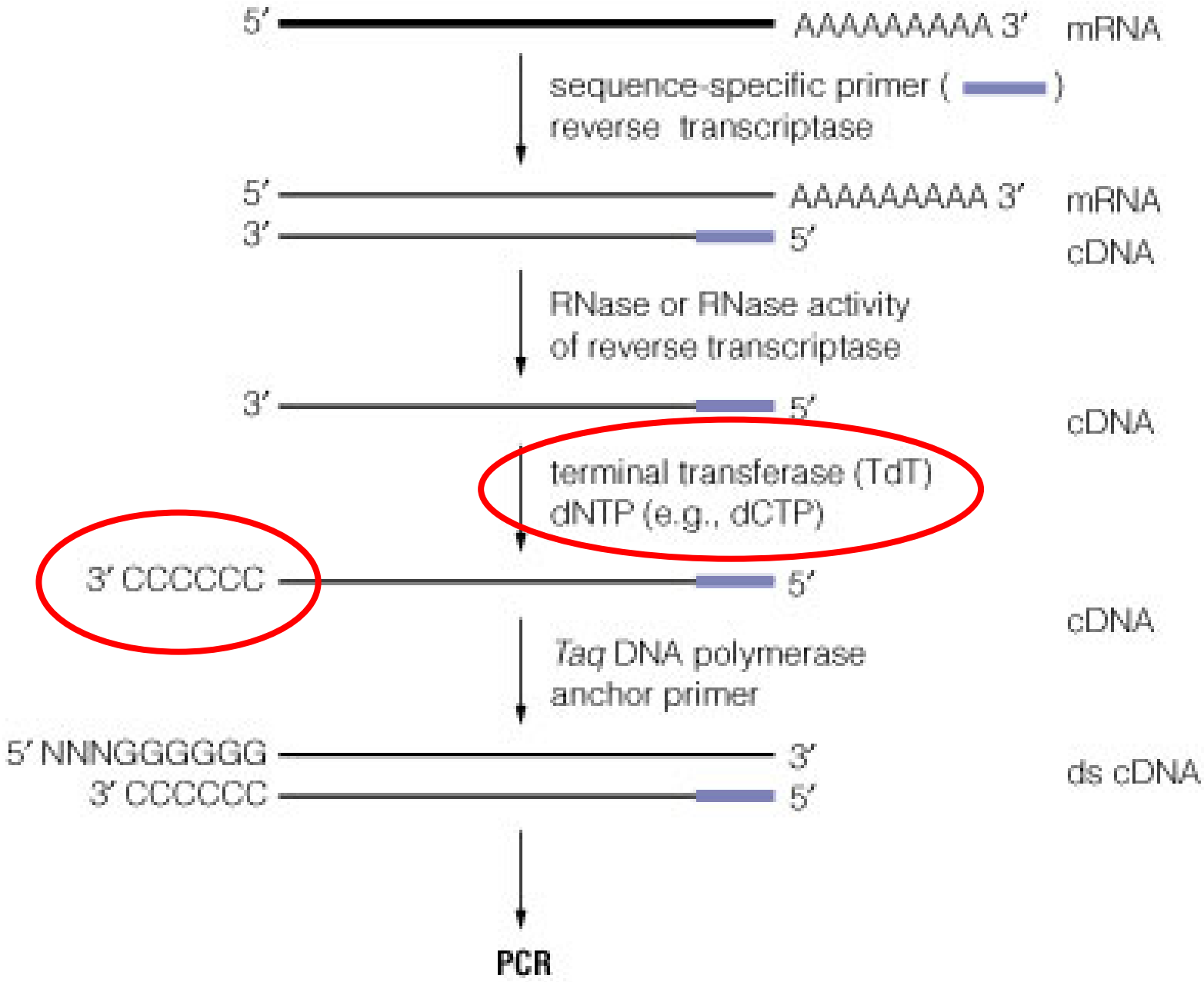
RACE-PCR is used for cloning Lnc13

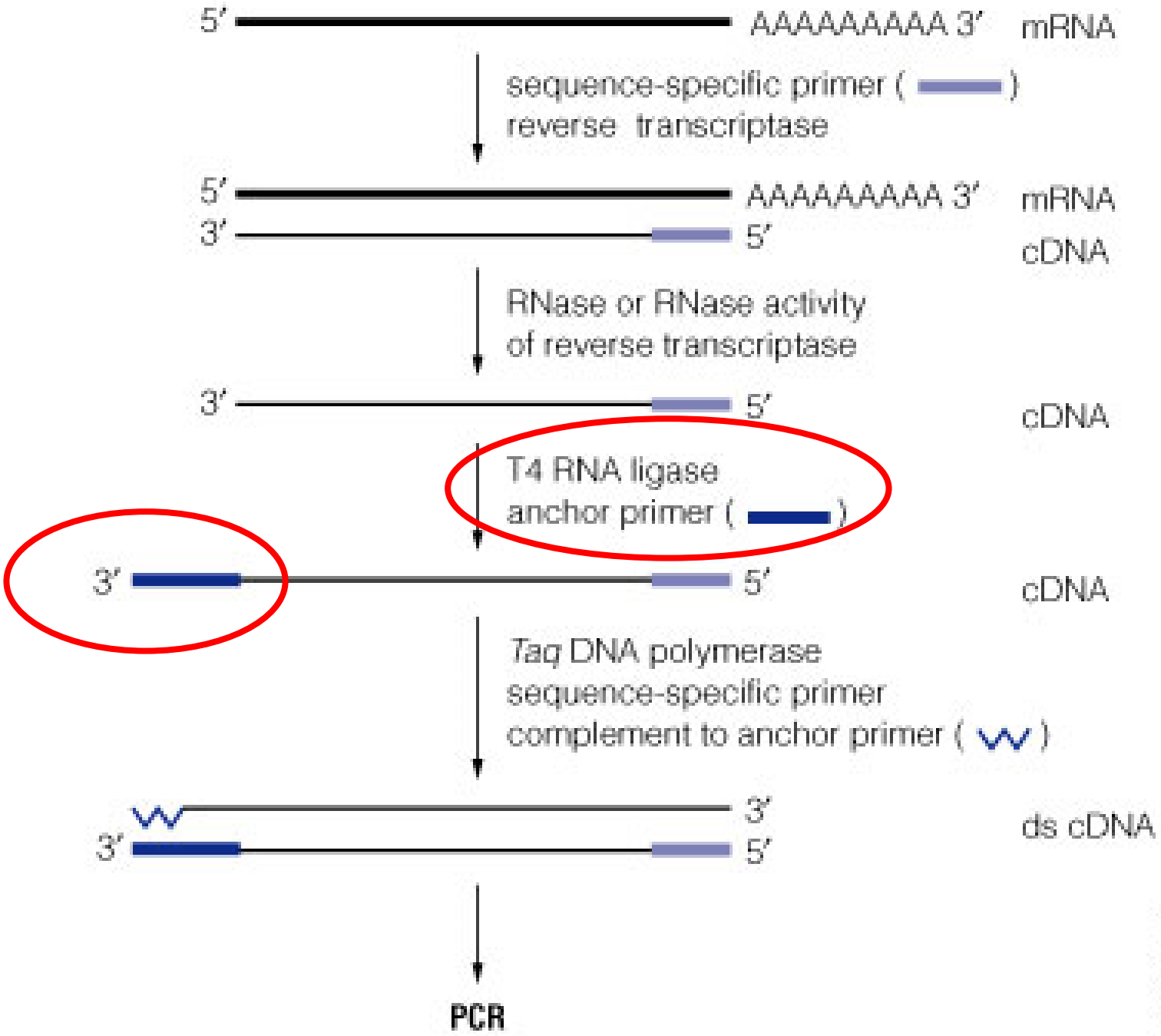




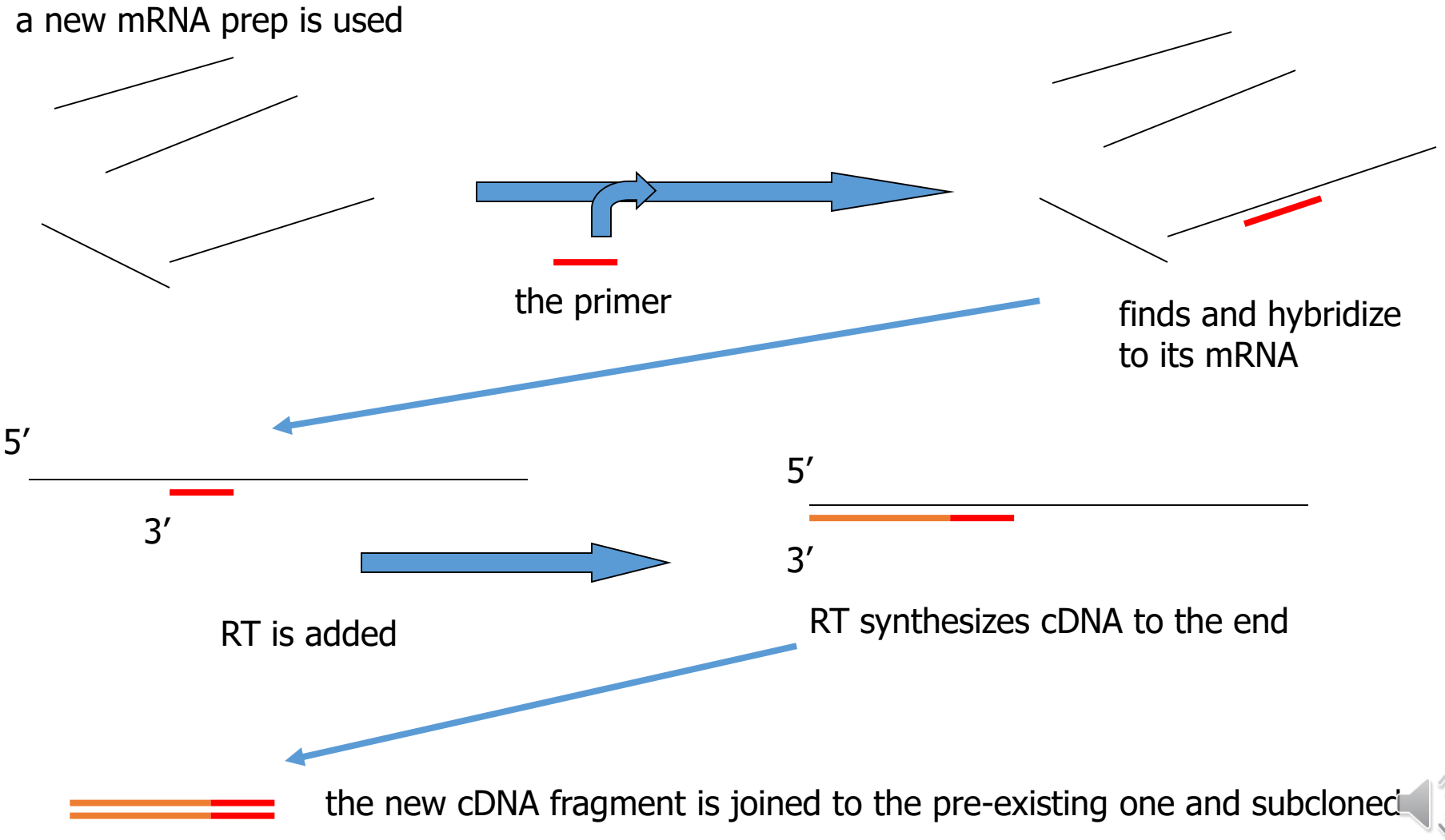
RACE

(rapid amplification of cDNA ends)


5′
AAAAAAAAA 3′
mRNA
sequence-specific primer ( )
reverse transcriptase
5′
AAAAAAAAA 3′
mRNA
3′
5′
cDNA
RNase or RNase activity
of reverse transcriptase
3′
5′
cDNA
terminal transferase (TdT)
dNTP (e.g., dCTP)
3′ CCCCCC
5′
cDNA
Taq DNA polymerase
anchor primer
5′ NNNGGGGGG
3′
3′ CCCCCC
5′
ds cDNA
PCR

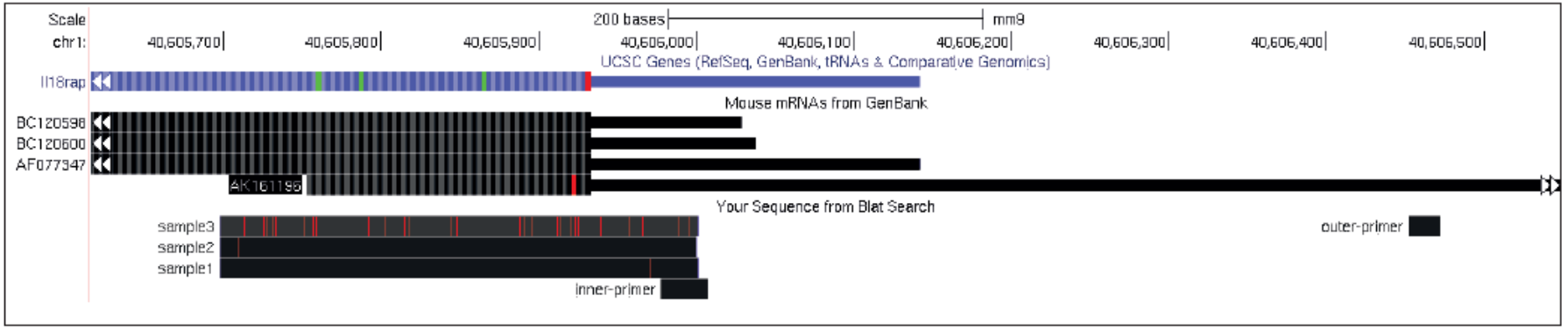

5′
AAAAAAAAA 3′
mRNA
sequence-specific primer ( )
reverse transcriptase
5′
AAAAAAAAA 3′
mRNA
3′
5′
cDNA
RNase or RNase activity
of reverse transcriptase
3′
5′
cDNA
T4 RNA ligase
anchor primer ( )
3′
5′
cDNA
Taq DNA polymerase
sequence-specific primer
complement to anchor primer ( )
3′
3′
5′
ds cDNA
PCR

An oligonucleotide primer is then prepared to match the very 5′ sequence of the cDNA, as determined by dideoxy-sequencing

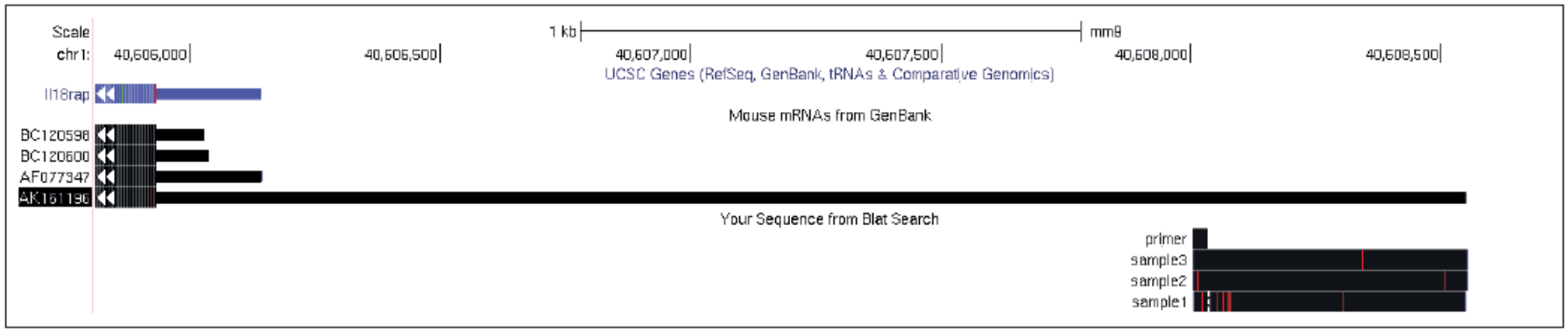


PCR bands were cloned, sequenced and blasted against the mouse genome using the UCSC genome browser. AK161196 (in a black box) is the annotated mouse lnc13

5'RLM RACE blat

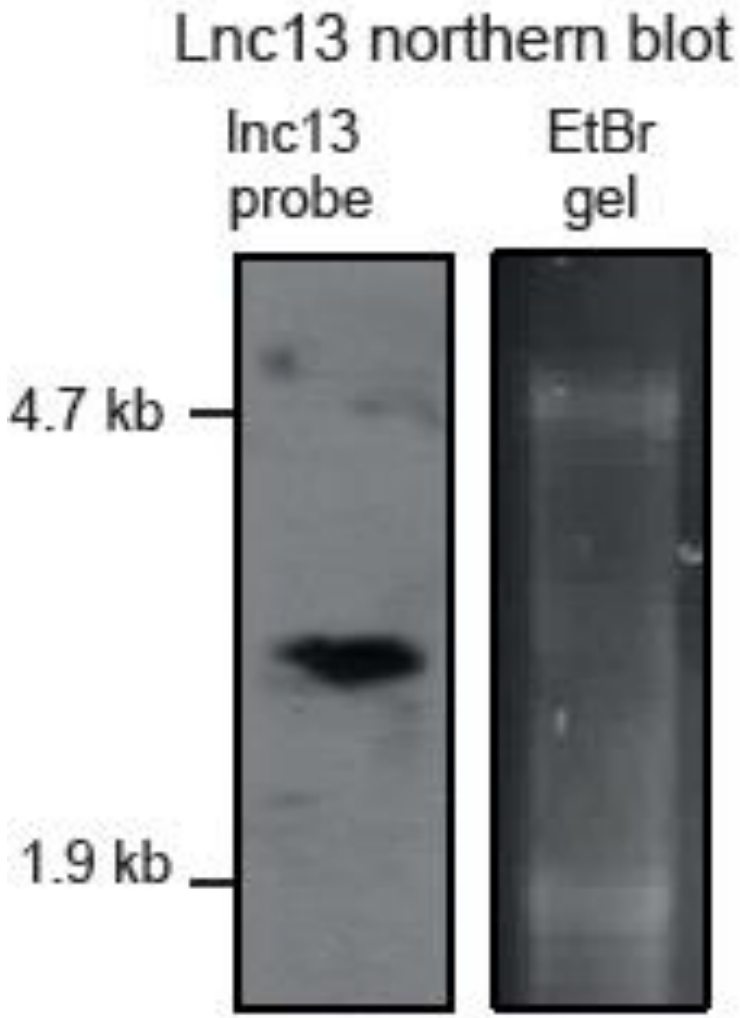


3' RACE blat

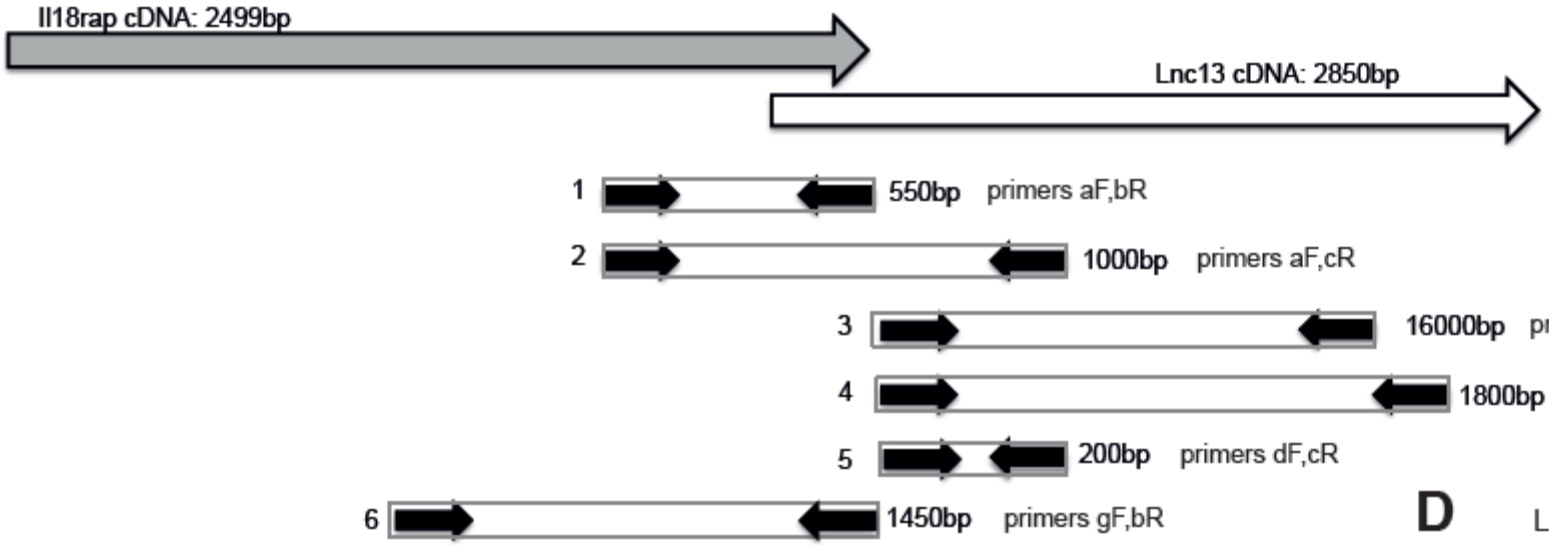


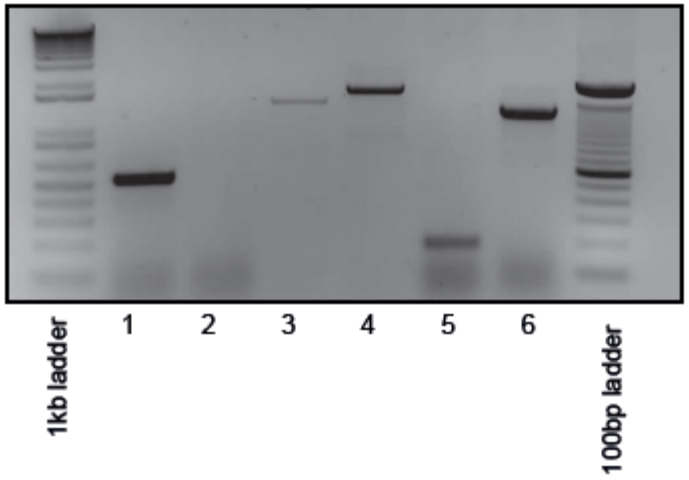
(fig. S1B)

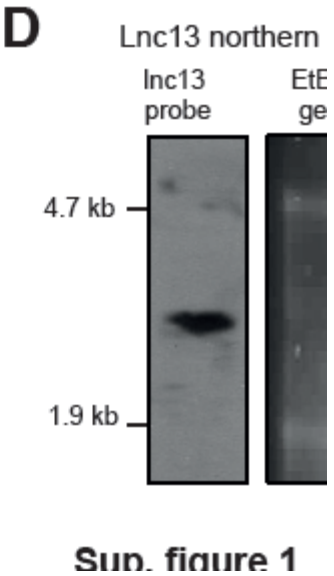
Northern blot of lnc13 expression in iBMM cells



IL18rap and Lnc13 transcripts are indipendent, two different transcripts







Sup. figure 1

fig. S1C

Data from small intestine sample showing the human region corresponding to mouse lnc13 (AK161196).
H3K4me3, H3K4me1, H3K27ac and H3K36me3 peaks denote areas actively transcribed. The chromatin signature points to enhancers (yellow) and actively transcribed chromatin (green). The region encompassing lnc13 is indicated by the line designated as AK161146.

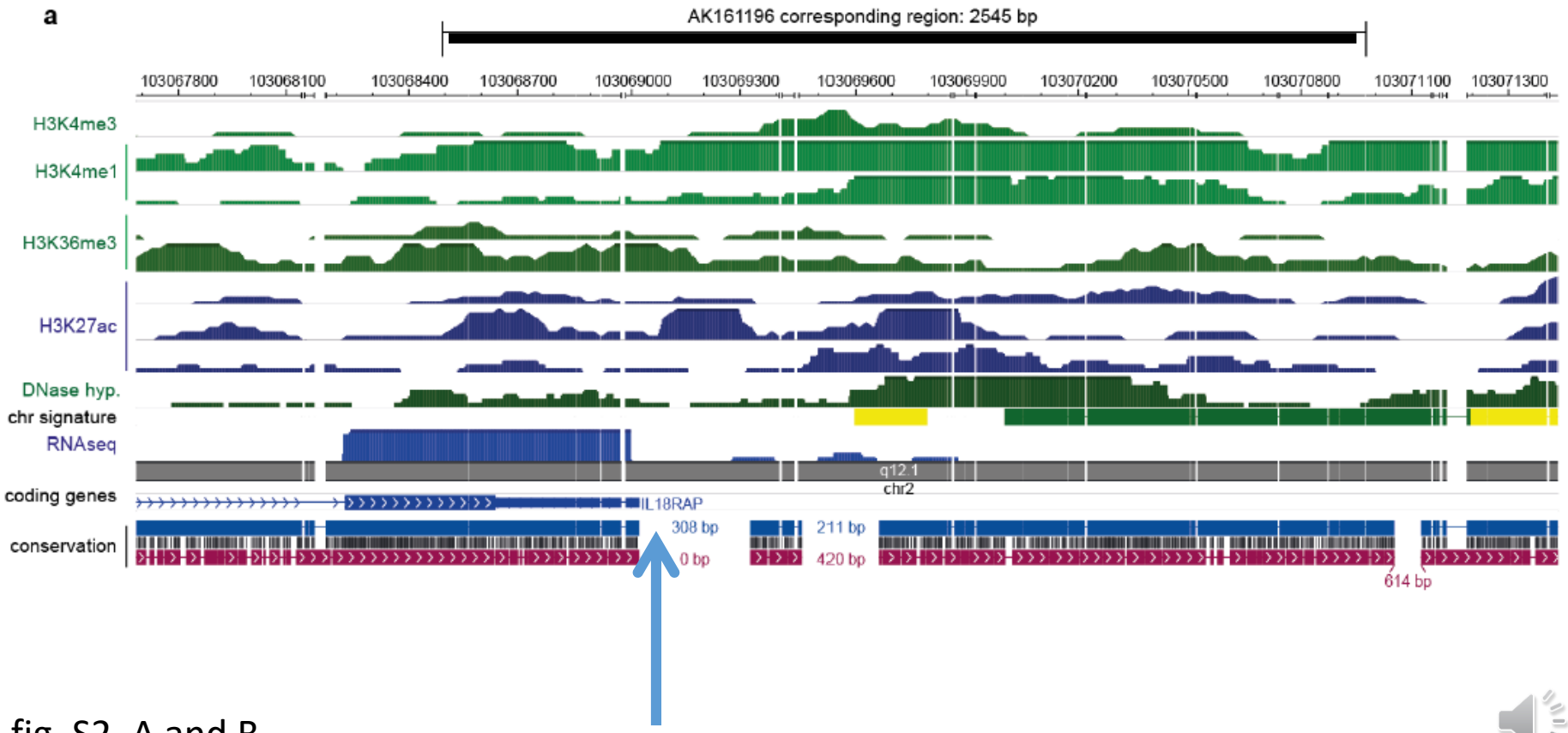
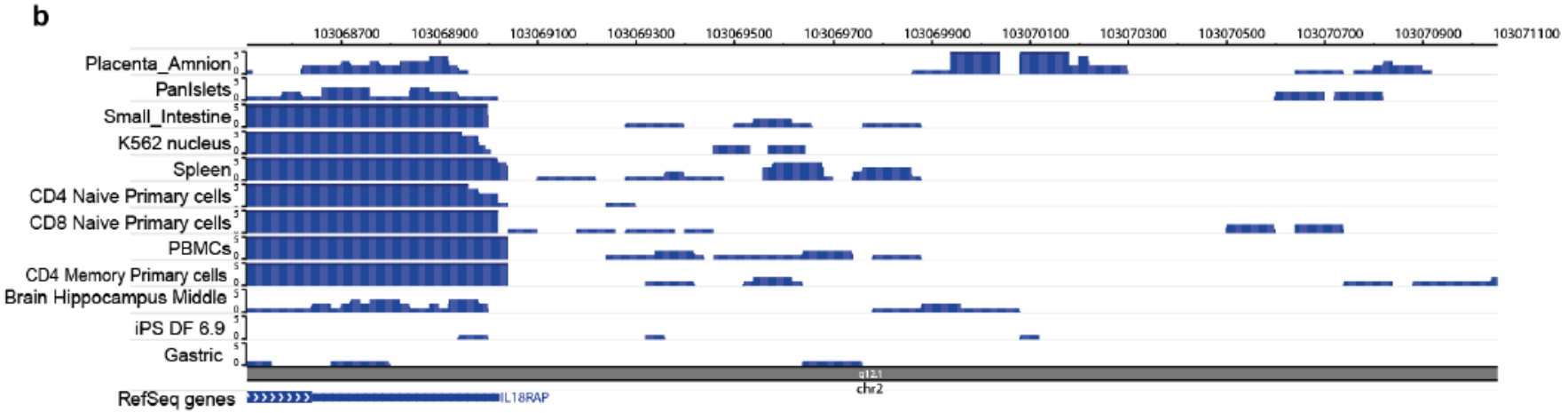
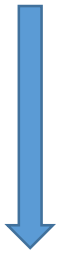


fig. S2, A and B

RNAseq signals in different tissues to the presence of a transcribed region.



The poor **Kozak consensus sequence** in the predicted open reading frame, and the absence of Pfam domains as assessed by the ATG pr and European Molecular Biology Laboratory–European Bioinformatics Institute InterPro prediction programs



lnc13 is unlikely to encode a protein product

Lnc13 expression in human intestinal lamina propria was detected by RNAscope technology.

Biopsies from CeD patients appeared to have substantially lower amounts of lnc13 compared with controls

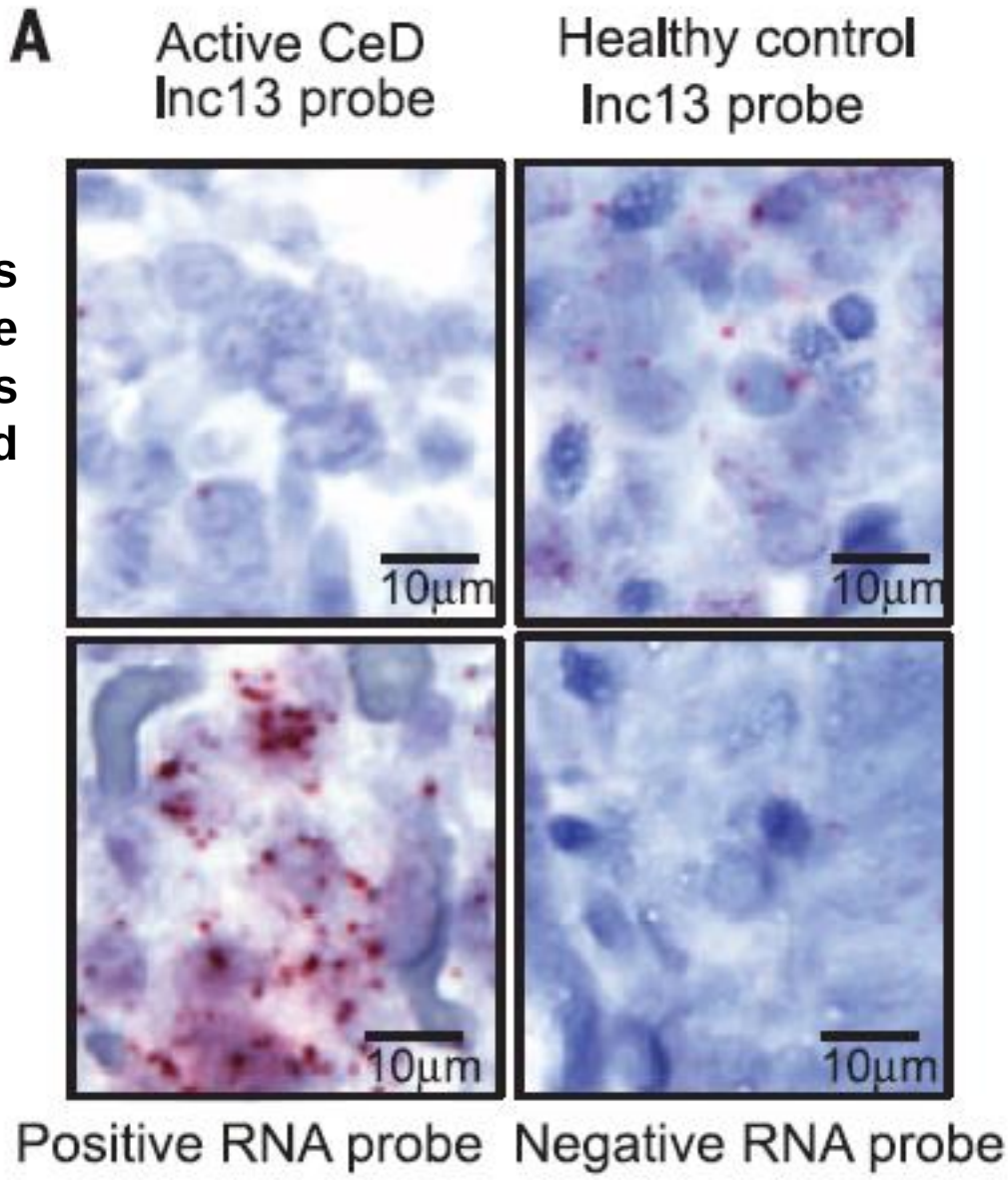


Fig. 1A

Lnc13 expression in human small intestinal biopsies was detected by RNAscope technology.

Biopsies from CeD patients appeared to have substantially lower amounts of lnc13 compared with controls

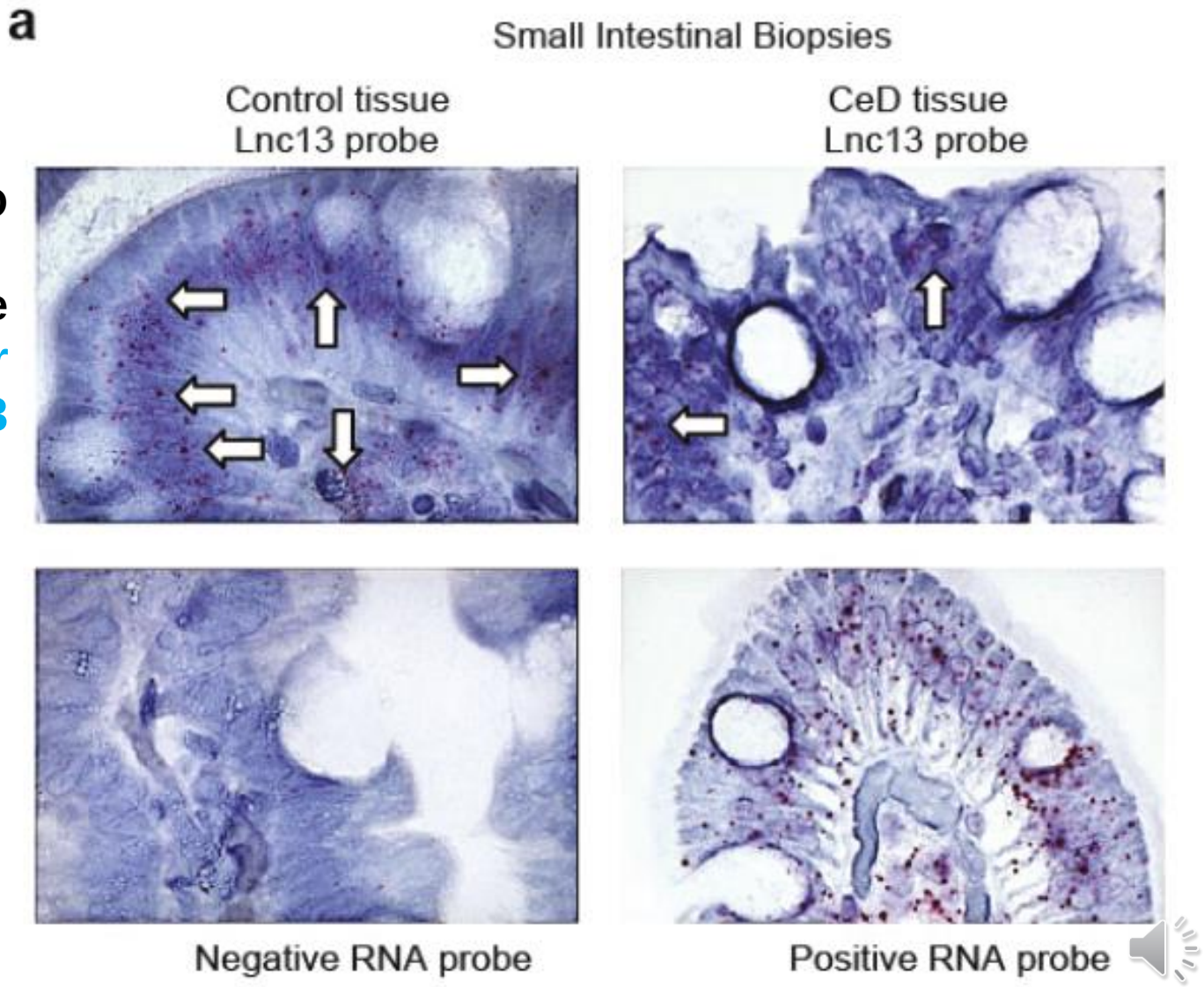


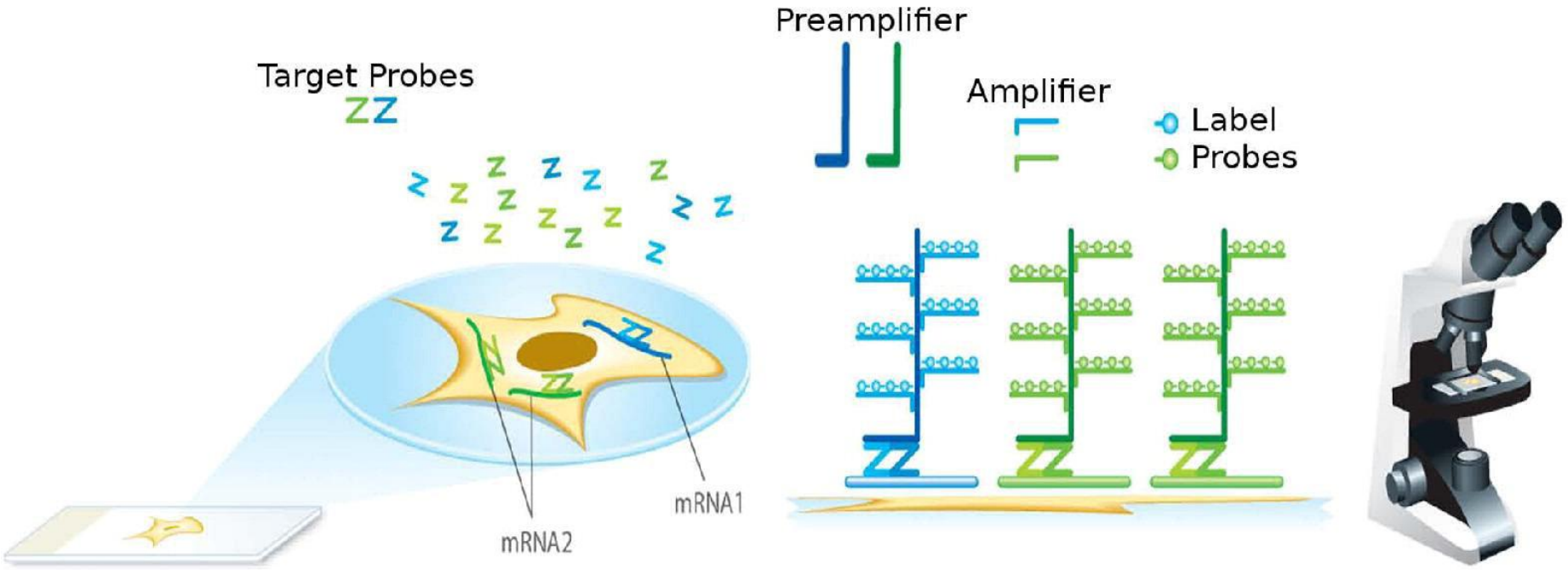
fig. S4A

RNA-SCOPE IS A MODIFIED METHOD OF FISH (IN SITU IBRIDIZATION RNA)

There are **two unlabeled tandem probes**. These probes contain a short complementary region (18-25 bases), a spacer sequence and a 14-base tail sequence.

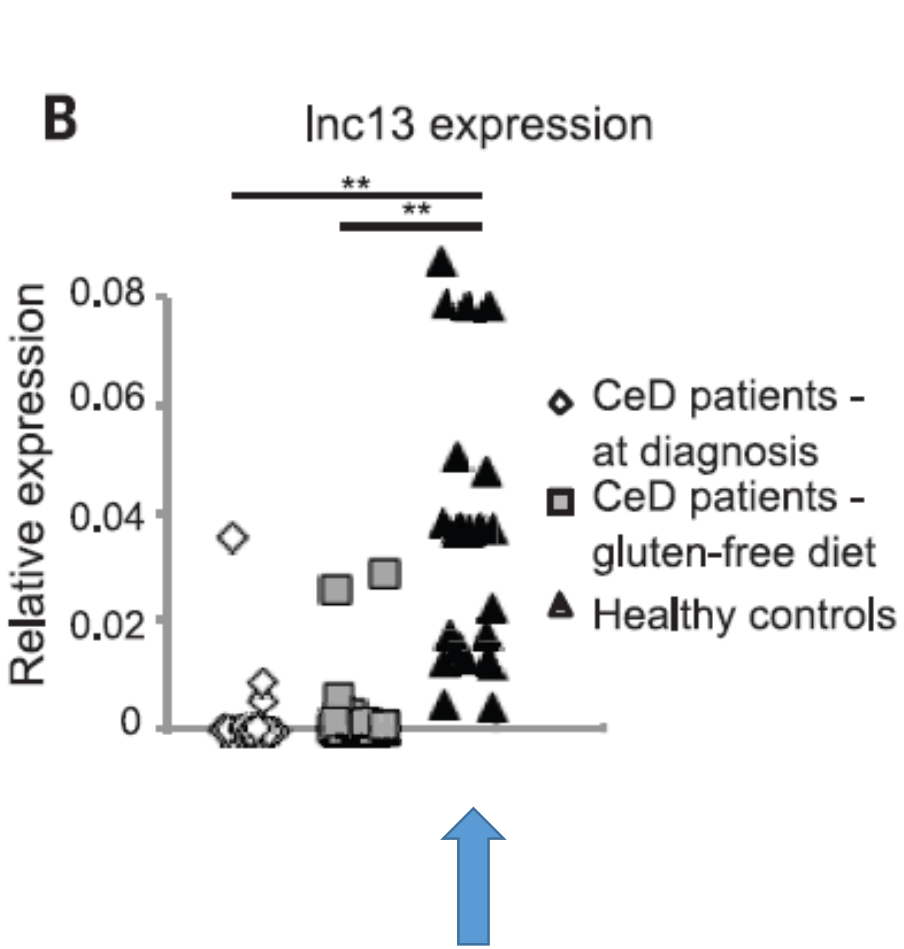
After hybridization with the target probes, comes a **second hybridization step with a pre-amplifier probe**. This is a long probe that contains a complementary sequence to the **28 bases of the two target probes tails** (14+14). So, only when the two target tails hybridize one next to the other the pre-amplifier will hybridize. The **pre-amplifier contains 20 binding sites for an amplifier probe** which in turn contains 20 binding sites for the labeled probe. Thus, for each target probe pair, we get 20×20=400 labeled probes.

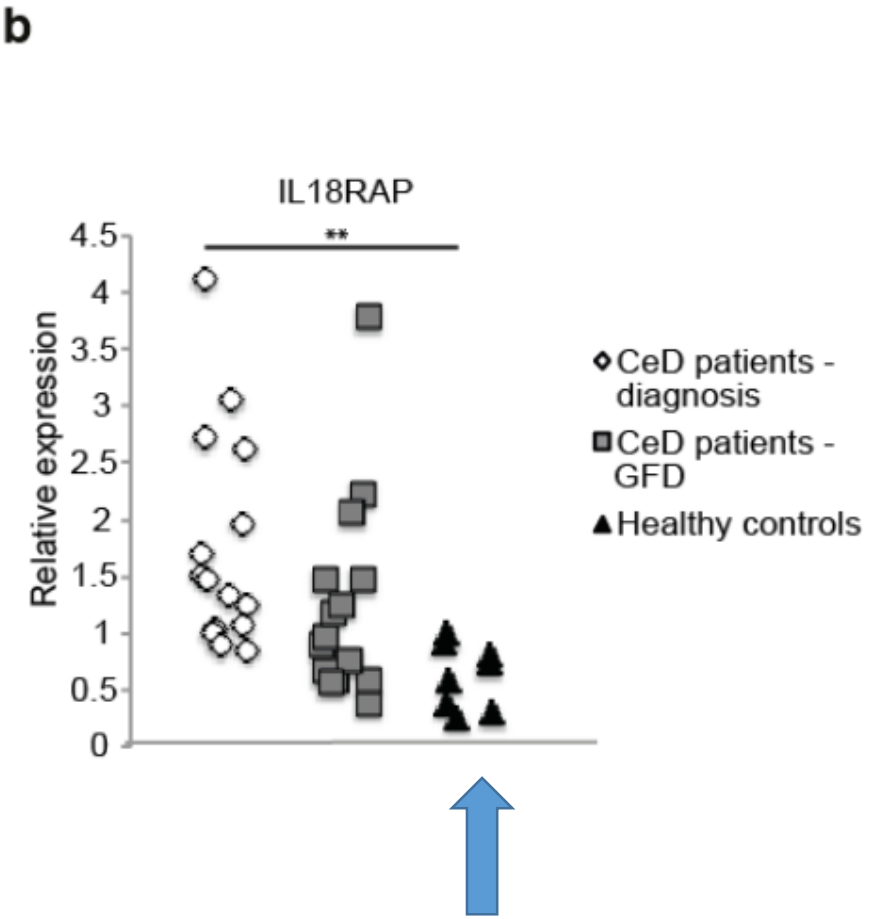
RNA-SCOPE IS A MODIFIED METHOD OF FISH (IN SITU IBRIDIZATION RNA)



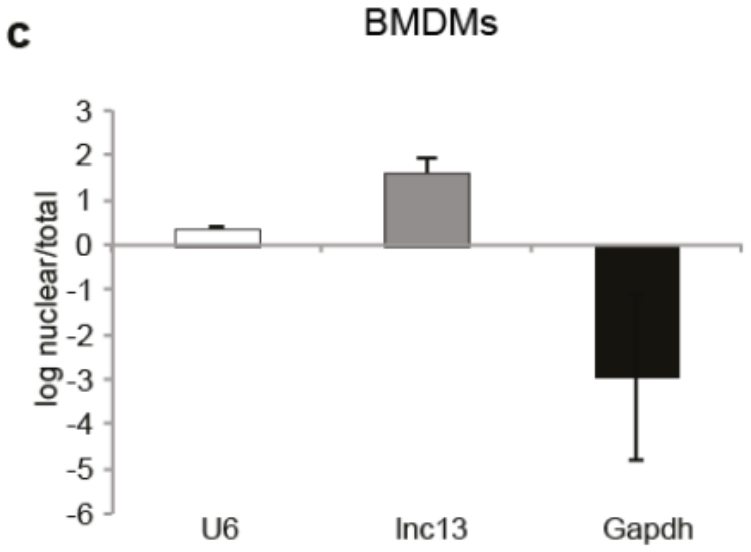
Step 1 Step 2 Step 3 Step 4

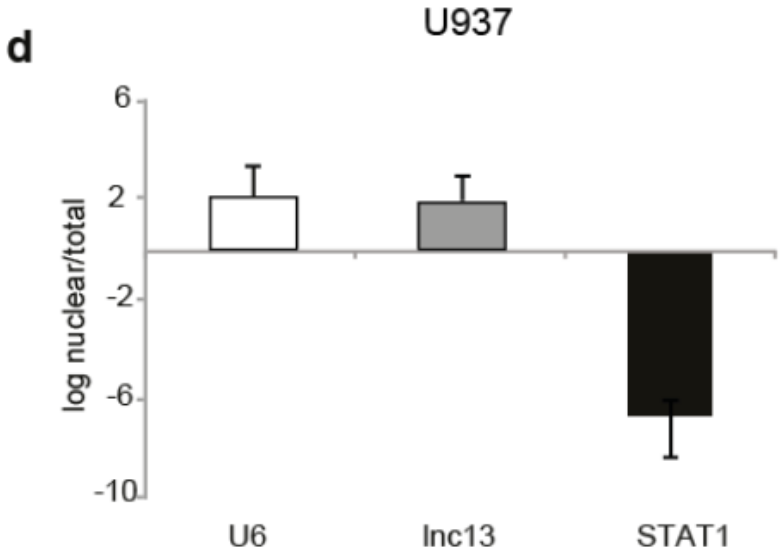
Lnc13 expression is higher in healthy donor respect to Celiac patients, while IL18RAP is low





Subcellular localization of lnc13 is primarily nuclear in both mouse and human macrophages





mouse macrophages

human macrophages

In situ hybridization on small intestinal biopsies showed nuclear localization of lnc13 in the mononuclear cells of the lamina propria

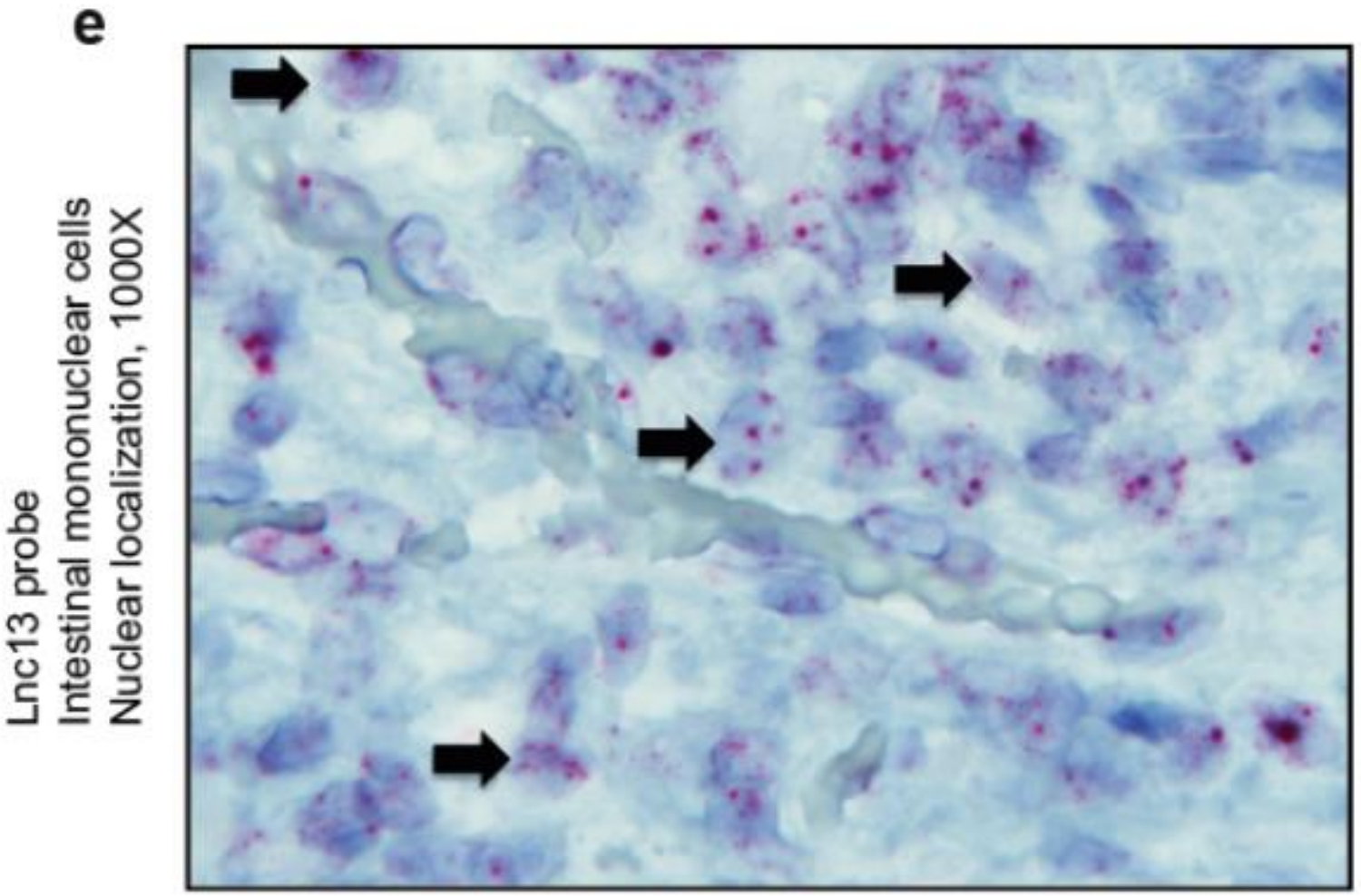
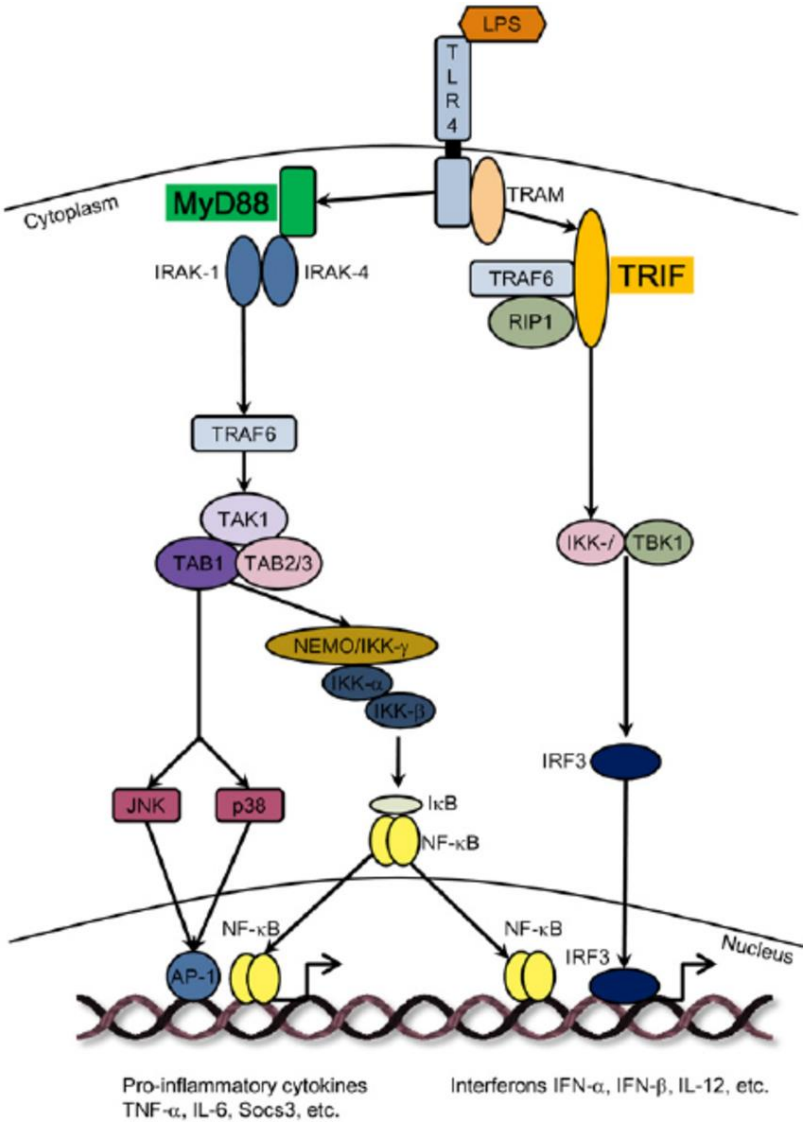


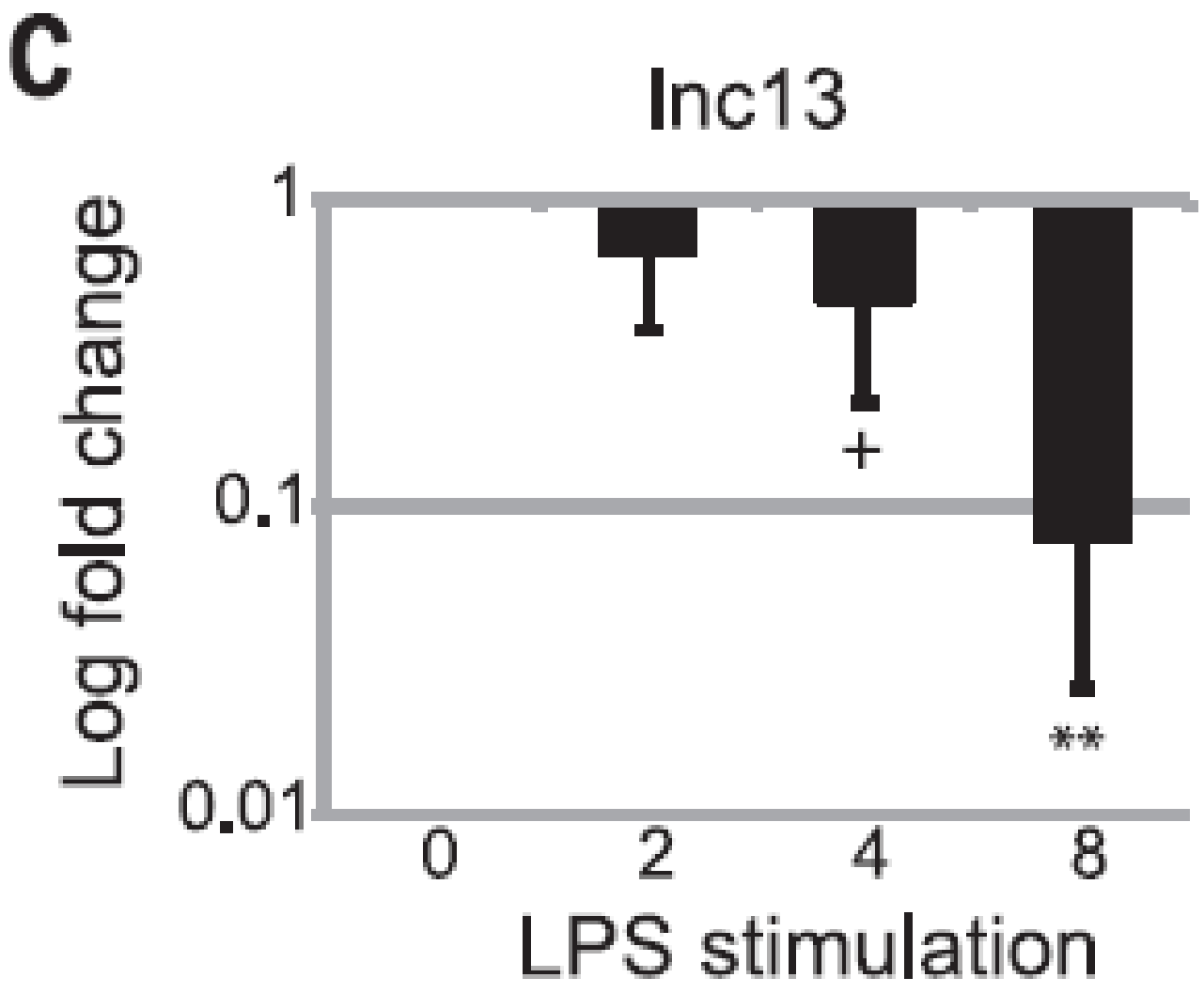
fig. S4E

How lnc13 is modulated in the inflammatory pathway

LPS pathway induces NF-kB activation via MyD88



The level of lnc13 decreased significantly upon LPS stimulation



Myd88 silencing increased lnc13 expression upon LPS treatment, LPS-induced IL12 is inhibited by siMyd88

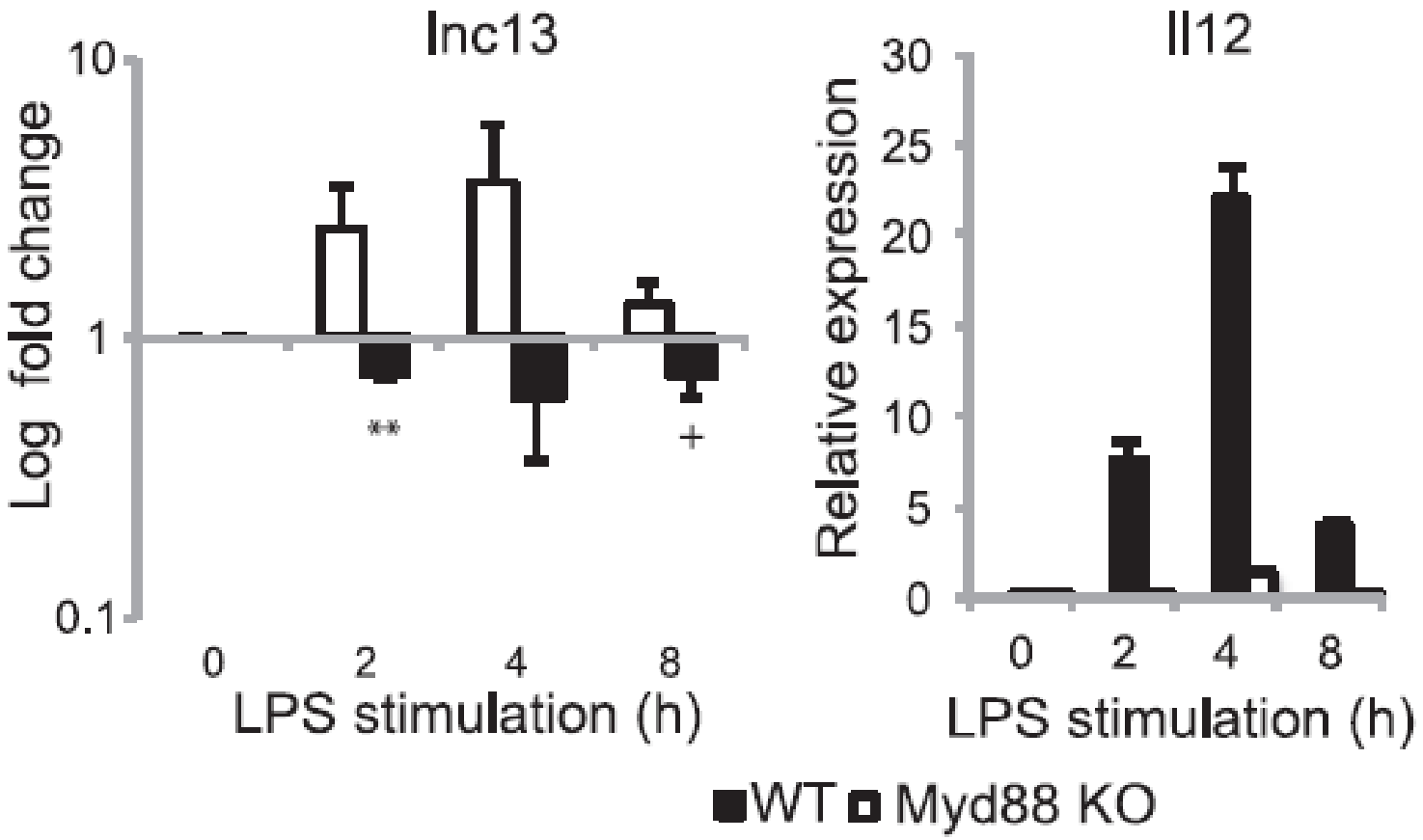
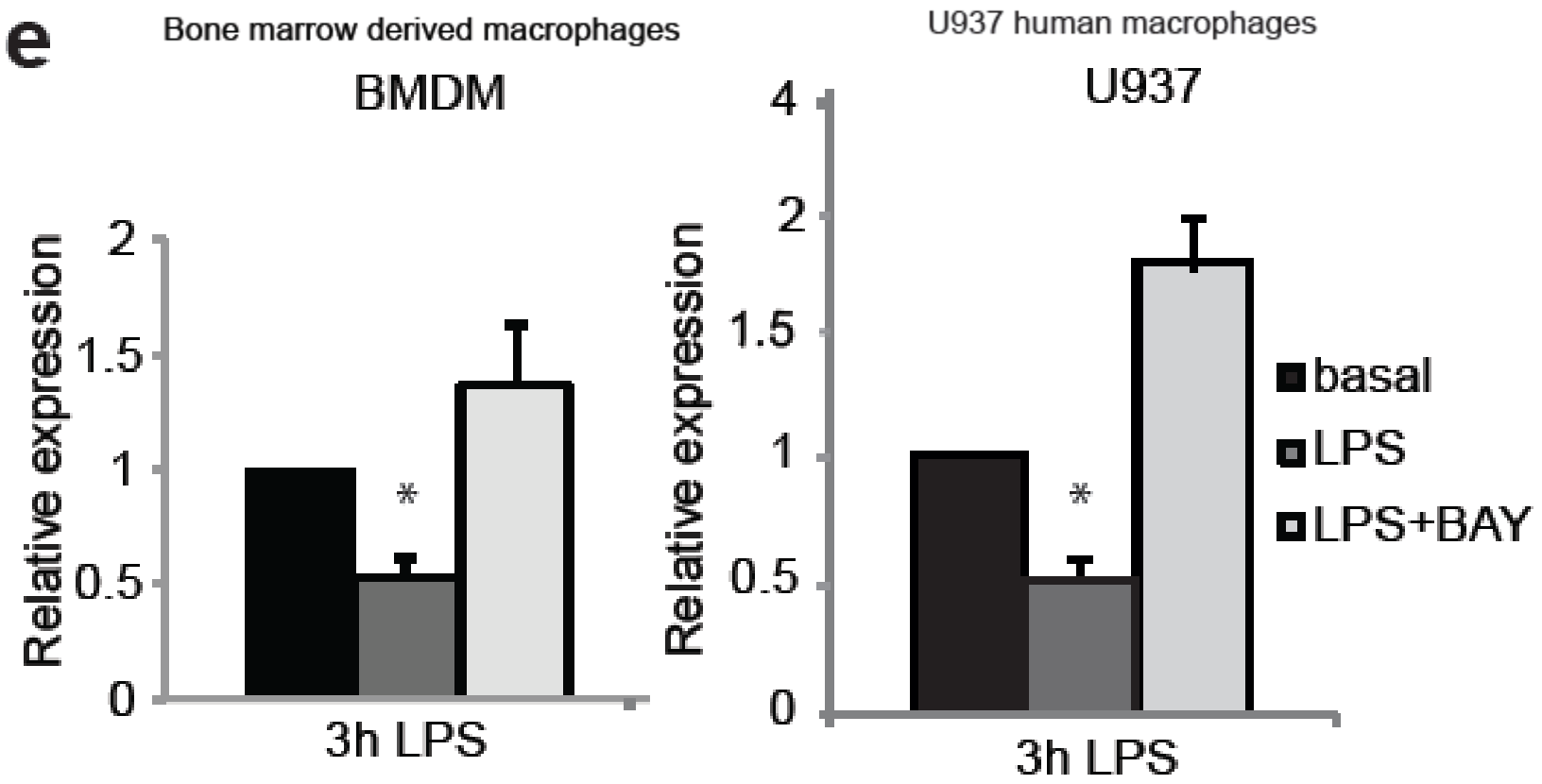


Fig. 1D

LPS induced lnc13 downregulation is blocked after NF-κB inhibition by BAY-11-7082 in mouse and human macrophages.



To identify the Lnc13 TARGET GENES,

gene expression analysis upon LPS stimulation

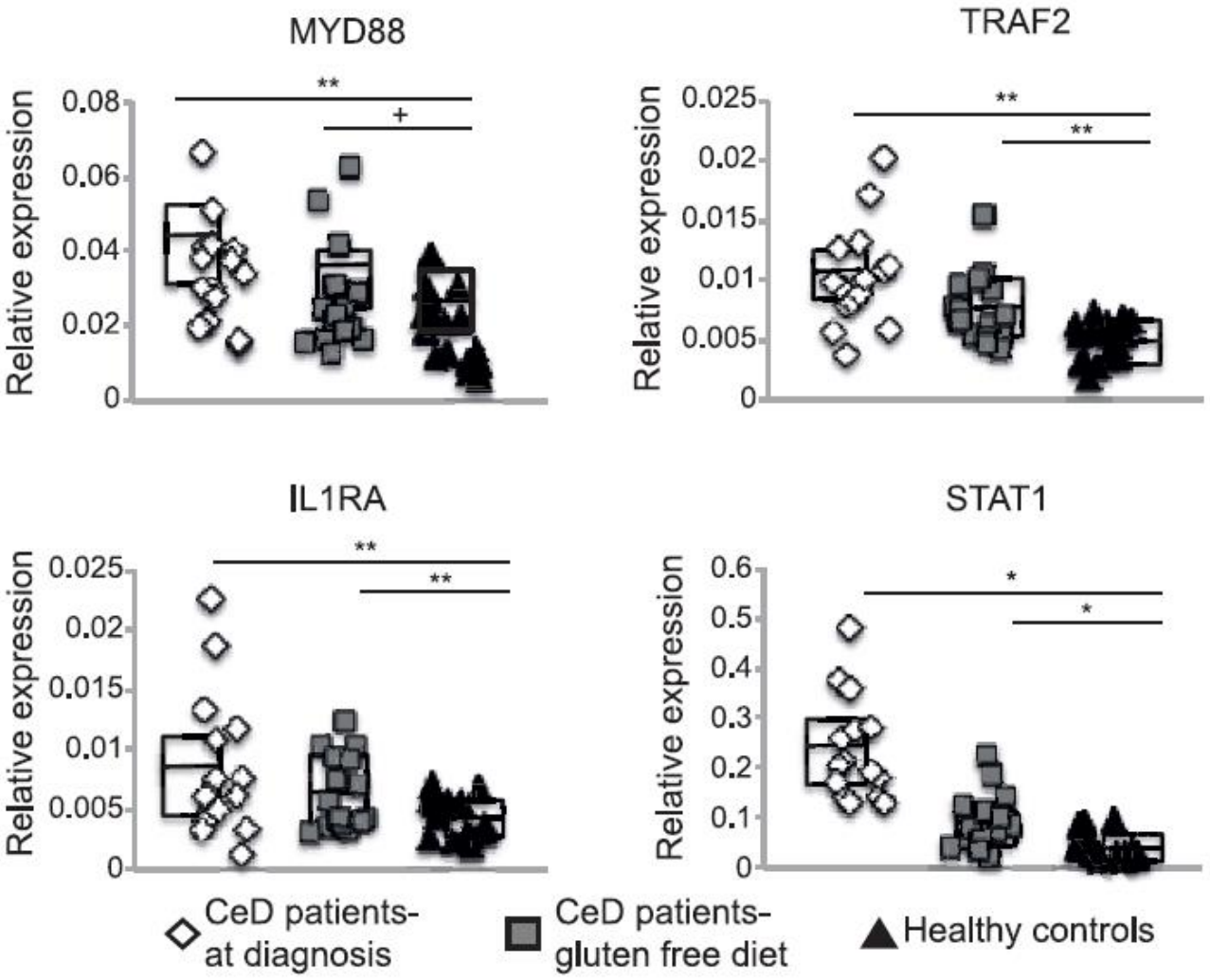
and selection of inflammatory pathway genes

The expression pattern of the Traf2, Stat1, Stat3, Tnfsf10, Il2ra, Ccl12, Myd88, Csf3, and Il1ra genes inversely correlated with changes in lnc13

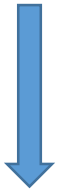
Heat map of the differential expression of a cluster of NF-kB–regulated genes that exhibit expression kinetics highly correlated with lnc13 expression kinetics in response to LPS stimulation (correlation coefficient R > 0.8) of BMDMs.

Fig. 2A

Expression of four of these genes (TRAF2, STAT1, IL1RA, and MYD88) was significantly increased in biopsies from CeD patients



Lnc13 affects the expression of genes involved in inflammatory pathway TRAF2, STAT1, IL1RA, and MYD88



In order to demonstrate that Lnc13 is a modulator of genes TRAF2, STAT1, IL1RA, and MYD88 is used two approaches:

a) **Overexpression**
b) **Downregulation**

Increased levels of lnc13 reduced the expression of the predicted targets

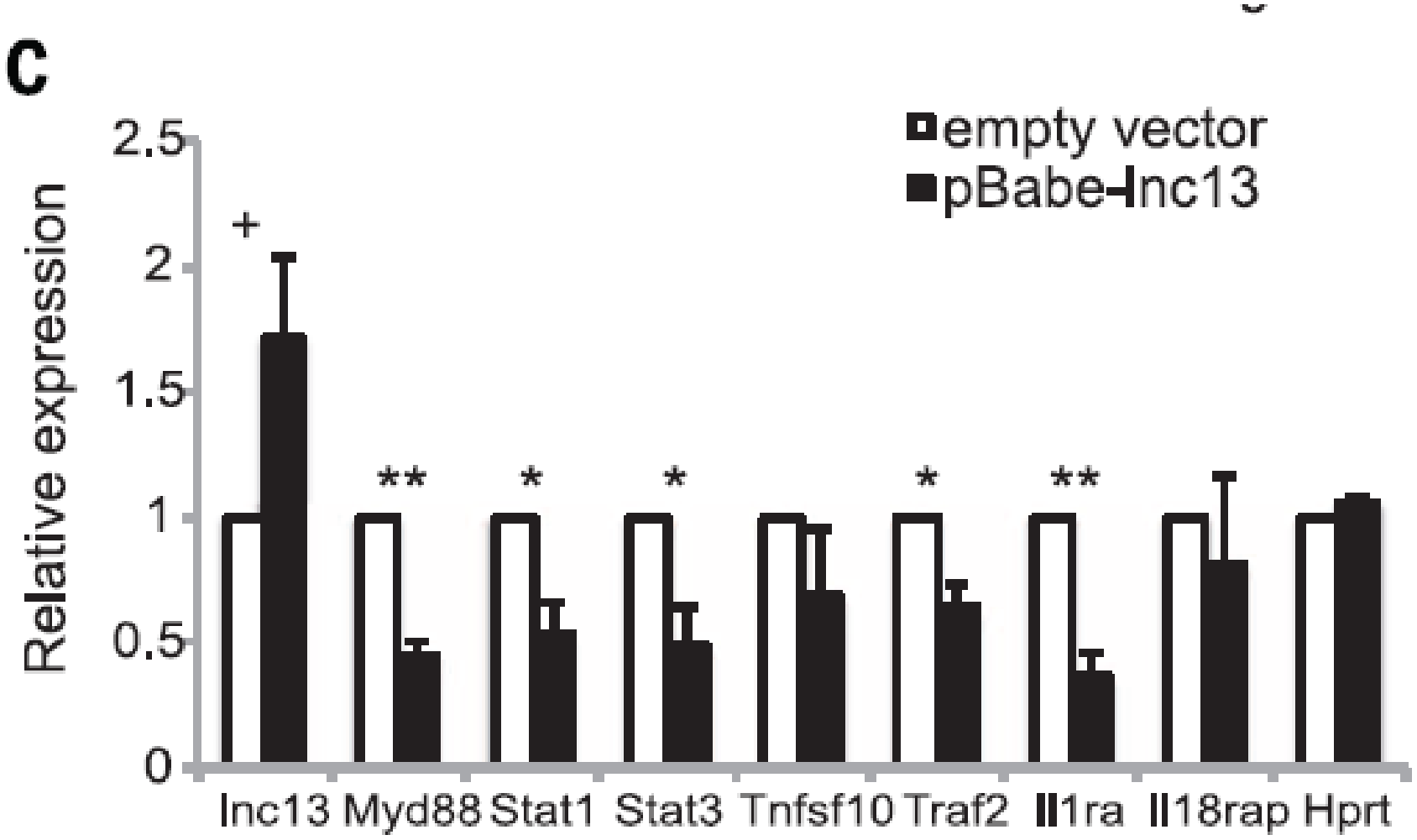


Fig. 2C

Silencing of lnc13 increased the expression of the predicted targets

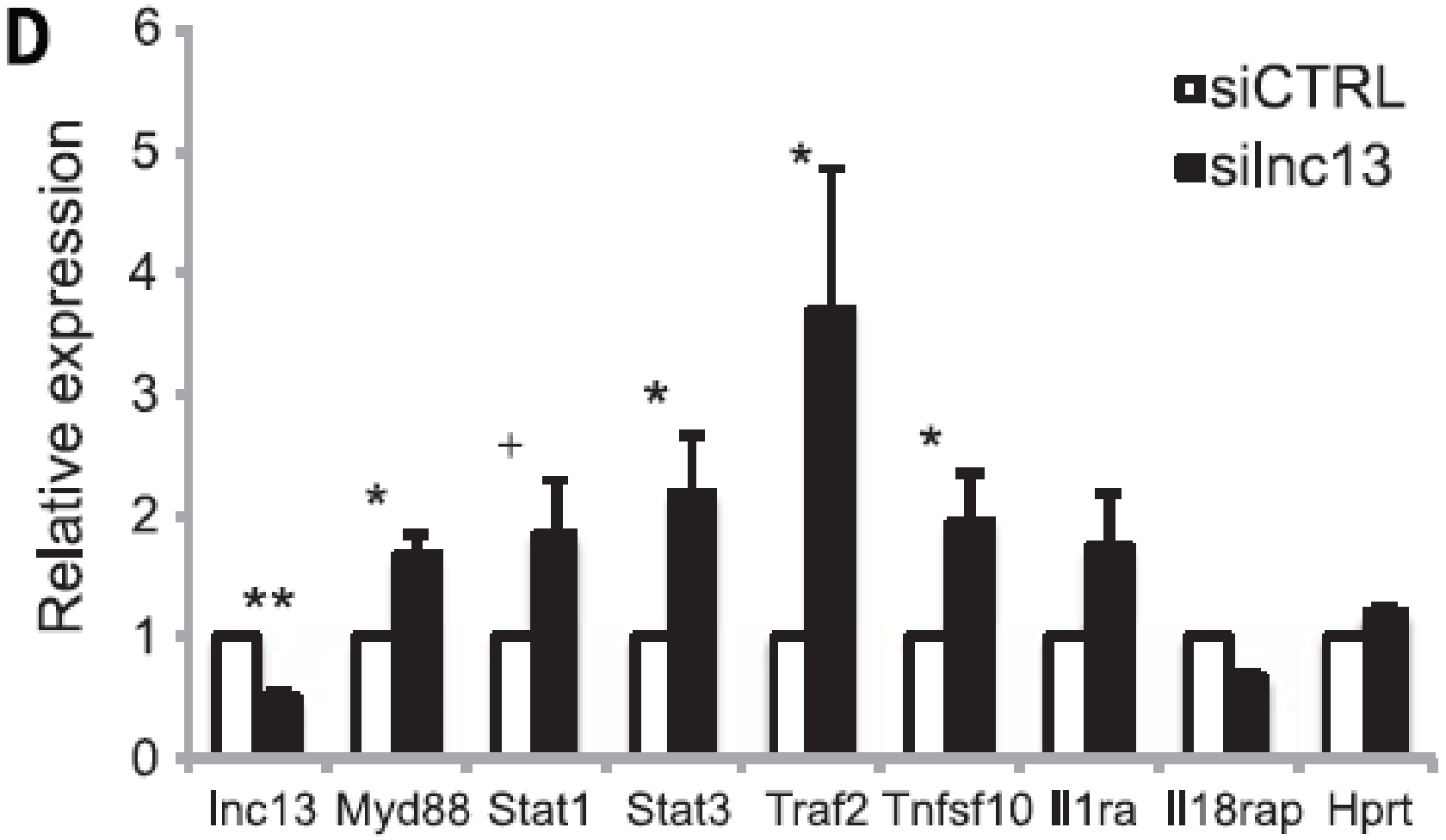


Fig. 2D

IN SUMMARY

- **Lnc13 is a long non coding RNA associated with Celiac Disease**

- **Lnc13 decreased upon inflammatory pathway activation**