



Citation for published version:

Roux, BT & Lindsay, MA 2015, 'LincRNA signatures in human lymphocytes', *Nature Immunology*, vol. 16, no. 3, pp. 220-222. <https://doi.org/10.1038/ni.3106>

DOI:

[10.1038/ni.3106](https://doi.org/10.1038/ni.3106)

Publication date:

2015

Document Version

Peer reviewed version

[Link to publication](#)

This is the author's accepted version of a paper published in: Roux, BT & Lindsay, MA 2015, 'LincRNA signatures in human lymphocytes' *Nature Immunology*, vol 16, no. 3, pp. 220-222., and available via: <http://dx.doi.org/10.1038/ni.3106>.

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Intergenic long non-coding RNAs emerge as potential regulators of the human adaptive immune response

Benoit T. Roux & Mark A. Lindsay

Benoit T. Roux and Mark A. Lindsay are in the Department of Pharmacy and Pharmacology, University of Bath, UK

e.mail: m.a.lindsay@bath.ac.uk

Sequencing studies have provided a comprehensive catalogue of the expression of intergenic long non-coding RNAs (lincRNAs) in 13 human T- and B-cells subtypes. A fraction of these lincRNAs are shown to be subtype selective and evidence is presented that linc-MAF-4 might regulate T-cell differentiation

It is now widely accepted that the microRNA family of short non-coding RNAs play an important regulatory role in virtually all aspects of the immune response including the proliferation, differentiation and activation of immune cells ¹. In contrast, little is known about long non-coding RNAs (lncRNAs), tens of thousands of which have been identified using high throughput sequencing. To address this issue, the report by Ranzani *et al* (REF) has catalogued the expression of the intergenic family of lncRNAs (lincRNA) in 13 human T- and B- subtypes, from which they have identified a group of 180 lincRNAs that are subtype selective. Significantly, a member of this group entitled *linc-MAF-4* was shown to promote a CD4⁺ T_H1 phenotype by acting *in cis* to suppress the expression of the CD4⁺ T_H2 promoting transcription factor, MAF-4.

Interest in this area of non-coding RNAs has been driven by the rapid advances in sequencing technology, which have shown that much of the genome is transcribed into lncRNAs (> 200 nucleotides) ². Since little is currently known about their function and mechanism, these tend to be arbitrarily grouped into intergenic, antisense and pseudogenes families, based upon their position relative to protein coding genes and/or sequence conservation. In this regard, the lincRNAs that are described by Ranzani *et al*. (REF) are located between protein-coding genes. These differ from the antisense lncRNAs that are transcribed across the exons of protein coding genes from the opposite strand and the transcribed pseudogenes lncRNAs that are generally produced when a gene loses its ability to produce a functional protein.

The first indication that lncRNAs might regulate the innate immune response was the identification of a lincRNA located downstream of *Cox2* (PTGS2), that was up-regulated in mouse macrophages following exposure to lipopolysaccharide ³. As a result of its

proximity to the *Cox2* gene, this was named *lincRNA-Cox2*. Subsequent studies have shown that *lincRNA-Cox2*, as well as other lincRNAs such as *LETHE*, *THRIL*, *NEAT1*, *PACER* and *IL-1 β -RBT46* regulate the production of inflammatory mediators although their physiological importance has yet to be confirmed using an *in vivo* model of innate immunity ^{4,5}.

By comparison, despite the fact that the T-cell associated lincRNAs, *NRON* (non-coding repressor of NFAT) and *Gas5* (growth arrest specific 5) were discovered nearly a decade ago, much less is known about the adaptive immune response ^{4,5}. Crucially, this situation that has been transformed by the publication of Hu *et al* ⁶ and the current report by Ranzani *et al.* (REF), that are primarily based upon sequencing studies. However, before proceeding with a detailed discussion on these papers, it is also important to highlight the genetic studies of Gomez *et al* ⁷, which are the first to have demonstrated that lincRNAs regulated the immune response *in vivo*. Thus, this seminal piece of work begins with the identification of a lincRNA named *NeST* (Nettoie Salmonella pas Theilers's) within a genetic locus (*Tmevp3*) that was historically associated with murine viral susceptibility ⁷. The authors then proceed to show that over-expression of *NeST* was responsible for the increased Theilers's viral persistence and reduced *Salmonella enterica* pathogenesis observed in mice, that was mediated through *NeST* induced up-regulation of IFN γ in CD8+ T-cells ⁷.

As already mentioned, investigations into lincRNAs commonly begin by using sequencing to identify those that are differentially expressed in different biological conditions. It is this approach that has been employed by Ranzani *et al.* (REF) to examine the profile of lincRNA expression in 13 lymphocytes populations including 7 CD4⁺ T-cells subsets, 3

CD8⁺ T-cell subsets and 3 B-cell subsets. Significantly, a similar approach was adopted to investigate the role of lincRNAs during T-cell development and differentiation, in an article by Hu *et al* ⁶ previously published in *Nature Immunology*. However, in contrast to the current report that isolated lymphocytes from human blood, Hu *et al* ⁶ employed mouse T-cells derived from the spleen (proliferation) or following differentiation of naïve CD4⁺ cells *in vitro*. The distinction between using human and mouse cells is crucial since lincRNAs are generally believed to exhibit low evolutionary sequence conservation, with estimates that only ~ 15% of mouse protein-coding genes have homologs in humans ⁸. It is therefore difficult to extrapolate profiling studies between these two species and underlies the primary impact of the current publication.

In order to catalogue the lincRNAs, the authors employed a combination of programmes to alignment (Tophat and STAR) and assemble (Cufflinks and Trinity) the sequence data. Following the removal of potential protein coding genes, Ranzani *et al.* (REF) identified 4754 lincRNAs including 563 novel transcripts, within the various human lymphocyte populations, a number comparable with the 1524 lincRNAs identified in mouse T-cells ⁶. Unfortunately, no attempt was made to compare the sequences of these two populations although both were located close to, and their expression correlated with, genes implicated in the immune response, implying a possible functional link. Having identified these lymphocyte association lincRNAs, cluster analysis showed that in comparison to mRNAs, these lincRNAs are expressed in a highly subtype specific manner, an observation that was previously made in mouse T-cells ⁶. Importantly, this included a smaller group of 180 lincRNAs that were selectively expressed in a single T- or B-cell population (> 2.5 fold increase compared to all other populations) (Figure).

Clearly, the selective expression of lincRNAs in different lymphocytes populations raised the question of whether or not they are regulating lymphocyte differentiation and/or function. To address this, the authors examined the role of *linc-MAF-4*, a lincRNA that was selectively expressed in CD4⁺ T_H1 cells and, as the name implies, was located upstream of *MAF*. This location is important since MAF-4 has been implicated in the differentiation of CD4⁺ T_H2 and the authors found an inverse correlation between the expression of *linc-MAF-4* and MAF-4. Significantly, knockdown of *linc-MAF-4* skewed the mRNA expression towards the CD4⁺ T_H2 phenotype indicating that this might contribute to the establishment of the T_H1 phenotype. Equivalent observations were made by Hu *et al.* ⁶ when examining *lincR-Ccr2-5'AS*, a lincRNA located adjacent to a cluster of chemokine receptor-encoding genes including *Ccr2*. This was selectively up-regulated during differentiation of CD4⁺ T_H2 cells and knockdown studies demonstrated that *lincR-Ccr2-5'AS* was not only required for T-cell migration into the lung but also appeared to regulate many genes associated with T-cell differentiation ⁶.

Having demonstrated repression of MAF-4 expression, Ranzani *et al.* (REF) then proceeded to investigate the mechanism of action of *linc-MAF-4*. To our mind, this is one of the most interesting aspects of the emerging lincRNA field, with evidence that the actions of individual lincRNAs are mediated both through protein binding and/or base pairing with RNA/DNA ⁹. Indeed, with the increasing numbers of immune-related lincRNAs that have been characterized, it is possible to divide these into 3 general mechanisms. The first can act as 'decoys' to prevent protein-DNA interaction and include those that bind to nuclear factor-κB and the glucocorticoid receptor ^{4,5}. Another group has been shown to bind to proteins such as the importin-beta family and signal transducer and activator of transcription 3, and antagonize their interaction with other proteins ^{4,5}. In

contrast to these, less is known about the final group that interact with components of chromatin modifying complexes including the polycomb repressor complex 2 (PRC2) ^{4,5}. Interestingly, it would appear that *linc-MAF-4* falls within this final group since its action required DNA looping to bring *linc-MAF-4* into proximity with the MAF-4 promoter, and the subsequent recruitment and activation of the PRC2 associated histone methyltransferases, EZH2. The histone demethylase LSD1 was also located in the complex although its activity appeared unchanged. As with all these chromatin-associated lincRNAs, additional studies are required to determine all the component of the chromatin complex, the kinetics of assembly and contribution of *linc-MAF-4*.

Overall, the report by Ranzani *et al.* (REF) has provided the first comprehensive catalogue of lincRNA expression in the various human T- and B-cells populations of circulating blood. As with *linc-MAF-4*, it is to be hoped that this will provide the foundations for the difficult and time-consuming process of identification of those lincRNAs that regulate the various aspects of the adaptive immune response. However, in moving forward there are a number of issues that will need to be addressed. As already emphasized, potentially the most significant hurdle is the lack of evolutionary sequence conservation, which makes it difficult to extrapolate from animal models to humans. This will need to be addressed through the development of more appropriate algorithms for the identification of the protein binding structures and/or the short RNA/DNA pairing regions that are implicated in lincRNA action ⁸. It will also be important to develop more efficient and high-throughput approaches to the identification of the functional lincRNAs amongst the thousands that are characteristically uncovered using high-throughput sequencing. In this regard, transcriptional regulation using CRISPR interference (CRISPRi) and over-expression (CRISPRa) appears like an exciting alternative

to using siRNA and antisense ¹⁰ and has already been employed to knockdown six common lncRNAs ¹¹. In summary, we are only at the beginning of our journey into understanding the role of the lincRNAs and the other families of lncRNAs, in the immune response. However, if this turns out to be as varied as that of miRNAs, there will be many exciting discoveries ahead.

Figure Legend

The illustration shows those human lymphocytes T- and B-cells subsets for which the profile of expression of long intergenic non-coding RNAs was determined and summarizes the potential interaction between *linc-MAF-4* and MAF-4 in the regulation of the T_H1 and T_H2 phenotype. The number in brackets describes how many lincRNAs were selective expressed in each subset, the names and genomic positions being available in Figure 2 and the Supplemental Figure 2 of the manuscript by Ranzani *et al.* (REF).

References

1. O'Connell, R. M., Rao, D. S. & Baltimore, D. microRNA regulation of inflammatory responses. *Annu. Rev. Immunol.* **30**, 295–312 (2012).
2. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
3. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
4. Heward, J. A. & Lindsay, M. A. Long non-coding RNAs in the regulation of the immune response. *Trends Immunol.* **35**, 408–419 (2014).
5. Atianand, M. K. & Fitzgerald, K. A. Long non-coding RNAs and control of gene expression in the immune system. *Trends Mol Med* **20**, 623–631 (2014).
6. Hu, G. *et al.* Expression and regulation of intergenic long noncoding RNAs during T cell development and differentiation. *Nat. Immunol.* **14**, 1190–1198 (2013).
7. Gomez, J. A. *et al.* The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon- γ locus. *Cell* **152**, 743–754 (2013).
8. Ulitsky, I. & Bartel, D. P. lincRNAs: Genomics, Evolution, and Mechanisms. *Cell* **154**, 26–46 (2013).
9. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. **14**, 699–712 (2013).
10. Bassett, A. R. *et al.* Considerations when investigating lncRNA function in vivo. *Elife* **3**, e03058 (2014).
11. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).

