HHS Public Access

Author manuscript

Sci Signal. Author manuscript; available in PMC 2021 June 08.

Published in final edited form as:

Sci Signal. ; 13(661): . doi:10.1126/scisignal.aba3244.

A small sustained increase in NOD1 abundance promotes ligand-independent inflammatory and oncogene transcriptional responses

Leah M. Rommereim[1,5,†], **Ajay Suresh Akhade**[1,†], **Bhaskar Dutta**[2,†], **Carolyn Hutcheon**[1,6], **Nicolas W. Lounsbury**[2,7], **Clifford C. Rostomily**[1], **Ram Savan**[3], **Iain D. C. Fraser**[2], **Ronald N. Germain**[2], **Naeha Subramanian**[1,3,4,*]

[1]Institute for Systems Biology, Seattle, WA, 98109, USA.

[2]Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, 20892-0421, USA.

[3]Department of Immunology, University of Washington, Seattle, WA, 98109, USA.

[4]Department of Global Health, University of Washington, Seattle, WA, 98109, USA.

[5]Current address: SEngine Precision Medicine, Seattle, WA, 98109, USA.

[6]Current address: Resolution Bioscience, Kirkland, WA, 98033, USA.

[7]Current address: Genentech, South San Francisco, CA, 94080, USA.

Abstract

Small, genetically determined differences in transcription (expression quantitative trait loci or eQTLs) are implicated in complex diseases through unknown molecular mechanisms. Here, we showed that a small, persistent increase in the abundance of the innate pathogen sensor NOD1 precipitated large changes in the transcriptional state of monocytes. A ~1.2–1.3 fold increase in NOD1 protein abundance resulting from loss of regulation by the microRNA cluster miR-15b/16 lowered the threshold for ligand-induced activation of the transcription factor NF-κB and the MAPK p38. An additional sustained increase in NOD1 abundance to 1.5-fold over basal amounts bypassed this low ligand concentration requirement, resulting in robust ligand-independent induction of proinflammatory genes and oncogenes. These findings reveal that tight regulation of NOD1 abundance prevents this sensor from exceeding a physiological switching checkpoint that promotes persistent inflammation and oncogene expression. Furthermore, our data provide insight into how a quantitatively small change in protein abundance can produce marked changes in cell state that can serve as the initiator of disease.

[*]Corresponding author: Office Ph: 206-732-1226, Cell: 240-205-6012, FAX: 206-732-1260, nsubrama@systemsbiology.org.
[†]Co-first authors

Author contributions: B.D., R.N.G and N.S. conceived the study. R.N.G and N.S. supervised the study and acquired funding. L.M.R, A.S.A. and N.S. designed and performed experiments. L.M.R, A.S.A. and N.S. analyzed the data. B.D., N.W.L., C.H. and C.C.R. performed informatic analysis of microarray and GEO data. L.M.R, A.S.A., C.H., C.C.R. and N.S. prepared data for publication. I.D.C.F provided reagents and input for construction of stable cell lines. R.S. advised miRNA target site mutagenesis and luciferase reporter assays. L.M.R. and N.S. wrote the manuscript. R.N.G. and N.S. edited the manuscript. All authors reviewed, revised or commented on the manuscript.

Competing interests: L. R. is an employee and stockholder of SEngine Precision Medicine. B. D. is an employee and stockholder of AstraZeneca. The other authors declare that they have no competing interests.

Introduction

The innate immune system employs a network of pattern recognition receptors to sense pathogens or danger signals and mount an inflammatory response for host defense. Although this ability to detect and respond to microbes or tissue damage is essential to ward off infections and trigger wound healing programs, unrestrained inflammation can lead to various diseases. Chronic low-level stimulation of the immune system can enforce mechanisms that lead to exaggerated reactivation responses (1) and cause uncontrolled tissue inflammation in genetically predisposed individuals, resulting in autoimmunity, autoinflammatory diseases, and cancers (2, 3).

Several mechanisms may underlie or predispose to chronic inflammation, and their contributions are yet to be resolved. In many cases, complex immune-related diseases are linked in genome-wide association studies (GWAS) to causal variants located primarily in non-coding regions of genes, generating eQTLs (expression quantitative trait loci). The extent of expression variation between susceptible and resistant genotypes is often in the 1.5–3 fold range, suggesting that small changes in gene or gene product expression might play an important role in functional immune dysregulation (4). Although changes in protein concentration or activity due to defects in single genes have been implicated in haploinsufficiencies and monogenic autoinflammatory or neurodegenerative syndromes (5), little information exists about whether modest changes in protein concentration of the type expected with cis- or trans-eQTL effects can lead to pathological changes in cell behavior that may promote eventual tissue or organ dysfunction. This latter question has special relevance for inflammation, which if sustained over time causes tissue damage and can also promote oncogenesis, and for the latter, facilitate the transformation process itself (6). Indeed, disease-promoting genomic modifications cluster upstream of functional master regulator proteins whose abnormal activity is necessary and sufficient for propagating a tumor cell state (7). The master regulator activity can be controlled post-transcriptionally, and may be dysregulated, for example, by modulation of microRNA (miRNA) activity (8, 9) suggesting that small changes in protein expression of master regulators may induce cellular transformation to a diseased state.

NOD1 is a ubiquitously expressed intracellular innate sensor of microbial infection that senses mesodiaminopimelic acid (iE-DAP, a component of bacterial peptidoglycan) (10, 11) and pathogen-induced alterations in cell state (12, 13) to trigger the induction of pro-inflammatory genes through nuclear factor-κB (NF-κB) and mitogen-activated protein kinase (MAPK). NOD1 activity is also intimately linked to gastric cancer. In some studies, genetic variants in *NOD1* are associated with gastric cancer risk and NOD1 expression is increased in gastric tumors (14, 15). In addition, chronic inflammation triggered by activation of NOD1 by *Helicobacter pylori* is an initiating event in gastric cancer (16). Pro-inflammatory cytokines including members of the TNF and IL-1 family contribute to tumor development in the gastric mucosa by promoting cell proliferation (17), development of an immunosuppressive microenvironment (18) and immune escape (19), suggesting that chronic inflammation can promote oncogenesis by multiple mechanisms.

Maintaining expression of innate sensors near a switching point to ensure a rapid and sensitive response to early signs of infection makes sense in evolutionary terms, but carries with it a requirement for tight regulation of the sensor signaling system and the risk of inadvertent activation if such regulation is disturbed. During a study of NOD1 signaling, we discovered that small but persistent increases in NOD1 expression spontaneously lead to large-scale gene activation. A big-data approach further revealed that *NOD1* is the most tightly regulated of those genes encoding innate sensors and suggested a physiological necessity to keep NOD1 expression under stringent control. Here we showed that a prolonged small (~1.5 fold) increase in NOD1 concentration within cells upregulated inflammatory genes and oncogenes in the absence of ligand exposure, and identified a miRNA-based circuit for stringent control of NOD1 expression in human monocytes. Our data highlight one way by which small changes in gene expression may impact chronic inflammation. These findings also suggest a mechanism that could promote cell-intrinsic oncogenic activity, suggesting that these observations have broad implications for understanding how small expression changes caused by eQTLs may shape the development of complex diseases like autoimmunity and cancer.

Results

A persistent small (1.5 fold) increase in NOD1 leads to large-scale gene activation

With an initial aim of investigating the quantitative relationship between NOD1 expression and its effects on cell state, we established a stable, lentiviral expression system in which 3X-FLAG-tagged *NOD1* coding sequence was transcribed under the control of an inducible tetracycline promoter (TRE; tetracycline response element) in THP-1, a human monocyte cell line (THP-1 NOD1 cells) (Fig. 1A). As controls, cells transduced with 3X-FLAG NLRP4 (THP-1 NLRP4 cells) or empty lentiviral vector (THP-1 Vector cells) were derived. NLR expression was induced with doxycycline (DOX, an analog of tetracycline) and changes in cellular transcriptome were analyzed by microarray. As expected, addition of DOX led to robust upregulation of *NOD1* and *NLRP4* expression (Fig. 1B). The microarray data were further subjected to pathway analysis using GAGE (Generally Applicable Gene-set Enrichment) (20), with pathway information derived from KEGG (Kyoto Encyclopedia of Genes and Genomes) (21). Multiple pairwise comparisons were conducted to exclude any effects of DOX (22) or the lentiviral expression vector (fig. S1). However, even in the absence of DOX, THP-1 NOD1 cells upregulated the expression of genes in a large number of cellular pathways similar to that seen upon DOX-induced strong over-expression of *NOD1* in the absence of ligand (Fig. 1C and Data File S1). We define these changes as “maximal” or “saturating” in the remainder of the paper. This behavior was not observed with THP-1 NLRP4 cells or THP1 vector-transduced cells treated with or without DOX, suggesting that it was not a general feature of NLRs, DOX treatment, or transduction of cells with the lentiviral vector per se. Even in the absence of DOX, there was a ~1.5 fold increase in *NOD1* and ~2-fold increase in *NLRP4* mRNA in the NLR-transduced cells, as compared to their endogenous levels in the parent cells or the control lentivirus-transduced cells (Fig. 1B). Although we did not anticipate this result at the conception of the study, this small increase in NLR expression was consistent with previous reports showing that even in the absence of tetracycline the reverse Tet transactivator (rtTA3) binds weakly to the TRE

promoter leading to a low-level of background activity (22). Quantitative RT-PCR using primers specific for endogenous *NOD1* or 3X-FLAG *NOD1* (table S1) confirmed that THP-1 NOD1 cells showed a small (0.5-fold) increase in expression of TRE-driven *NOD1* mRNA compared to THP-1 vector-transduced cells in the absence of DOX without affecting endogenous *NOD1* (Fig. 1D). Low-level expression of 3X-FLAG NOD1 in the absence of DOX was also observed at the protein level by immunoblot at high exposures and by flow cytometry (Fig. 1E–F). Our flow cytometry data (Fig. 1F) showed that the small increase in NOD1 expression was distributed across the population, suggesting that the gene expression effects were not due to 'jack-potting' with very high expression of NOD1 in only some of the THP-1 cells and emphasizing that it was the modest 1.5-fold increase in NOD1 concentration that underlies the response phenomenon. Correlation analysis revealed a high concordance in genes differentially expressed upon low-level (that is, in the absence of DOX) and DOX-induced expression of NOD1 (r^2=0.9) but not in genes differentially expressed between low-level and DOX-induced expression of NLRP4 (r^2=0.01) (Fig. 1G). These data suggest that a persistent small (1.5-fold) increase in expression of NOD1 leads to saturating gene activation in a ligand-independent, switch-like manner.

Several pro-inflammatory genes and negative feedback regulators associated with ligand-induced activation of NOD1 including *IL1B*, *JUN*, *NFKB1*, *NFKBIA*/*IκBα*, and *TNFAIP3*/*A20* (23) were upregulated in cells that had sustained 1.5-fold increase in NOD1 expression (Data File S2 and fig. S2A). This finding suggests that ligand-free activation of NOD1 by 1.5-fold over-expression triggers the expression of classical ligand-dependent inflammatory genes. Furthermore, native gel analysis showed that persistent 1.5-fold over-expression of NOD1 induced its oligomerization similar to that seen upon ligand treatment of cells with normal levels of NOD1 (fig. S2B). Together, these data indicate that NOD1 may be activated in a switch-like manner whereby a persistent small increase in its expression at or above a 1.5-fold threshold of expression mimics full activation.

A prolonged small increase in NOD1 also triggers an oncogene transcriptional response

In addition to pro-inflammatory genes, several genes including proto-oncogenes (*ALX1*, *C-KIT*, *CAV1*, *CNN2*, *CD68* and *GPX8*) not conventionally associated with NOD1 activation were highly correlated with *NOD1* expression (r^2>0.7, fold change≥4) (Fig. 1G and fig. S3A) (24–27). This finding indicated that a sustained small increase in NOD1 could lead to gene activation related to cell transformation processes, in addition to the elaboration of pro-inflammatory mediators. Quantitative RT-PCR, immunoblot, and flow cytometry based analysis showed that low-level persistent expression of NOD1 was sufficient to induce maximal expression of the oncogenes *ALX1* and *C-KIT*, similar to that observed upon DOX-induced upregulation of NOD1 (Fig. 2A–C). Similar to NOD1, the increase in ALX1 and C-KIT was distributed across the cell population (Fig. 2C) indicating that effects on gene expression were not due to high expression in only a subset of cells, and that it was the small increase in NOD1 protein concentration that achieved a maximal increase in these oncogenic proteins. To determine if a small increase in NOD1 also activated other mediators of oncogenesis, we analyzed the expression of C-MYC and activation (that is, phosphorylation) of AKT which are considered to be major drivers in the pathogenesis of many cancers (28, 29). Phosphorylation of AKT and expression of C-MYC were both

increased in cells that had a persistent small increase in NOD1 (Fig. 2D). Trace peptidoglycan contaminants in serum can contribute to NOD1-dependent signaling in cell culture work (30). However, we observed similar NOD1-dependent expression of oncogenes and inflammatory genes (fig. S3B–C) and similar degradation of IκBα regardless of the presence or absence of serum (fig. S3D), indicating that effects on gene expression observed in response to low-level NOD1 over-expression were independent of trace levels of peptidoglycan present in fetal calf serum. An exception was *C-KIT* whose expression was induced to a much lower but significant level in serum-free medium compared to serum-containing medium, in line with previous work showing that serum factors can promote *C-KIT* expression (31). Collectively, these data suggest that a sustained 1.5-fold increase in NOD1 can upregulate oncogenic processes independent of ligand triggering.

We next asked if long-term ligand stimulation could mimic the effect of a prolonged 1.5-fold increase in NOD1 in causing pathological gene induction. We stimulated THP-1 cells repeatedly with ligand and monitored the expression of *JUN* as a classical ligand-responsive gene and *C-MYC* as a representative oncogene. As expected, ligand treatment led to acute upregulation of *JUN* expression as early as 1 h followed by a gradual tolerization by 48–72 h, with decreasing but substantial levels of *JUN* induction in response to recurrent ligand stimulation (Fig. 2E). A similar trend was observed with the other acute ligand-responsive genes, *IL1B* and *TNFAIP3* (which encodes A20) (fig. S4). The tolerization effect was also mirrored in cells with persistent 1.5-fold over-expression of NOD1, which showed significantly elevated but lower abundance of these inflammatory gene transcripts compared to that observed at peak ligand stimulation (fig. S2A). Kinetically, upregulation of C-*MYC* after recurrent ligand stimulation always followed peak *JUN* induction and showed a delayed surge between 52–60 h (Fig. 2E). These data suggest that prolonged ligand activation of NOD1 can lead to oncogene induction. The later surge in *C-MYC* may reflect a time delay cell-intrinsic switch to *C-MYC* induction, or alternatively may suggest a mechanism whereby chronic inflammation leads to pathologic gene activation.

We next sought to determine if persistent 1.5-fold over-expression of NOD1 was necessary and sufficient for the upregulation of pathologic transcriptional responses. We specifically ablated TRE-driven 3X-FLAG *NOD1* over-expression in THP-1 NOD1 cells using CRISPR-Cas9 based gene editing while leaving expression of endogenous *NOD1* intact (effectively restoring NOD1 expression in these cells from 1.5-fold to basal values). We validated CRISPR targeting in three independent single cell clones by sequencing and observed a unique indel in each clone that resulted in early termination of FLAG-*NOD1* (fig. S5A–B). Ablation of FLAG-NOD1 reduced the expression of *ALX1*, *C-KIT*, and C-MYC and the phosphorylation of AKT (Fig. 2F–H), indicating that the sustained low-level increase in NOD1 expression caused the upregulation of these genes and of AKT activity. Inflammatory NF-κB and MAPK signaling in response to NOD1 activation requires the adaptor protein RIPK2. To determine whether RIPK2 was required for the abnormal transcriptional signature observed in response to a sustained 1.5-fold increase in NOD1, we deleted *RIPK2* in THP-1 NOD1 cells by gene editing and derived three independent single cell clones with unique indels in *RIPK2* that disrupted gene expression (fig. S5C–D). Deletion of RIPK2 reduced the expression of oncogenes to near-baseline levels (Fig. 2I) indicating that

signaling through RIPK2 is responsible for pathologic gene induction in cells with a persistent 1.5-fold increase in NOD1.

NOD1 is the most tightly regulated human innate sensor

The finding that a sustained small increase in NOD1 abundance could lead to disproportionately large increases in gene expression prompted us to compare the regulation of *NOD1* expression to that of other innate immune sensors. We analyzed publicly available gene expression data from >70,000 samples from the GEO database corresponding to five widely-used human microarray platforms (GPL96, 97, 571, 5175 and 6480; table S2) and concluded that *NOD1* exhibits the least variation in expression among genes encoding known innate sensors regardless of the experimental context. Unlike traditional microarray analysis, which often focuses on studying specific cell types or diseases, our approach of calculating expression variance drew strength from combining publicly available data from diverse experimental conditions. To make datasets generated in different laboratories cross comparable, genome-wide expression data from each sample was converted to robust z-scores (Fig. 3A). For each available probe corresponding to the innate sensors, the variability in distribution of z-score across all the samples was estimated from the standard deviation. From this unbiased analysis, probes targeting the *NOD1* gene consistently showed minimal deviation in gene expression across all microarray platforms (Fig. 3B), implying that *NOD1* expression is under especially stringent control.

MicroRNAs tightly control NOD1 expression and are controlled by NOD1 activation

Because a small increase in NOD1 protein amount per cell leads to saturating changes in gene expression and NOD1 gene expression is also very tightly regulated, we next sought to uncover the mechanisms of *NOD1* expression control. One class of regulatory elements in the human genome that can exert tight but subtle control of protein abundance usually in the less than 2 fold range (32) are small noncoding RNAs called miRNAs. Correlation analysis showed that the expression of the pre-miRNAs mir-15b, mir-16–1 and mir-16–2 and miR-106a was negatively correlated with the expression of *NOD1* but not that of *NOD2* or *NLRP4* (Fig. 4A and Data File S3). Furthermore, the mature forms of these miRNAs (miR-15b, miR-16 and miR-106a) were predicted to bind to the 3'-UTR of *NOD1* in two independent miRNA-target prediction databases, TargetScan and miRanda (fig. S6A; note that the seed regions in miR-15b and miR-16 that are complementary to the *NOD1* 3'-UTR are identical (33)). Consistent with the negative correlation, quantitative RT-PCR showed that expression of miR-15b, 16 and 106a was reduced in THP-1 NOD1 cells, and moreover that miR expression was reduced to a similar extent in cells with low-level expression of *NOD1* or DOX-induced over-expression of *NOD1* when compared to empty vector cells (Fig. 4B). The effects on miR expression were similar regardless of the presence or absence of serum in culture media (fig. S6B). Expression of miR-191, an endogenous control, was unchanged.

To test if the observed reduction in miRNA abundance under conditions of low-level *NOD1* expression was due to NOD1 activation, we analyzed expression of miRNAs upon activation of NOD1 with the ligand iE-DAP in non-transduced, control THP-1 cells. Such activation triggers self-oligomerization of endogenous NOD1 resulting in formation of a protein

complex called the nodosome that activates NF-κB and MAPK. Indeed, treatment with iE-DAP led to an early decrease in miRNA expression by 60 min that was followed by a restoration of miRNA expression at later times (Fig. 4C); this later upswing may be indicative of a ligand-induced negative feedback response. In a reciprocal manner to the miRNAs, *NOD1* expression increased slightly early after ligand stimulation, after which NOD1 decreased at later times both at the RNA (Fig. 4C) and the protein level (Fig. 4D), which coincided with an increase in miR-15b, 16 and 106a expression. Activation of NF-κB showed similar kinetics and peaked by 60 min after iE-DAP stimulation followed by a decrease by 4 h (fig. S6C). These data suggest that ligand activation of NOD1 leads to a transient decrease in miRNA expression.

Because miR-15b, 16 and 106a were predicted to bind the 3'-UTR of *NOD1* (fig. S6A), we next examined whether these kinetic relationships in turn represented miRNA control of *NOD1* expression. To test this, we used highly specific locked nucleic acids (LNAs) to achieve short-term inhibition of miR activity in THP-1 cells and observed a small (20–30% or 1.2–1.3 fold) increase in NOD1 protein levels in cells transfected with miR-15b, miR-16 or miR-106a LNA compared to cells transfected with control LNA (Fig. 4E–F). LNA targeting miR-191 had no effect. This phenomenon was also observed in primary CD14+ monocytes isolated from human blood although miR-106a had a lesser effect compared to miR-15b and miR-16 in these cells (fig. S6D–E). Based on TargetScan and Miranda, binding sites for miR-15b/16 and miR-106a in the *NOD1* 3'-UTR are not conserved between humans and mice. Mouse *Nod1* was not predicted to be targeted by these miRNAs and consistent with this prediction, inhibition of miRNA activity in mouse BMDMs did not result in an increase in NOD1 protein levels (fig. S6F) despite abundant expression of these miRNAs in mouse cells at levels similar to those in THP-1 cells (fig. S6G). These data indicate that miR-15b/16 and miR-106a control of NOD1 protein expression is specific to human cells.

miR-15b/16 act directly on the NOD1 3'-UTR

MicroRNAs can bind the 3'-UTRs of multiple genes. To test if the observed effects of miRNAs on NOD1 expression directly depended on their predicted binding site in the *NOD1* 3'-UTR, we used a luciferase reporter assay system in which luciferase activity was controlled by an intact *NOD1* 3'-UTR (WT) or *NOD1* 3'-UTRs in which miR binding sites had been mutated (Fig. 5A). Compared to a minimal 3'-UTR, the WT *NOD1* 3'-UTR showed reduced luciferase activity when transfected into HEK-293 cells. Mutation of only the miR-15b/16 binding site (miR-15b/16mut) almost completely rescued this effect, and to an extent similar to that seen upon mutation of both the miR-15b/16 and miR-106a binding sites (miR-mut) in the *NOD1* 3'-UTR (Fig. 5B). Co-transfection of luciferase constructs with miR-15b or miR-16 mimics inhibited luciferase activity of the WT *NOD1* 3'-UTR (Fig. 5C, left) but, as expected, had no effect on *NOD1* 3'-UTRs in which either the miR-15b/16 binding site or all miR binding sites had been mutated (Fig. 5C, middle and right). Consistent with these data, LNA inhibitors to miR-15b or miR-16 increased luciferase activity of the WT *NOD1* 3'-UTR (Fig. 5D). On the other hand, neither miR-106a mimic nor LNA inhibitor had an effect on luciferase activity of the WT *NOD1* 3'-UTR (Fig. 5C, left and Fig. 5D). This result was in contrast to our observation that miR-106a LNA

increases the expression of endogenous NOD1 (Fig. 4E). We therefore think that the effect of miR-106a on regulating endogenous NOD1 may involve long-range interactions dictated by the secondary and tertiary structure of the full-length mRNA that includes the NOD1 coding region, 3'-UTR and 5'-UTR. Because luciferase reporters are minimalistic and only test the effect of NOD1 3'-UTR, such long-range interactions would not occur in our luciferase assay. Consistent with miR-15b/16 LNA-based experiments (Fig. 5D and 4E), co-transfection of luciferase constructs with a miR-15b/16 target site blocker (TSB), an antisense oligonucleotide designed to block the miR-15b/16 target site only in the *NOD1* 3'-UTR and not predicted to bind miR-15b/16 sites in any other known human 3'-UTR, increased luciferase activity of the WT *NOD1* 3'-UTR but not the mutated 3'-UTRs (Fig. 5E). Together, these results show that miR-15b and miR-16 directly exert their effect on *NOD1* expression through their specific binding sites in the *NOD1* 3'-UTR.

miR-15b/16 set the threshold for NOD1 ligand-mediated MAPK and NF-κB signaling

Because miR-15/16 played a role in regulating NOD1 expression through its 3'-UTR, we hypothesized that loss of miRNA control might lower the threshold for activation of pro-inflammatory signaling through NOD1. Short-term inhibition of miRNA activity in THP-1 cells with LNA led to a small 1.2–1.3X increase in NOD1 expression (Fig. 6A) that for most cells was below the 1.5-fold increase that led to spontaneous NOD1 signaling events (Fig. 1B–D). Subsequent treatment with iE-DAP, however, led to robust phosphorylation of the MAPK p38 in cells treated with miR-15b/16 LNA as compared to those treated with scramble LNA across all ligand concentrations tested, and a delayed degradation of IκBα only at a sub-saturating, otherwise inert concentration of ligand (10 ng iE-DAP) (Fig. 6B). Augmentation of NF-κB activation at a lower, suboptimal concentration of the ligand indicated that higher doses of ligand have a saturating effect on NF-κB activation. This agrees with work showing that distinct thresholds exist for NF-κB and MAPK signaling downstream of a common stimulatory ligand, with the threshold for NF-κB activation being lower than that for MAPK activation (34).

To further test if the observed effects of miR-15b/16 on MAPK and NF-κB activation directly depended on their binding to the *NOD1* 3'-UTR as opposed to indirect effects through binding of these miRNAs to unrelated target genes, we treated cells with miR-15b/16 TSB and examined p38 and NF-κB activation in response to ligand. As observed with LNA inhibitors, THP-1 cells treated with miR-15b/16 TSB for 24 h showed a subtle (~1.3 fold) increase in NOD1 expression (Fig. 6C). This small increase in NOD1 enhanced ligand–induced p38 phosphorylation at all ligand concentrations and IκBα degradation only at a suboptimal ligand concentration (10 ng iE-DAP) (Fig. 6D), suggesting that miR-15b/16 exert their effect on NOD1 expression and downstream inflammatory signaling through their predicted binding site in the *NOD1* 3'-UTR. Together, these data indicate that a small increase in NOD1 primes cells for heightened inflammatory signaling in response to sub-saturating concentrations of ligand, a characteristic of the system that might function as a double-edged sword. On the one hand, such a state may permit ambient concentrations of ligand derived for example from commensals, to activate inflammatory responses to a greater degree than when miRNA control is intact, setting the stage for persistent inflammation and pathological consequences. On the other hand, because NF-κB

and MAPK responses activated through NOD1 are critical for host defense against bacterial pathogens, this property of the system might also be important for functional rather than pathologic responses during infection because inflammatory responses to low amounts of ligand from invading pathogens may be readily mounted, providing a feedforward boost to the immune response as soon as bacterial ligand is present.

Prolonged disruption of miR-15b/16 control of NOD1 expression causes spontaneous pathologic gene activation

We next asked if a sustained small increase in endogenous NOD1 by loss of miR-15b/16 control triggers pathologic responses. First, we induced prolonged disruption of miR-15b/16-mediated control of NOD1 by treatment of cells with miR-15b/16 TSB that inhibits miRNA binding specifically to the *NOD1* 3'UTR. This led to a small increase in endogenous NOD1 protein by 24 h which was sustained until day 8 of TSB treatment and was accompanied by a significant induction of *C-KIT, ALX1* and *C-MYC* expression by day 3 (Fig. 7A–B). A similar upregulation of NOD1 and induction of oncogenes by day 3 upon inhibition of miR-15b/16 was also observed in primary CD14+ monocytes purified from human blood (Fig. 7C). Secondly, we genetically targeted miR-15b and miR-16 in THP-1 cells by CRISPR-Cas9 gene editing to achieve a long-term reduction in the expression of these miRNAs and upregulation of NOD1. miR-15b and miR-16 are present contiguously within an intronic region of the same gene (*SMC4*) and we found that gRNAs targeting miR-15b also significantly reduced miR-16 expression (Fig. 7D, left and fig. S7). Prolonged reduction in miR-15b/16 expression in two independent single cell clones for ~4–6 weeks led to a 1.5 to 2 fold increase in endogenous *NOD1* (Fig. 7D, left) and spontaneous induction of *C-KIT, ALX1* and *C-MYC* (Fig. 7D, right) as well as a tolerization of inflammatory gene expression (Fig. 7E). Together, these results indicate that a sustained small increase in NOD1 expression by disruption of miR-15b/16 function leads to pathologic gene activation. Overall, our data suggest that miR-15b/16 constitute a critical control mechanism that keeps NOD1 expression in check in normal cells. A prolonged small increase in NOD1 at or above a 1.5-fold threshold of expression leads to spontaneous induction of inflammatory genes and oncogenes, underscoring the importance of stringent control mechanisms needed to restrain NOD1 expression.

Discussion

The human genome contains many regulatory elements that modulate gene activity at different points in the progression from gene transcription to translation. Alterations in non-coding regions of the genome are linked to inflammatory and autoimmune diseases through GWAS (4), suggesting that persistent small changes in gene or gene product expression might play an important role in immune dysregulation. Here we provide evidence that a prolonged small increase in expression of NOD1, a ubiquitously expressed cytosolic sensor of bacterial infection, resulted in a large impact on cellular transcription state including both inflammatory and especially oncogene expression in the absence of ligand-driven activity. To avoid such spontaneous activity, NOD1 protein expression was tightly controlled in human cells by at least two miRNAs, miR-15b and miR-16. Our data identified an innate regulatory circuit involving miRNAs that kept NOD1 expression within a narrow window

below a digital switching point, which when exceeded can trigger persistent inflammation and induce pathologic gene expression (fig. S8). They also support the important role already assigned to miRNA dysregulation in some cancers and add to our understanding of the mechanisms by which aberrant control of innate immune responses may lead to the origin of oncogenic activity.

How might a sustained ≥1.5-fold increase in NOD1 expression lead to activation? Because NOD1 needs to oligomerize to signal, we hypothesize that it may operate in a manner akin to a sol-gel transition, whereby a small increase in its protein concentration triggers a molecular change from a monomeric species to a macromolecular or polymeric gel-like NOD1 signaling complex that can induce activation of downstream RIPK2-dependent signaling cascades. Such a phenomenon has been previously observed in multivalent signaling systems including T cell receptor signaling (35, 36). Because the protein concentration required for phase transition depend on physical properties of the monomeric species such as valency and affinity, such a molecular conversion is likely to underlie oligomerization of a large, multivalent entity like NOD1 that has an inherent propensity to not only self-associate but also interact with its downstream signaling adapter protein RIPK2 through homotypic protein-protein interactions. Furthermore, proteins in solution are believed to exhibit marked fluctuations in their three-dimensional structures, a movement referred to as protein ‘breathing’ (37). Maintaining a supra-physiologic concentration of NOD1 might allow the normal opening and re-closing of each molecule (that is, molecular breathing) to lead to gel transition that would not take place when the molecules are more separated and hence, more likely to self-close and cover the oligomerization domain before they associate with another open molecule in the cytosol. Our finding that NOD1 is the most tightly regulated intracellular bacterial sensor (Fig. 3B) and also that sustained small increases in its intracellular protein concentration drive inflammatory and proto-oncogene expression suggest that evolution has led to NOD1 being maintained at just below the triggering concentration. This yields a highly sensitive detector, at the risk of pathologic activation with small disturbances in expression control.

Why does a short-term increase in NOD1 below the 1.5-fold threshold by disruption of miRNA control not lead to auto-activation but rather only sensitizes cells to ligand (Fig. 6A–D)? We posit that this could be a safety mechanism built into cells because such small differences in expression may be within the range of random variation of mRNA or protein level within a given cell at different times or between different cells in a clonal population. It therefore seems plausible that cells would not want to spontaneously activate with very small (<1.5-fold) short-term increases in NOD1 unless another threat (for example, a bacterial ligand) is imminent or unless miRNA dysregulation persists so that a critical expression threshold at which auto-activation occurs is reached over time. Indeed, a small prolonged increase in NOD1 at or above 1.5 fold either by NOD1 transgene expression or by disruption of regulatory control of endogenous NOD1 by miR-15b/16, led to auto-oligomerization similar to that caused by ligand treatment (fig. S2B), spontaneous inflammatory gene activation (fig. S2A and Data File S2) and in particular a major effect on oncogene expression (Fig. 2A–D, fig. S3A–C and Fig. 7A–D). Such a prolonged increase in gene or gene product expression, which may be achieved by eQTLs or genetic variants in

humans, may potentially be one way by which a diseased state is initiated over time in cells of genetically predisposed individuals.

Persistent inflammation is believed to instigate oncogenesis in many ways, including triggering the transformation process itself as well as providing a suitable milieu for the proliferation of transformed cells. In our system, oncogene expression was substantially reversed by deletion of RIPK2 (Fig. 2I), suggesting that it may indeed be dependent on persistent inflammatory signaling through the NOD1-RIPK2 pathway. However, at least two signals are thought to be required to transform a normal cell into a tumor cell (38, 39), suggesting that to take advantage of this inflammatory microenvironment, the cells themselves need more than one hit to acquire such unconstrained oncogenic potential. We think it is unlikely that short-term activation of NOD1 alone would be sufficient to spur oncogenic changes in otherwise normal cells because ligand-induced, scaled activation of NOD1 appears to trigger negative feedback mechanisms that strongly repress NOD1 expression (Fig. 4C) and should therefore prevent unchecked inflammation. Thus, under normal conditions, NOD1 activation is counterbalanced by repression of its expression. To shift this balance towards chronic NOD1 expression and sustained inflammation, it is likely that a disruption of miR-15b/16 and/or other regulatory mechanisms that restrict NOD1 expression is required. Indeed, our data showed that a small increase in NOD1 due to loss of miR-15b/16 repression brought NOD1 levels closer to the ligand-independent threshold and sensitized cells to inflammatory responses in response to otherwise inert concentrations of ligand (Fig. 6A–D), a state that might permit usually innocuous concentrations of commensal products to activate inflammatory responses to a greater degree than when miRNA control is intact. Disruption of one or more mechanisms that sustain miR-15b/16 expression thus may constitute an initial trigger that unleashes the inflammatory potential of NOD1, and possibly also dysregulates the expression of other genes controlled by these miRNAs, creating a microenvironment that drives oncogenesis (3, 6).

Recent studies across several cancer types suggest a tumor regulatory architecture wherein functional master regulator proteins whose abnormal activity is necessary and sufficient for implementing a tumor cell state, integrate the effects of multiple and heterogeneous upstream genomic alterations. These master regulators are not themselves mutated but lie downstream of mutant genes that change the expression of the master regulator at the protein level by 1.5–3 fold, which in turn is believed to cause dysregulation of downstream genes that control cell growth and metastasis (7). Our data in monocytes show that a persistent 1.5-fold increase in NOD1 is sufficient to trigger a prolonged switch to oncogenes. In this regard, NOD1 may itself act as a master regulator whose altered activity propagates a tumor cell state. However, because NOD1 is expressed ubiquitously in the body, such oncogene induction could also be potentially relevant to expression of NOD1 in non-hematopoietic cell types such as epithelial cells. Therefore, in a physiological context, a small dysregulation of NOD1 expression, for example by defective miR-15b/16 control, could lead to a ‘double whammy’ effect whereby a trans effect of chronic inflammation involving monocytes promotes pro-malignant changes in epithelial cells that synergizes with a cis effect within epithelial cells themselves. Genetic variants in *NOD1* are associated with gastric cancer risk (14) and chronic inflammation triggered by activation of NOD1 by *Helicobacter pylori* is an initiating event in gastric cancer (16). Expression of miR-15b/16,

which we identify here as regulators of NOD1 expression, is reduced in gastric tumor cells (40, 41) and NOD1 expression is increased in biopsies of patients with gastritis and gastric cancer (15). Additionally, MAPK activation and macrophage-derived inflammatory mediators contribute to tumor development in the gastric mucosa (17, 42). Together, these studies raise the intriguing possibility that the ‘dangerous’ set-point behavior for NOD1 that we describe here, wherein a small persistent increase in its expression causes spontaneous cell signaling without ligand, could be connected to initiation of carcinogenesis by genetic risk factors in gastric cancer.

Overall, our data emphasize that very small prolonged changes in protein concentration can have large effects on cellular transcriptional state inciting inflammatory and potentially oncogenic behavior. Rather than the ‘Knockout = large effect’ studies that have captivated the field for decades, these data emphasize the need to consider in a more quantitative and subtle way how chronic inflammation and associated oncogenic activity may evolve in susceptible hosts over years through a sustained small shift in gene or gene product expression that translates into aberrant population responses over time, with perhaps the occurrence of such events in several genes necessary to escape multilayered control and manifest full-blown clinical disease. Future studies should consider the possible ramifications of such seemingly minor but persistent changes in gene expression on the origin of complex immune-related diseases like cancer and autoimmunity in humans.

Materials and Methods

Cells and reagents

THP-1 cells were obtained from ATCC and cultured in RPMI (Lonza) supplemented with 2 mM glutamine (Gibco), 10 mM HEPES (Corning) and 10% FCS (RPMI-10) in a humidified atmosphere of 5% CO_2 at 37°C. For serum-free experiments, THP-1 cells were cultured in Macrophage-SFM (serum free media; Thermo Fisher) containing L-glutamine for 2 weeks. Parent THP-1 cells and all stable cell lines generated in this study were regularly tested for potential mycoplasma contamination to ensure authenticity. Blood from healthy adult donors was obtained from BenTech under an MTA between BenTech and ISB approved by ISB’s Institutional Review Board (IRB). Vectors for gateway cloning namely pEN_TmiRc3 entry vector and pSLIK_Neo destination vector were kindly provided by Dr. Iain Fraser, NIAID, NIH. Primers and probes used for detection of endogenous and 3X-FLAG NOD1 and NLRP4 are listed in table S1.

Constructs and stable cell lines

For generation of stable cell lines, 3X-FLAG-tagged NOD1 and NLRP4 were cloned into an entry vector (pEN_TmiRc3) driven by a tetracycline-inducible (TRE) promoter and recombined into a lentiviral expression vector (pSLIK_Neo). Lentiviruses were produced in modified HEK-293T cells (provided by Dr. Iain Fraser) by co-transfecting plasmid DNA of interest along with pRSV, pVSV and pMDL plasmids as previously described (43). Virus was concentrated from the supernatant and used to infect THP-1 cells at low copy to ensure <30% infection frequency such that majority of the transduced cells contain a single viral integration. Stable cell lines were selected with 1 mg/ml G418 (Invivogen). Expression of

NOD1 and NLRP4 was induced by treatment with 1 μg/ml doxycycline (DOX; Sigma) for 6 h.

Microarray data analysis

THP-1 cells were seeded at a density of 50,000 cells per well in a 96-well plate and treated with 1 μg/ml DOX for 6 hours to induce NOD1 or NLRP4 expression. Total RNA was isolated from 4 replicates per treatment using the RNeasy miniprep kit (Qiagen) and microarray analysis performed using the Affymetrix GeneChip Human Gene 1.0 ST Arrays. Expression data was analyzed by GAGE (Generally Applicable Gene-set Enrichment) (20) using the 'gage' package in 'R' (R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/). Pathway information was derived from KEGG (Kyoto Encyclopedia of Genes and Genomes). Correlations between normalized expression values of miRNAs and *NOD1*, *NOD2* and *NLRP4* were determined using the "rcorr" function with Spearman's method in R.

Big data approach for analysis of expression of innate sensor genes from the GEO database

Genome-wide microarray expression data was downloaded from Gene Expression Omnibus (GEO) to the NIAID high-performance computing cluster by OMics Compendia Commons (OMiCC) project (44). Details about quality control, normalization, and annotation are provided in the 'Gene-expression data and pre-processing' section with Supplementary Note 2 of the OMiCC manuscript. Five different human microarray platforms with large numbers of samples (more than 6000 samples per platform) from different vendors were selected for the analysis. These microarray samples were generated from hundreds of microarray experiments. The list of platforms with details on numbers of experiments and samples analyzed from each of these platforms is shown in table S2. In order to compare data across experiments generated by different laboratories, we normalized each gene of interest within each experiment by computing robust z-scores. We chose to use a robust metric to ensure that outlier expression values, if any, had a less significant effect on the rescaled data. The robust z-scores were computed by taking each vector x of gene values in a given experiment and calculating: $z(x) = \frac{x - MED(x)}{MAD(x)}$, where $MED(x)$ is the median of the vector x, and $MAD(x)$ is the median absolute deviation of x defined as: $MAD(x) = MED(|x - MED(x)|)$. Then, the variability in the distribution of robust z-scores for each probe in a microarray platform across all samples was determined by computing the standard deviation (Fig. 3A). These steps were applied for each platform separately and a platform specific standard deviation for each probe was obtained. Finally, probes were mapped to gene symbols using annotation data downloaded from the OMiCC server. All the data and code can be downloaded upon request.

Purification of monocytes from human blood

PBMCs were isolated from heparinized venous blood of healthy adult donors (BenTech) by Ficoll-Paque (GE Healthcare) density gradient centrifugation. CD14+ monocytes were isolated from PBMCs by MACS using Monocyte Isolation Kit II (Miltenyi Biotec). Cell

purity was determined by staining with anti-human CD14 (61D3, eBioscience) and analyzed by FACS.

Transfection of cells with LNA inhibitors

MiRCURY LNA™ microRNA power inhibitors were purchased from Exiqon. The following LNA inhibitors were used: hsa-miR-15b-5p, hsa-miR-16–5p, hsa-miR-106a-5p (these LNAs target both human and mouse miRNA), hsa-miR-191–5p, hsa-miR-15b/16 NOD1 3'-UTR specific TSB and scramble negative control A. Briefly, THP-1 cells or ex vivo differentiated mouse BMDM were plated at a density of 2.5×10^5 per well in a 24-well plate in RPMI-10 the day prior to transfection. On the following day, cells were transfected with LNA inhibitors at a final concentration of 50 μM using HiPerfect transfection reagent (Qiagen). Transfection complex was prepared by combining 200 μl OPTI-MEM (Gibco) with 6 μl HiPerfect transfection reagent (Qiagen) and 1μL LNA inhibitor (50 nM stock), and incubated for 20 min at room temperature prior to addition to cells. NOD1 expression was analyzed at 24–36 h post transfection with LNAs. For experiments with ligand, cells were first treated with LNA-based power inhibitors at a final concentration of 75 μM without transfection reagent (to achieve the same inhibition efficacy as that seen with 50 μM inhibitor in presence of transfection reagent), incubated for 24 h and then ligand was transfected using HiPerfect transfection reagent. For long-term experiments with LNA (as in Fig. 7A–B), 1.5×10^6 THP-1 cells were seeded in a 12-well tray, incubated overnight and then treated with 8 μl LNA inhibitor (50 nM stock; without transfection reagent). After 24 h, half of the culture was removed for experiments and an equal amount of fresh media containing LNA (4 μl; 50 nM stock) was added to the well in order to maintain a consistent concentration of cells and LNA inhibitor in each well. This process was repeated for 8 days. For long term LNA experiments with primary monocytes (as in Fig. 7C), cells were treated with LNA at a final concentration of 50 uM without transfection reagent and analyzed 3 days after treatment.

Activation of NOD1 with ligand

THP-1 or CD14+ PBMC cells were seeded at a density of 0.5×10^6/mL overnight in a 24-well plate in RPMI-10 prior to exposure to C12-iE-DAP (InvivoGen). Ligand was transfected using HiPerfect (Invitrogen). The transfection complex was prepared by combining 200 μl OPTI-MEM with 6 μL HiPerfect and 1 μL ligand (from a 1 mg/mL stock), then incubated for 20 min at RT followed by dropwise addition to each well. Unless otherwise specified, cells were stimulated with 1 μg C12-iE-DAP. At specified times post treatment with ligand, cells were analyzed by flow cytometry, quantitative RT-PCR or cell lysates were prepared and analyzed by immunoblot.

NOD1 3'-UTR Luciferase Reporter Assays

Luciferase reporter constructs were engineered by cloning the WT *NOD1* 3'-UTR, a *NOD1* 3'-UTR with the miR-15b/16 binding site mutated (miR15b/16-mut), or *NOD1* 3'-UTR with miR-15b/16 and miR-106a binding sites mutated (miR-mut) (gBLOCK, IDT) into the pGL3 vector (Promega). For luciferase assays, HEK-293 cells were seeded at 4×10^4 cells/well in a 48-well plate overnight. Cells were transfected with 200 ng NOD1 WT or miR-15b/16-mut 3'-UTR luciferase reporter constructs and 10 ng eGFP using FuGENE HD transfection

reagent (Promega). Where noted cells were co-transfected with either miRNA mimics, LNA inhibitors or TSB at a final concentration of 100 nM using HiPerfect (Invitrogen) for 36–48 h. Cells were washed with PBS and lysed in cell lysis buffer (9803, Cell Signaling Technology). Luciferase activity was measured with the Steady-Glo Luciferase Assay System (Promega) using a microplate reader (Biotek Synergy H4 Plate reader). Luciferase activity was normalized for transfection efficiency (eGFP) and plotted as a fold change over the respective control.

Immunoblot analysis

Immunoblots were prepared using Bolt Bis-Tris gel systems (Invitrogen) and probed overnight at 4°C with antibodies to NOD1 (3545, Cell Signaling), IκBα (4814, Cell Signaling), p-p38 (4511, Cell Signaling), p-AKT (4060, Cell Signaling), pan-AKT (4691, Cell Signaling), C-MYC (13987, Cell Signaling), C-KIT (Ab 81, Santa Cruz), ALX1 (ab181101, abcam) or ERK2 (C-14, Santa Cruz). Anti-FLAG antibody conjugated to HRP (A8592, Sigma) was probed at room temperature for 30 min. Blots were visualized using SuperSignal West Dura chemiluminescent substrate (ThermoFisher) at a high or low exposure as indicated. For native gel analysis of NOD1 oligomers, cells were lysed using lysis buffer containing 1% NP-40. Cell debris were separated by centrifugation at 11000g for 15 min. Native sample buffer was added to the cell lysate and run on a 3–12% native Bis-Tris gel as per the manufacturer's instructions (ThermoFisher). Proteins were transferred to a nitrocellulose membrane followed by immunoblotting.

Quantitative RT-PCR

Total RNA was isolated from THP-1 cells or CD14+ PBMCs using the miRNeasy Mini Kit (QIAGEN) and cDNA was generated using SuperScript III First-Strand Synthesis System (ThermoFisher). Quantitative RT-PCR was performed using FAM-labeled TaqMan MGB probes (Applied Biosystems) for *NOD1* (hs00196075_m1), *C-KIT* (hs00174029_m1), *ALX1* (hs00232518_m1), *IL1B* (hs01555610_m1), *JUN* (hs01103582_s1), *NFKBIA* (hs00355671_g1), *TNFAIP3* (hs00234713_m1), *GAPDH* (hs02786624_g1), and *ACTB* (hs01060665_g1). Quantitative RT-PCR for hsa-miR-191–5p, hsa-miR-15b-5p, hsa-miR-16–5p and hsa-miR-106a-5p was performed using TaqMan MicroRNA Assays (Applied Biosystems). Gene mRNA levels were normalized to the housekeeping genes *GAPDH* or *ACTB* and microRNA levels were normalized to the endogenous control hsa-miR-191–5p.

Analysis of endogenous and transgenic NOD1 expression by flow cytometry

THP-1 or CD14+ PBMC cells were seeded in a 24-well tray at 2.5×10^5 cells/well in RPMI-10 prior to treatment and were exposed to DOX (1 μg /mL, Sigma) or media for 24 h to induce the expression of NOD1. Alternatively, cells were treated with LNA inhibitors and ligand as described elsewhere in these methods. Endogenous NOD1 expression was assessed using anti-rabbit NOD1 (H-176, Santa Cruz; B-4, Santacruz) and transgenic NOD1 expression was assessed using anti-FLAG M2 antibody (Sigma). Briefly, cells were fixed using CytoFix/CytoPerm (BD Biosciences) followed blocking with Human IgG (Pierce) and staining for primary and secondary antibodies. Cells were analyzed with a BD FACSCalibur flow cytometer (Becton Dickinson). Fluorescence of 10^4 to 10^5 cells per sample was

acquired in logarithmic mode. Expression was quantified by calculating the mean or median fluorescence intensity of each sample.

Measurement of phospho-p65, ALX1 and C-KIT by flow cytometry

THP-1 or CD14+ PBMC cells were seeded in a 24-well tray at 2.5×10^5 cells/well in RPMI-10 prior to treatment and were treated with LNA inhibitors and ligand as described elsewhere in these methods. The following antibodies were used: anti-rabbit phospho-p65 (Ser536) conjugated to Alexa Fluor 647 (93H1, Cell Signaling), anti-ALX1 (Sigma, HPA018905) and anti-C-KIT (eBioscience, 17–1178). Briefly, cells were fixed using 4% paraformaldehyde (Alfa Aesar), permeabilized with MeOH followed by blocking with Human IgG (Pierce) and staining with primary and secondary antibodies. After washing, cells were analyzed on a BD FACSCalibur flow cytometer (Becton Dickinson). Fluorescence of 10^4 to 10^5 cells per sample was acquired in logarithmic mode. Data was analyzed using FlowJo software.

CRISPR mediated ablation of *RIPK2* and *FLAG NOD1*

For CRISPR/Cas9 targeting of *RIPK2* or FLAG *NOD1*, the guide RNA targeting *RIPK2* or FLAG *NOD1* (IDT) was cloned into a single self-inactivating lentivirus plasmid pRRL-U6-empty-gRNA-MND-Cas9-t2A-Puro, that expresses a Cas9-T2A-puromycin resistance cassette controlled by an MND promoter (45). Expression of guide RNA was controlled by the U6-promter. THP-1 NOD1 cells were transduced with lentivirus and stably transduced cells were selected with puromycin (5 μg/ml; Invivogen). Gene targeting events were validated by sequencing and western blot for RIPK2 or FLAG NOD1 to identify RIPK2$^{-/-}$ or NOD1-FLAG$^{-/-}$ THP-1 cells. The sequence of the gRNA target site is as follows, where a (G) denotes a nucleotide added to enable transcription off the U6 promoter: gRNA targeting *FLAG-NOD1*: (G)AGATCATGATATCGATTACA; gRNA targeting *RIPK2*: (G)AGCGCAGGTCGGCGAGTTTG.

CRISPR targeting of miR-15b/16

Guide RNAs (IDT) were incubated with a fluorescently tagged tracrRNA (IDT) to make a stable duplex, which was then incubated with the Cas9 nuclease (IDT) to make the RNP complex. THP-1 cells were transfected with the RNP complex by nucleofection (Amaxa; SG Cell Line 4D Nucleofector Kit). Cells were sorted 24 h post transfection by FACS, selecting for cells expressing the tracrRNA fluorescent dye and then subcloned to obtain single cell clones. Gene targeting events were validated by sequencing and quantitative RT-PCR for miR-15b/16. Of the clones reported in this study clone 1 was from a targeting event by miR-15b gRNA #2, clone 2 was from a targeting event by miR-15b gRNA #1. The sequence of the gRNA target site is as follows: gRNA targeting *miR-15b* #1: TGTGCTGTCTACAGTACTTTA; gRNA targeting *miR-15b* #2: AGTACTGTAGCAGCACATC.

Statistical analyses

Data are presented as means ± standard error of the mean (SEM) or as indicated in the figure legend. Statistical analyses were performed using a Student’s unpaired t test (two-tailed)

without assuming a consistent standard deviation in either Prism Graphpad or with the “t test” function in R. p values for Spearman’s r-values (Fig. 4A and Data File S3) were computed using the R function cor.test(). p values < 0.05 were considered significant. Statistical parameters such as the value of n, the number of replicates, precision measures and statistical significance are reported in the figures and figure legends.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Dr. Steve Porcella and the Genomics Unit at Rocky Mountain Laboratories for microarray support, Dr. Daniel Stetson’s lab at the University of Washington for assistance with gene editing and Drs. James Heath and Nitin Baliga for critical reading of the manuscript.

Funding: This work was funded by the Institute for Systems Biology (ISB), U.S. National Institutes of Health grant R21-AI138258, and in part by the Intramural Research Program of NIAID, NIH and funds from the Steven and Alexandra Cohen Foundation to N.S. A.S.A. was supported by an American Association of Immunologists (AAI) Careers in Immunology Fellowship. All work excluding initial acquisition of microarray data and OMiCC analysis was performed at ISB.

Data and materials availability:

Cell lines and plasmids described in this study are available to the scientific community under an MTA with the Institute for Systems Biology upon request to N.S. Microarray data generated in this study is available at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE99550. All other data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

References and Notes

1. Clarke TB et al., Recognition of peptidoglycan from the microbiota by Nod1 enhances systemic innate immunity. Nat Med 16, 228–231 (2010). [PubMed: 20081863]
2. Doria A et al., Autoinflammation and autoimmunity: bridging the divide. Autoimmun Rev 12, 22–30 (2012). [PubMed: 22878274]
3. Colotta F, Allavena P, Sica A, Garlanda C, Mantovani A, Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. Carcinogenesis 30, 1073–1081 (2009). [PubMed: 19468060]
4. Farh KK et al., Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature 518, 337–343 (2015). [PubMed: 25363779]
5. Stoffels M, Kastner DL, Old Dogs, New Tricks: Monogenic Autoinflammatory Disease Unleashed. Annu Rev Genomics Hum Genet 17, 245–272 (2016). [PubMed: 27362340]
6. Coussens LM, Werb Z, Inflammation and cancer. Nature 420, 860–867 (2002). [PubMed: 12490959]
7. Califano A, Alvarez MJ, The recurrent architecture of tumour initiation, progression and drug sensitivity. Nat Rev Cancer 17, 116–130 (2017). [PubMed: 27977008]
8. Poliseno L et al., A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature 465, 1033–1038 (2010). [PubMed: 20577206]
9. Sumazin P et al., An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. Cell 147, 370–381 (2011). [PubMed: 22000015]
10. Chamaillard M et al., An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. Nat Immunol 4, 702–707 (2003). [PubMed: 12796777]

11. Laroui H et al., L-Ala-gamma-D-Glu-meso-diaminopimelic acid (DAP) interacts directly with leucine-rich region domain of nucleotide-binding oligomerization domain 1, increasing phosphorylation activity of receptor-interacting serine/threonine-protein kinase 2 and its interaction with nucleotide-binding oligomerization domain 1. The Journal of biological chemistry 286, 31003–31013 (2011). [PubMed: 21757725]
12. Keestra AM et al., Manipulation of small Rho GTPases is a pathogen-induced process detected by NOD1. Nature 496, 233–237 (2013). [PubMed: 23542589]
13. Keestra-Gounder AM et al., NOD1 and NOD2 signalling links ER stress with inflammation. Nature 532, 394–397 (2016). [PubMed: 27007849]
14. Li ZX et al., NOD1 and NOD2 Genetic Variants in Association with Risk of Gastric Cancer and Its Precursors in a Chinese Population. PLoS One 10, e0124949 (2015). [PubMed: 25933107]
15. Allison CC et al., Nucleotide oligomerization domain 1 enhances IFN-gamma signaling in gastric epithelial cells during Helicobacter pylori infection and exacerbates disease severity. J Immunol 190, 3706–3715 (2013). [PubMed: 23460743]
16. Suarez G et al., Modification of Helicobacter pylori Peptidoglycan Enhances NOD1 Activation and Promotes Cancer of the Stomach. Cancer Res 75, 1749–1759 (2015). [PubMed: 25732381]
17. Oguma K et al., Activated macrophages promote Wnt signalling through tumour necrosis factor-alpha in gastric tumour cells. EMBO J 27, 1671–1681 (2008). [PubMed: 18511911]
18. Tu S et al., Overexpression of interleukin-1beta induces gastric inflammation and cancer and mobilizes myeloid-derived suppressor cells in mice. Cancer Cell 14, 408–419 (2008). [PubMed: 18977329]
19. Bennett MW et al., Expression of Fas ligand by human gastric adenocarcinomas: a potential mechanism of immune escape in stomach cancer. Gut 44, 156–162 (1999). [PubMed: 9895372]
20. Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ, GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinformatics 10, 161 (2009). [PubMed: 19473525]
21. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M, KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 44, D457–462 (2016). [PubMed: 26476454]
22. Baron U, Bujard H, Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. Methods Enzymol 327, 401–421 (2000). [PubMed: 11044999]
23. Masumoto J et al., Nod1 acts as an intracellular receptor to stimulate chemokine production and neutrophil recruitment in vivo. J Exp Med 203, 203–213 (2006). [PubMed: 16418393]
24. Yuan H et al., ALX1 induces snail expression to promote epithelial-to-mesenchymal transition and invasion of ovarian cancer cells. Cancer Res 73, 1581–1590 (2013). [PubMed: 23288509]
25. Sotgia F et al., Caveolin-1 and cancer metabolism in the tumor microenvironment: markers, models, and mechanisms. Annu Rev Pathol 7, 423–467 (2012). [PubMed: 22077552]
26. Liang J et al., The C-kit receptor-mediated signal transduction and tumor-related diseases. Int J Biol Sci 9, 435–443 (2013). [PubMed: 23678293]
27. Xu Z et al., Frequent KIT mutations in human gastrointestinal stromal tumors. Sci Rep 4, 5907 (2014). [PubMed: 25080996]
28. Wang X et al., Recurrent amplification of MYC and TNFRSF11B in 8q24 is associated with poor survival in patients with gastric cancer. Gastric Cancer 19, 116–127 (2016). [PubMed: 25618371]
29. Sasaki T, Yamashita Y, Kuniyasu H, AKT plays a crucial role in gastric cancer. Oncol Lett 10, 607–611 (2015). [PubMed: 26622541]
30. Molinaro R, Mukherjee T, Flick R, Philpott DJ, Girardin SE, Trace levels of peptidoglycan in serum underlie the NOD-dependent cytokine response to endoplasmic reticulum stress. The Journal of biological chemistry 294, 9007–9015 (2019). [PubMed: 30996003]
31. Tomeczkowski J et al., Expression and regulation of c-kit receptor and response to stem cell factor in childhood malignant T-lymphoblastic cells. Leukemia 12, 1221–1229 (1998). [PubMed: 9697876]
32. Baek D et al., The impact of microRNAs on protein output. Nature 455, 64–71 (2008). [PubMed: 18668037]
33. Ellwanger DC, Buttner FA, Mewes HW, Stumpflen V, The sufficient minimal set of miRNA seed types. Bioinformatics 27, 1346–1350 (2011). [PubMed: 21441577]

34. Gottschalk RA et al., Distinct NF-kappaB and MAPK Activation Thresholds Uncouple Steady-State Microbe Sensing from Anti-pathogen Inflammatory Responses. Cell Syst 2, 378–390 (2016). [PubMed: 27237739]
35. Li P et al., Phase transitions in the assembly of multivalent signalling proteins. Nature 483, 336–340 (2012). [PubMed: 22398450]
36. Su X et al., Phase separation of signaling molecules promotes T cell receptor signal transduction. Science 352, 595–599 (2016). [PubMed: 27056844]
37. Makowski L et al., Molecular crowding inhibits intramolecular breathing motions in proteins. J Mol Biol 375, 529–546 (2008). [PubMed: 18031757]
38. Su T et al., Two-signal requirement for growth-promoting function of Yap in hepatocytes. Elife 4, (2015).
39. Hahn WC et al., Creation of human tumour cells with defined genetic elements. Nature 400, 464–468 (1999). [PubMed: 10440377]
40. Xia L et al., miR-15b and miR-16 modulate multidrug resistance by targeting BCL2 in human gastric cancer cells. Int J Cancer 123, 372–379 (2008). [PubMed: 18449891]
41. Zhang J et al., Circulating MiR-16–5p and MiR-19b-3p as Two Novel Potential Biomarkers to Indicate Progression of Gastric Cancer. Theranostics 5, 733–745 (2015). [PubMed: 25897338]
42. Yang M, Huang CZ, Mitogen-activated protein kinase signaling pathway and invasion and metastasis of gastric cancer. World J Gastroenterol 21, 11673–11679 (2015). [PubMed: 26556994]
43. Shin KJ et al., A single lentiviral vector platform for microRNA-based conditional RNA interference and coordinated transgene expression. Proc Natl Acad Sci U S A 103, 13759–13764 (2006). [PubMed: 16945906]
44. Shah N et al., A crowdsourcing approach for reusing and meta-analyzing gene expression data. Nat Biotechnol 34, 803–806 (2016). [PubMed: 27323300]
45. Gray EE et al., The AIM2-like Receptors Are Dispensable for the Interferon Response to Intracellular DNA. Immunity 45, 255–266 (2016). [PubMed: 27496731]
46. Xia Z et al., Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. Nat Commun 5, 5274 (2014). [PubMed: 25409906]

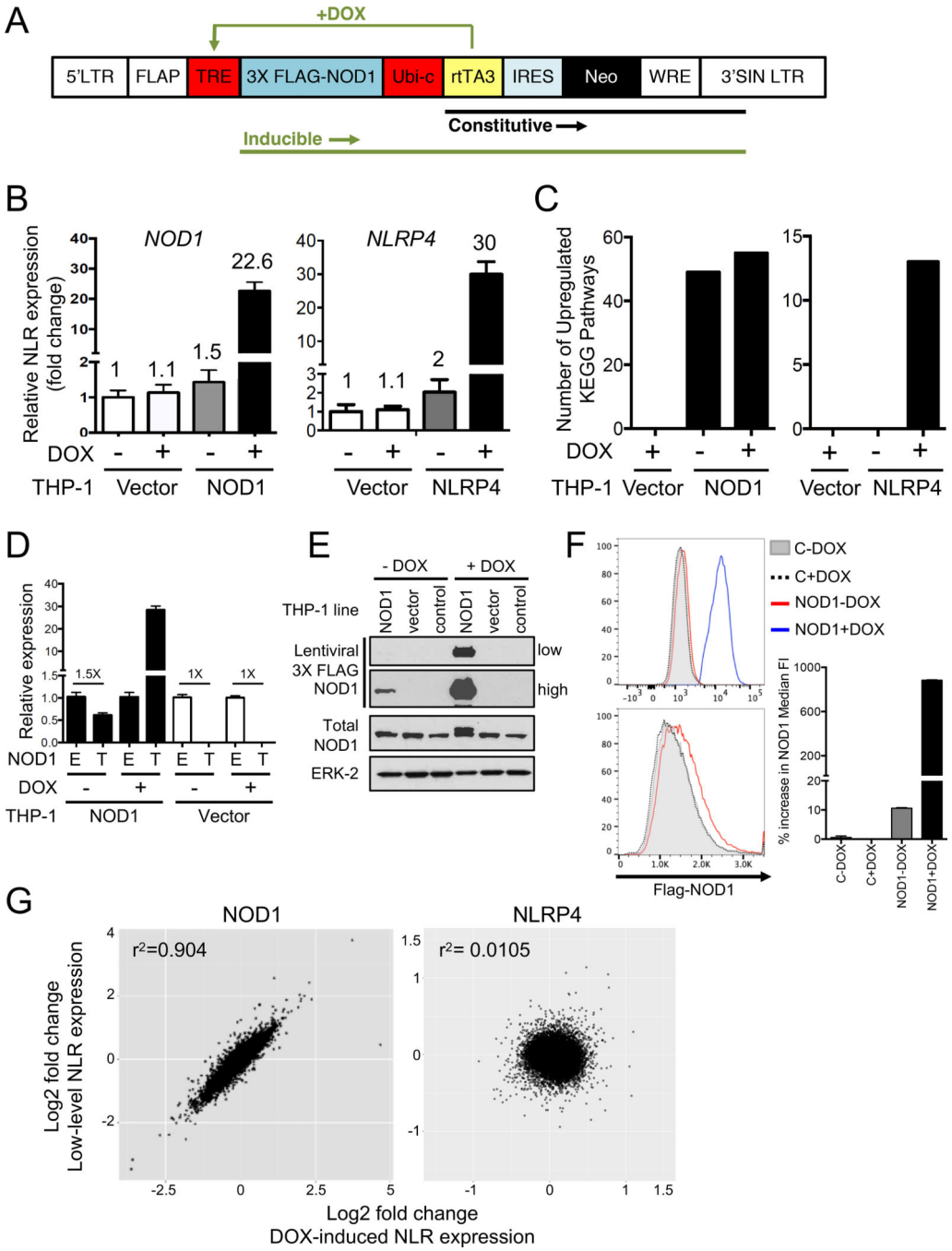

Fig. 1. NOD1 exhibits a switch-like behavior.

(A) Lentiviral system for DOX-inducible expression of NOD1 in THP-1 cells. **(B)** Microarray expression of *NOD1* and *NLRP4* in THP-1 NOD1, NLRP4 and Vector cells treated with (+) or without (–) DOX for 6 h. n=4 biological replicates per condition. **(C)** Number of KEGG pathways upregulated in the indicated THP-1 cell lines treated as in (B). **(D)** qPCR showing expression of endogenous NOD1 (E) or TRE-driven FLAG-tagged NOD1 (T) in the indicated THP-1 lines treated with or without DOX for 6 h. Labels above bars indicate total NOD1 expression. **(E-F)** Immunoblot (E) and flow cytometry (F) (plots:

left; quantification: right) for NOD1 protein in the indicated THP-1 lines with or without DOX. FACS plots show NOD1 expression on a log scale (top) and a linear scale (bottom). **(G)** Correlation plot of genes differentially expressed (fold change>1) in THP-1 NOD1 (left) and THP-1 NLRP4 (right) cells treated as in (B). In (E), 'control' refers to THP-1 cells with no lentivirus transduction. Data in D (qPCR), E (immunoblot) and F (FACS) are representative of at least three independent experiments and the bar graph in F includes pooled data from two independent experiments. Error bars are mean±SEM of four biological replicates.

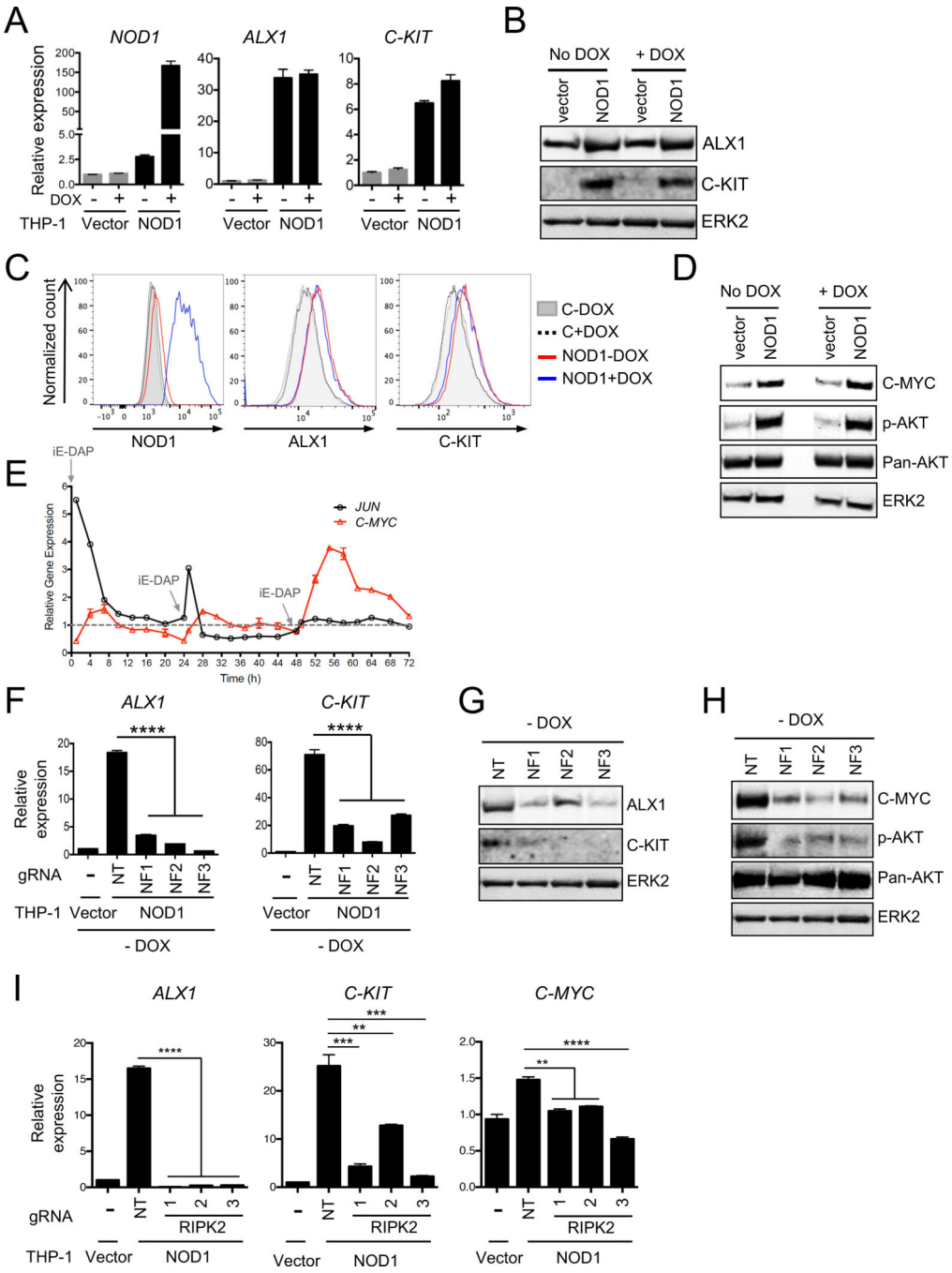

Fig. 2. Proto-oncogenes are induced by a sustained small increase in NOD1 expression.

(A-C) qPCR (A), immunoblot (B) and flow cytometry plots (C) showing oncogene expression in THP-1 NOD1 cells in absence and presence of DOX. **(D)** Immunoblot showing the abundance of the oncogene C-MYC and the phosphorylation (p) of AKT in THP-1 NOD1 cells in presence or absence of DOX. **(E)** Kinetics of *C-JUN* and *C-MYC* expression following ligand treatment as measured by qPCR. Grey arrows indicate times of recurring ligand addition (0, 24 and 48 h). Expression at each time point is represented relative to the untreated control at that timepoint (dashed line set to 1). **(F-H)** qPCR (F) and

immunoblots (G, H) for the indicated genes in THP-1 NOD1 cells following CRISPR/Cas9 mediated ablation of FLAG-*NOD1* in three independent single cell clones (NF1, NF2 and NF3). **(I)** qPCR for the indicated genes in THP-1 NOD1 cells following CRISPR/Cas9 mediated ablation of *RIPK2* in three independent single cell clones. NT: non-targeting gRNA. NF: gRNA targeting FLAG-NOD1. RIPK2: gRNA targeting RIPK2. Data in (A) to (I) are representative of three independent experiments. Error bars on graphs are mean±SEM of at least three biological replicates and where not visible in (E) are shorter than the height of the symbol. p values from an unpaired t test (two-tailed) without assuming a consistent standard deviation are shown. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

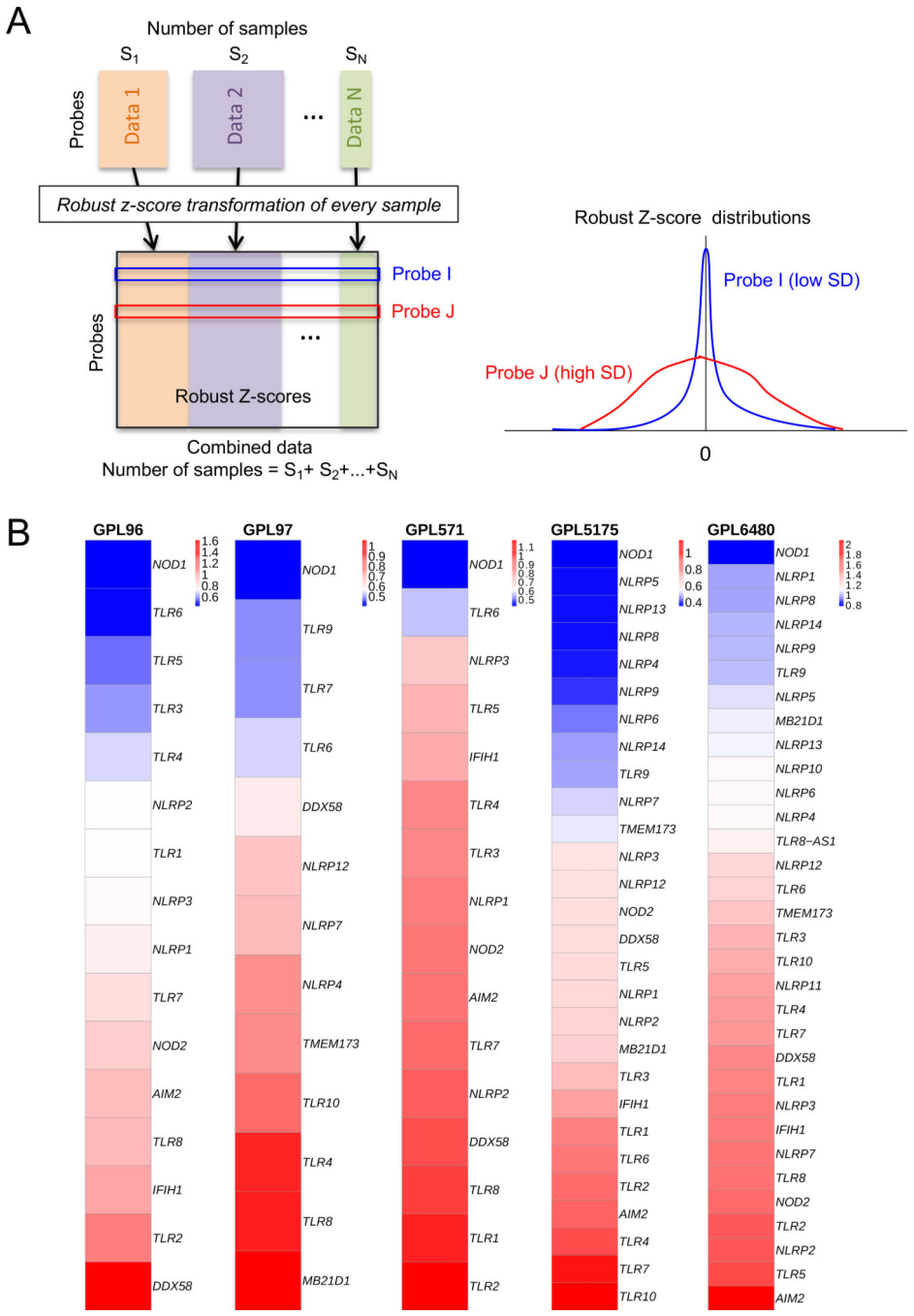

Fig. 3. *NOD1* shows minimal change in expression regardless of experimental context.

(A) Schematic for analysis of publicly available microarray data from the GEO database. Genome-wide expression data from 70,753 samples from five different microarray platforms namely GPL96, 97, 571, 5175 and 6480 were analyzed and transformed into robust z-scores. For each probe corresponding to innate sensors, the variability in distribution of z-score across all samples was estimated from the standard deviation (SD). Probe I and probe J refer to probes for two hypothetical genes on the microarray with low and high standard deviation in gene expression respectively. **(B)** Heatmaps showing standard deviation in expression of

innate sensors from the analyses conducted in A. Genes are arranged in ascending order of standard deviation from top to bottom (blue to red). Not all genes are present on all microarray platforms; for each platform, only genes for which probes were present on that platform are shown.

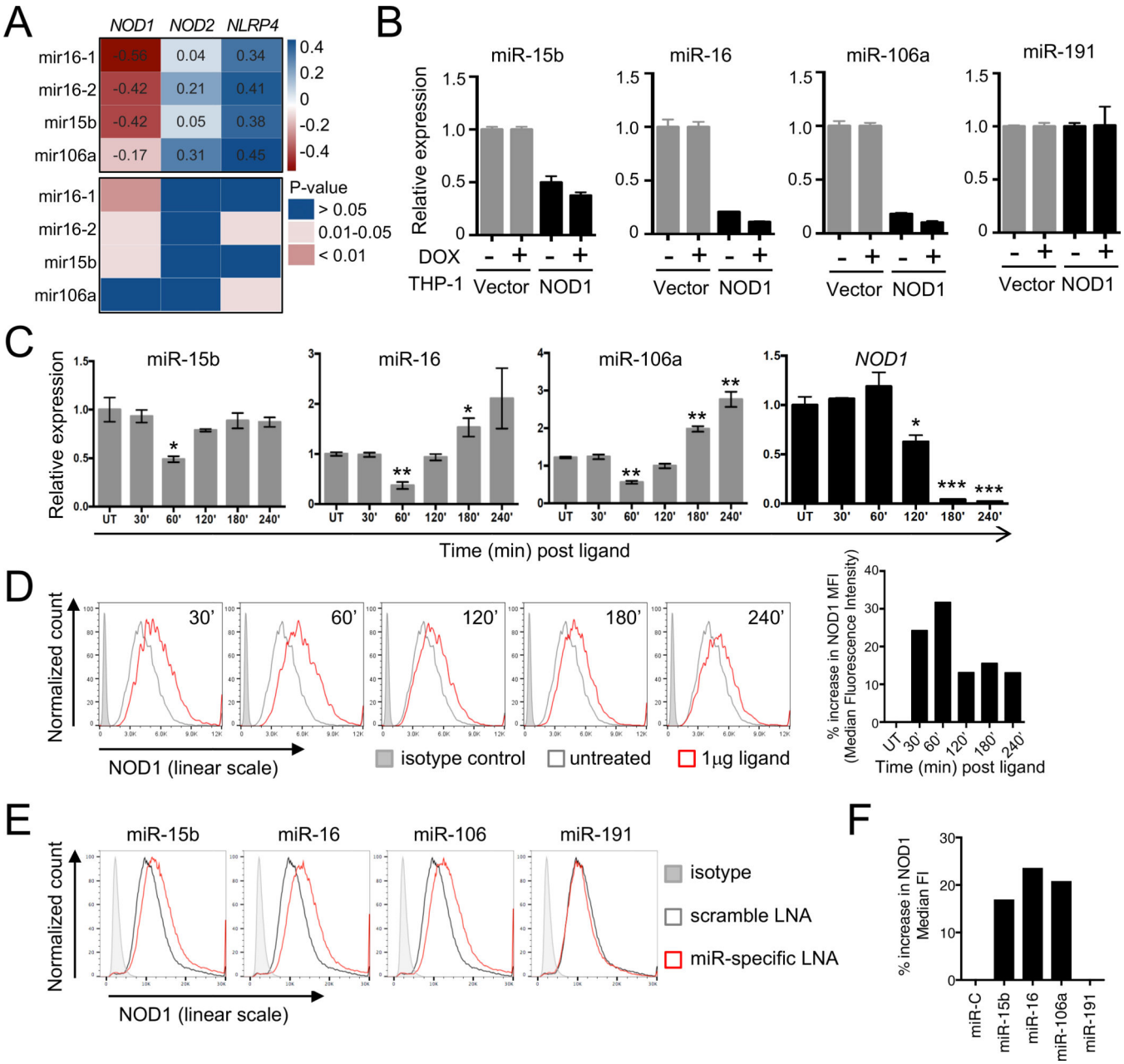

Fig. 4. miRNAs tightly regulate NOD1 expression and are in turn regulated by its activation.

(A) Top: Correlation (Spearman's) of mir-15b, mir-16 and mir-106a with *NOD1*, *NOD2* and *NLRP4*. Data originate from biological quadruplicates of a microarray experiment in THP-1 cells. Numbers indicate r-values (determined using the "rcorr" function with Spearman's method in R) and color depicts strength of correlation. Bottom: p-values corresponding to the r-values above. **(B)** q-PCR of miRNA expression in THP-1 NOD1 and Vector cells treated with (+) or without (–) DOX. **(C-D)** qPCR for miRNAs and *NOD1* (C) and FACS for NOD1 (plots, left; quantification, right) (D) in THP-1 cells treated with 1 μg C12-iE-DAP for the indicated times. **(E)** FACS for NOD1 protein in THP-1 cells transfected with LNA specific to the indicated miRNAs. **(F)** Quantification of histograms in (E). Data in B-E are

representative of three independent experiments. Error bars on graphs are mean±SEM of at least three biological replicates and p values from an unpaired t test (two-tailed) without assuming a consistent standard deviation are shown. * $p<0.05$, ** $p<0.01$, *** $p<0.001$. In C, each 30', 60', 120', 180' and 240' value was tested against the untreated (UT) condition.

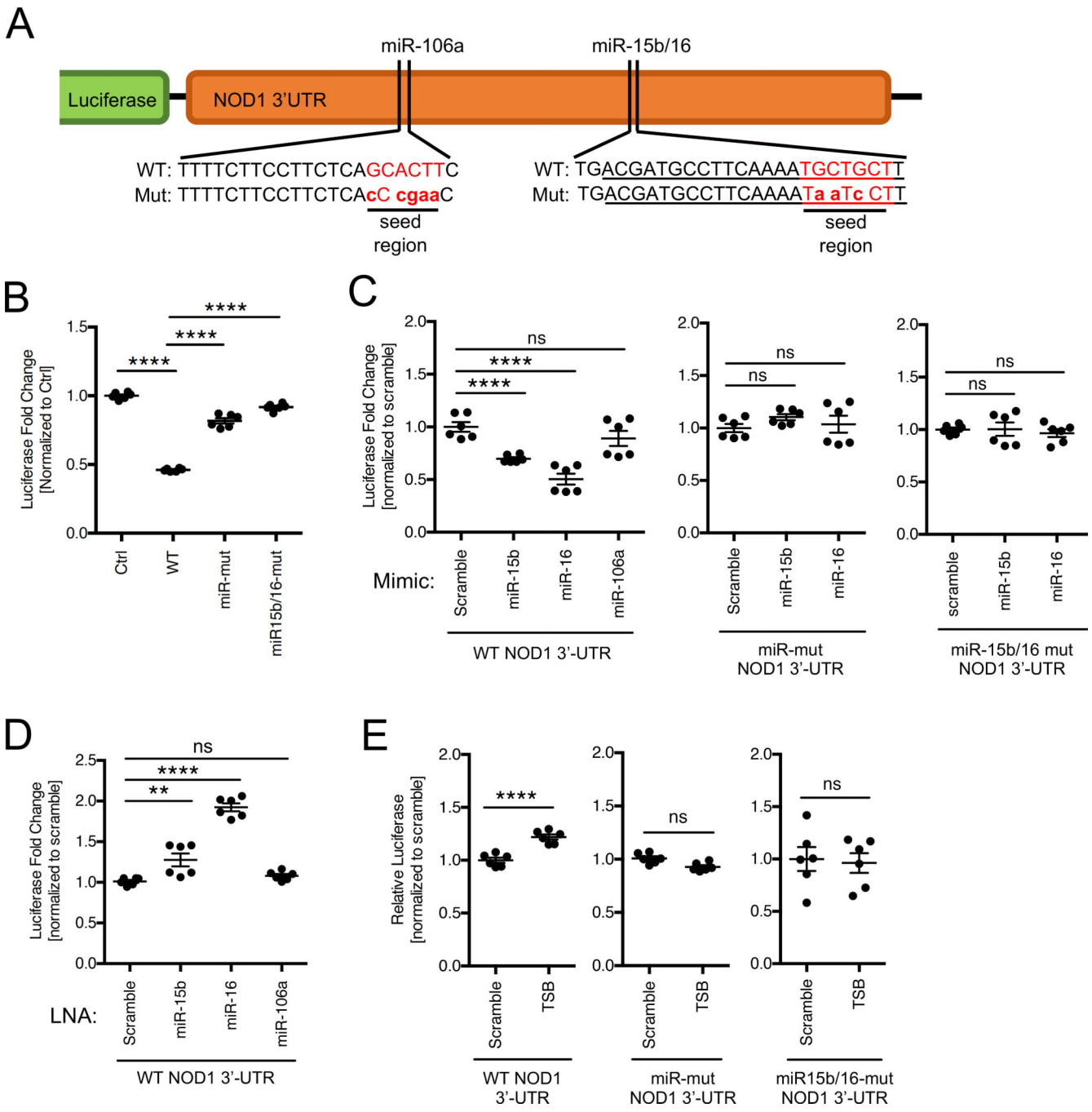

Fig. 5. miR-15b and miR-16 exert their activity through their binding site in the *NOD1* 3'-UTR.

(A) Schematic of luciferase reporter constructs showing the miR-15b/16 and miR-106a binding sites in the *NOD1* 3'-UTR (WT) and mutations introduced to obtain mutated 3'-UTRs (Mut). Seed regions are in red and mutations are in lowercase. **(B)** Luciferase reporter assay in HEK-293 cells showing decreased luciferase activity of the WT *NOD1* 3'-UTR compared to a minimal 3'-UTR (Ctrl; a 3'-UTR with no known regulatory sequences), a *NOD1* 3'-UTR in which the miR15b/16 binding site was mutated (miR-15b/16 mut) or a *NOD1* 3'-UTR in which all miRNA binding sites were mutated (miR-mut). **(C)** Luciferase

activity of the WT *NOD1* 3'-UTR (left), miR-mut 3'-UTR (center) or miR15b/16-mut 3'-UTR (right) in response to the indicated miRNA mimics. HEK-293 cells were co-transfected with the indicated 3'-UTR luciferase constructs and miRNA mimics. (**D**) Luciferase activity of the WT *NOD1* 3'-UTR in response to miR-15b, miR-16, miR-106a or scramble LNA. HEK-293 cells were co-transfected with the WT *NOD1* 3'-UTR luciferase construct and the indicated LNAs. (**E**) Luciferase activity of the WT *NOD1* 3'-UTR (left) and the mutated 3'-UTRs (center and right) in response to miR-15b/16 target site blocker (TSB). HEK-293 cells were co-transfected with the indicated 3'-UTR luciferase constructs and TSB. Data in B-E are representative of at least three independent experiments. Error bars on graphs are mean ±SEM of the indicated number of biological replicates. p values from an unpaired t test (two-tailed) without assuming a consistent standard deviation are shown. ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$, ns: not significant.

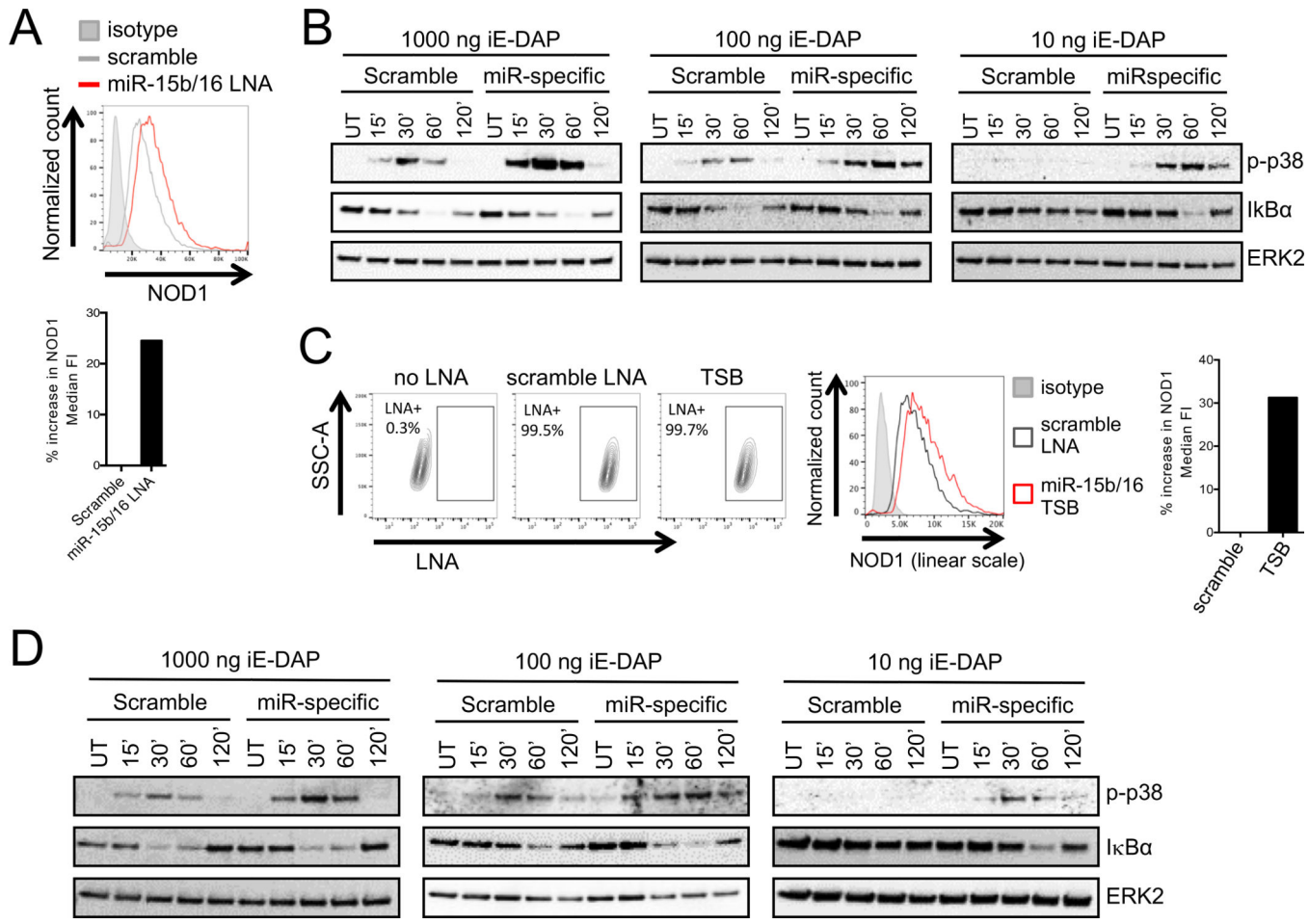

Fig. 6. A small increase in NOD1 by disruption of miR-15b/16 function sensitizes cells to ligand-induced pro-inflammatory signaling.

(A) Histogram (top) and quantification (bottom) of NOD1 protein in THP-1 cells transfected with miR-15b/16 LNA or scramble LNA. **(B)** Immunoblots of p-p38 and IκBα in THP-1 cells treated with scramble or miR-15b/16 LNA and then left untreated (UT) or treated with the indicated concentrations of C12-iE-DAP for the indicated times. **(C)** Representative FACS plots with gating strategy (left), histogram (center) and corresponding quantification (right) for NOD1 protein in THP-1 cells transfected with scramble TSB or TSB specific for the miR-15b/miR-16 binding site in the NOD1 3'-UTR. **(D)** Immunoblots of p-p38 and IκBα in THP-1 cells treated with scramble or miR-15b/16-specific TSB and either left untreated (UT) or treated with the indicated concentrations of C12-iE-DAP for the indicated times. Data in A and C are one representative of three independent experiments, and in B and D are one representative of two independent experiments.

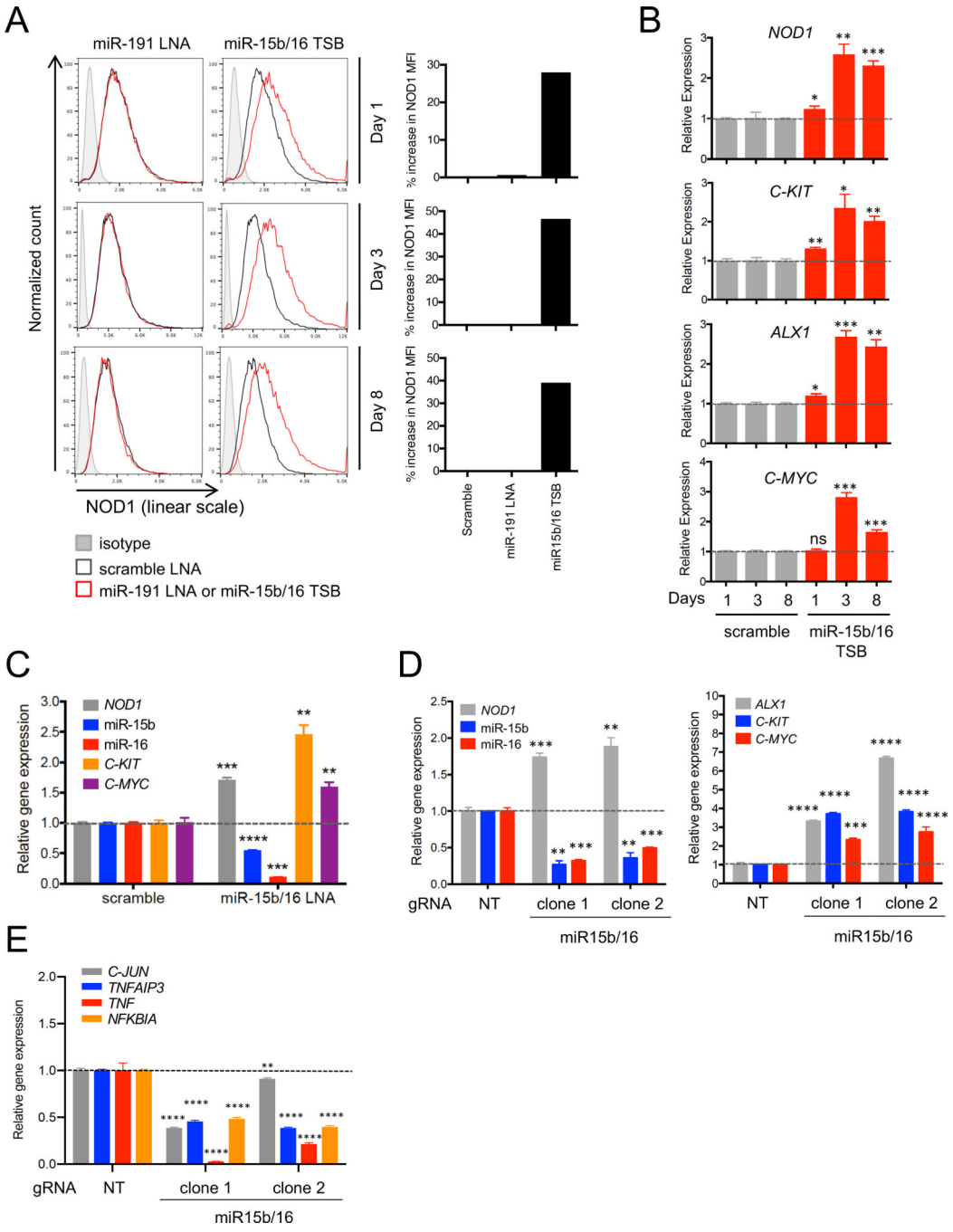

Fig. 7. Prolonged loss of miR-15b/16 control of NOD1 expression induces spontaneous upregulation of proto-oncogenes.

(A) Histograms (left) and quantification (right) of NOD1 protein in THP-1 cells treated with scramble or miR-191 LNA or miR-15b/16 TSB. **(B)** qPCR for the indicated oncogenes in cells treated with miR-15b/16 TSB. Gene expression at each timepoint is represented relative to the scramble LNA-treated control at that timepoint (dashed line set to 1). **(C)** qPCR for miR-15b, miR-16, *NOD1* and oncogenes in ex vivo human monocytes following inhibition of miR-15b and miR-16 with LNA. Gene expression is represented relative to that

in the scramble LNA control (dashed line set to 1). **(D)** qPCR for miR-15b, miR-16 and *NOD1* (left) and the indicated oncogenes (right) in THP-1 cells following CRISPR/Cas9 mediated reduction of miR15b/16 in two independent clones analyzed 4–6 weeks after miR-15b/16 targeting and single cell cloning. Gene expression is represented relative to that in the non-targeting (NT) gRNA control (dashed line set to 1). **(E)** qPCR for the indicated inflammatory genes in THP-1 cells following CRISPR/Cas9 reduction of miR15b/16. Data in A, B and D are representative of three independent experiments, and in E are representative of two independent experiments with three biological replicates per condition. In C, one representative of two independent donors is shown with three replicates per condition. Error bars on graphs are mean±SEM and p values from an unpaired t test (two-tailed) without assuming a consistent standard deviation are shown. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$, ns: not significant.