

**Supplemental Information**

**Dual RNA-Seq of Human Leprosy Lesions  
Identifies Bacterial Determinants Linked  
to Host Immune Response**

**Dennis J. Montoya, Priscila Andrade, Bruno J.A. Silva, Rosane M.B. Teles, Feiyang Ma, Bryan Bryson, Saheli Sadanand, Teia Noel, Jing Lu, Euzenir Sarno, Kristine B. Arnvig, Douglas Young, Ramanuj Lahiri, Diana L. Williams, Sarah Fortune, Barry R. Bloom, Matteo Pellegrini, and Robert L. Modlin**

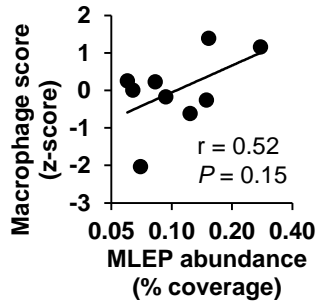
	LL1	LL2	LL3	LL4	LL5	LL6	LL7	LL8	LL9
<b>Total raw exonic reads</b>	44,712,768	57,381,932	55,656,315	43,284,470	58,924,537	74,403,791	46,699,318	44,636,954	55,489,167
<i>M. leprae</i> raw exonic reads	918,256	541,624	718,047	1,705,199	831,029	3,282,448	619,919	762,590	533,522
<b>Human raw exonic reads</b>	43,794,512	56,840,308	54,938,268	41,579,271	58,093,508	71,121,343	46,079,399	43,874,364	54,955,645
<i>M. leprae</i> exonic coverage ((reads x 50bp)/sum of length of all exons)	18.3X	10.8X	14.3X	34.0X	16.6X	65.5X	12.4X	15.2X	10.6X
<b>Percent <i>M. leprae</i> genes with ≥ 5 reads</b>	96%	93%	94%	98%	96%	98%	93%	94%	92%
<b>Total DESeq2 norm counts</b>	54,850,139	53,686,021	53,626,983	60,909,854	54,600,511	54,140,182	57,979,524	49,273,521	54,026,617
<b>Human DESeq2 norm counts</b>	53,723,695	53,179,282	52,935,117	58,510,300	53,830,465	51,751,697	57,208,427	48,431,719	53,507,157
<i>M. leprae</i> DESeq2 norm counts	1,126,445	506,739	691,866	2,399,554	770,046	2,388,485	771,098	841,802	519,460
<b>Human transcriptional abundance (Norm counts/gene length)</b>	9,719	12,807	12,923	24,013	17,262	11,787	17,071	10,691	13,900
<i>M. leprae</i> transcriptional abundance (Norm counts/gene length)	1,753	873	1,333	4,207	1,293	4,530	1,541	1,507	887
<b><i>M. leprae</i> total RNA abundance (log2(<i>M. leprae</i>/ (Human+<i>M. leprae</i>))+1))</b>	<b>0.15</b>	<b>0.06</b>	<b>0.09</b>	<b>0.15</b>	<b>0.07</b>	<b>0.28</b>	<b>0.08</b>	<b>0.12</b>	<b>0.06</b>
<b>Total <i>M. leprae</i> mRNA abundance (Norm counts/gene length)</b>	1,080	419	503	1,927	684	1,854	523	699	429
<b>23S (1343455-1343551) abundance (Norm counts/gene length)</b>	202	122	426	734	244	1,240	552	363	173
<b><i>M. leprae</i> mRNA:rRNA log2((mRNA/23S)+1))</b>	<b>2.67</b>	<b>2.15</b>	<b>1.12</b>	<b>1.86</b>	<b>1.93</b>	<b>1.32</b>	<b>0.96</b>	<b>1.55</b>	<b>1.80</b>

**Table S1. Related to Figure 1. Summary of in situ dual RNAseq of host-pathogen transcriptomes.** Raw read counts of exonic reads mapped to either the human or *M. leprae* genomes for each L-lep patient. Statistics of exome, gene coverage, and subsequent normalization also given. Specific values of the *M. leprae* abundance or mRNA:rRNA ratio are in bold font.

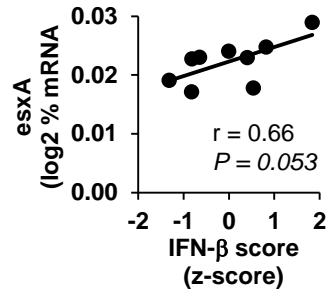
RANK	Gene ID (MLBr4923)	DESeq norm counts	Transcriptional Abundance (norm counts/gene length)	Gene Name (MLTN)	Description	Type
1	<b>MLBr_s02</b>	<b>177956</b>	<b>484.9</b>	<b>ssr</b>	<b>Transfer messenger RNA</b>	<b>ncRNA</b>
2	<b>MLBr_r2</b>	<b>117615</b>	<b>450.6</b>	<b>23S</b>	<b>23S ribosomal RNA</b>	<b>ncRNA</b>
3	<b>MLBr_s02</b>	<b>64270</b>	<b>154.1</b>	<b>RnaseP</b>	<b>Ribonuclease P</b>	<b>ncRNA</b>
4	MLBr_00049	6147	21.4	esxA	Probable 6 kDa early secretory antigenic target homolog EsxA (ESAT-6-like protein) (L-ESAT)	CDS
5	MLBr_00050	4817	16.0	esxB	Possible 10 KDA culture filtrate antigen homolog EsxB (lhp) (cfp10)	CDS
6	MLBr_00406	17102	10.1	espC	ESX-1 SECRETION-ASSOCIATED PROTEIN ESPC	CDS
7	MLBr_01795	3240	8.7	hsp18	18 KDA antigen Hsp18 (HSP 16.7)	CDS
8	MLBr_00804	2326	9.2	whiB1	Probable transcriptional regulatory protein WhiB-like WhiB1	CDS
9	MLBr_00198	3863	7.8	cspA	Probable cold shock protein A CspA	CDS
10	MLBr_02526	1588	5.2	ML2526	Conserved hypothetical protein (pseudogene)	CDS (pseudo)
11	MLBr_02686	446	4.9	ML2686c	Conserved hypothetical protein (pseudogene)	CDS (pseudo)
12	MLBr_02400	1251	4.1	ML2400c	PROBABLE CONSERVED TRANSMEMBRANE PROTEIN	CDS
13	MLBr_02042	2479	4.2	ahpC	ALKYL HYDROPEROXIDE REDUCTASE C PROTEIN AHPC (ALKYL HYDROPEROXIDASE C)	CDS
14	MLBr_00317	1529	3.3	groEL2	65 kDa chaperonin 2 GroEL2 (Protein Cpn60-2) (groEL protein 2) (65 kDa antigen) (Heat shock protein 65) (Cell wall protein A) (Antigen A).	CDS
15	MLBr_02135	3497	3.9	ML2135c	possible transmembrane protein	CDS
16	MLBr_01746	3962	4.2	ML1746c	Probable short-chain dehydrogenase/reductase (pseudogene)	CDS (pseudo)
17	MLBr_00405	4374	3.7	espA	Conserved hypothetical protein	CDS
18	MLBr_00380	531	3.1	groES	10 kD chaperonin GroES (Protein Cpn10) (groES protein) (BCG-A heat shock protein) (10 kDa antigen)	CDS
19	MLBr_00738	5298	3.6	pstA2	Phosphate-transport integral membrane ABC transporter PstA2 (pseudogene)	CDS (pseudo)
20	MLBr_00551	877	3.0	ML0551c	Putative methyltransferase	CDS
21	MLBr_01852	930	3.1	ML1852c	Probable conserved transmembrane protein (pseudogene)	CDS (pseudo)
22	MLBr_00410	928	3.1	PE3	PE-family protein	CDS
23	MLBr_01331	2504	2.5	tatA	PROBABLE SEC-INDEPENDENT PROTEIN TRANSLOCASE	CDS
24	MLBr_01420	954	2.7	ML1420	MEMBRANE-BOUND PROTEIN TATA hypothetical protein	CDS
25	MLBr_02038	3224	2.5	bfrA	PROBABLE BACTERIOFERRITIN BFRA	CDS
26	MLBr_01961	727	2.5	rpmJ	PROBABLE 50S RIBOSOMAL PROTEIN L36 RPMJ	CDS
27	MLBr_02563	281	2.4	fadE5	PROBABLE ACYL-CoA DEHYDROGENASE FADE5	CDS
28	MLBr_00407	1178	2.3	espD	ESX-1 SECRETION-ASSOCIATED PROTEIN ESPD	CDS
29	MLBr_00411	4420	2.6	ML0411	serine-rich antigen	CDS
30	MLBr_00938	1526	2.3	ML0938	hypothetical protein	CDS

**Table S2. Related to Figure 2. Top 30 *M. leprae* genes expressed in L-lep skin lesions.** Most abundantly expressed *M. leprae* genes as calculated by mean expression across the nine L-lep samples divided by gene length. Noncoding RNA (ncRNA) genes are in bold font. Fractional transcriptional abundance calculated by divided the mean expression of gene by total abundance of all *M. leprae* genes.

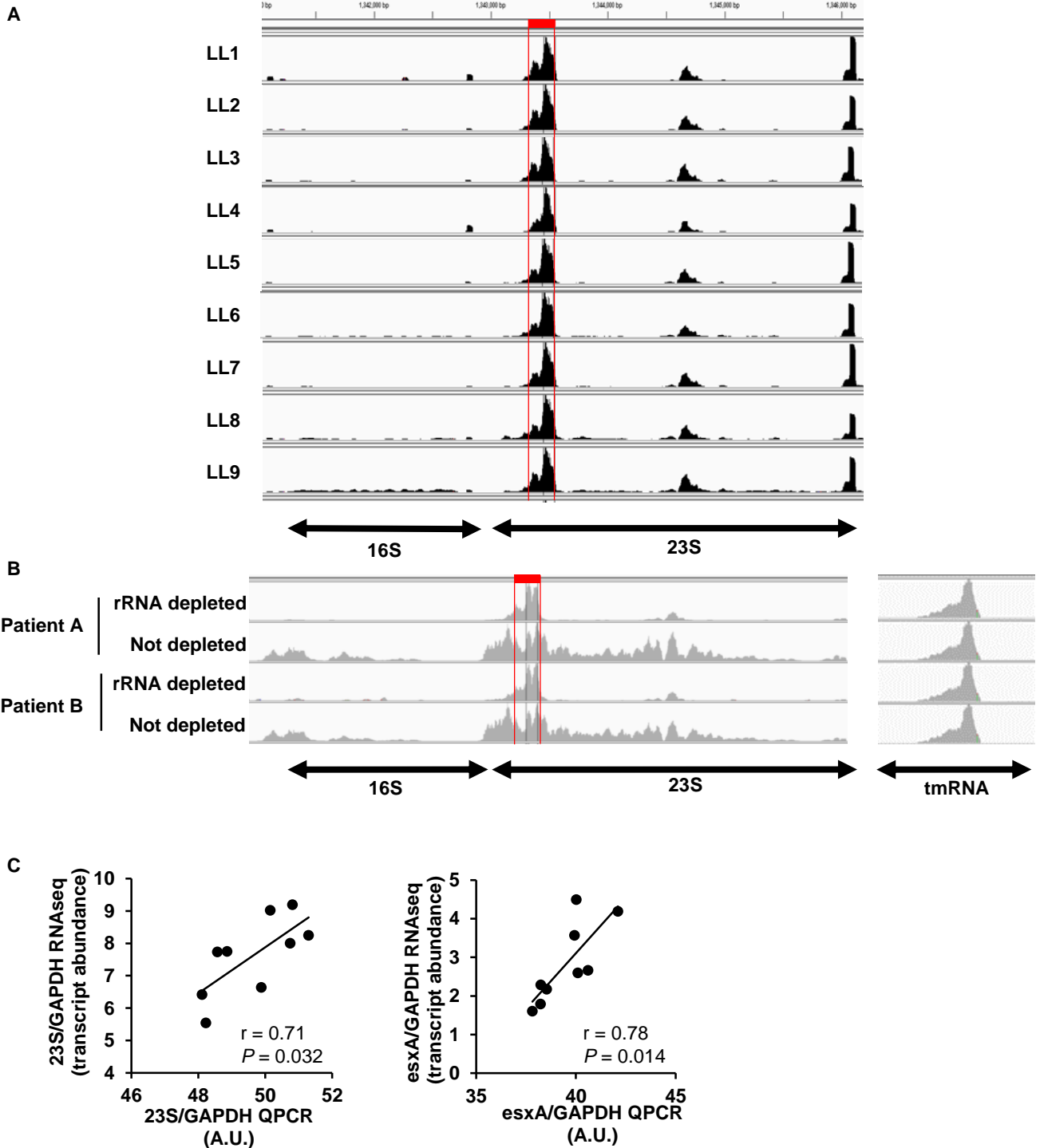
A



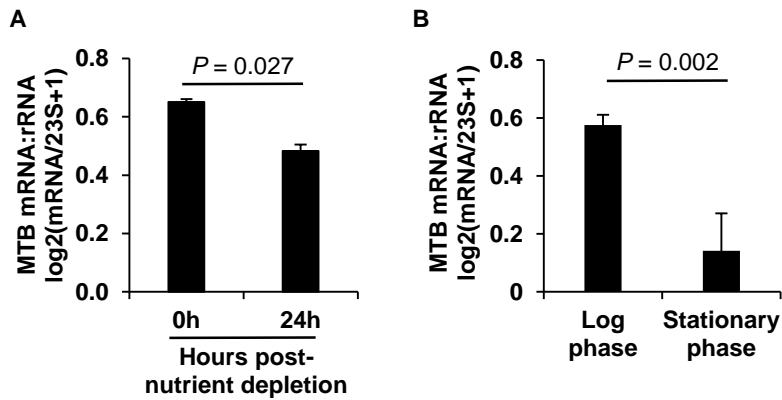
B



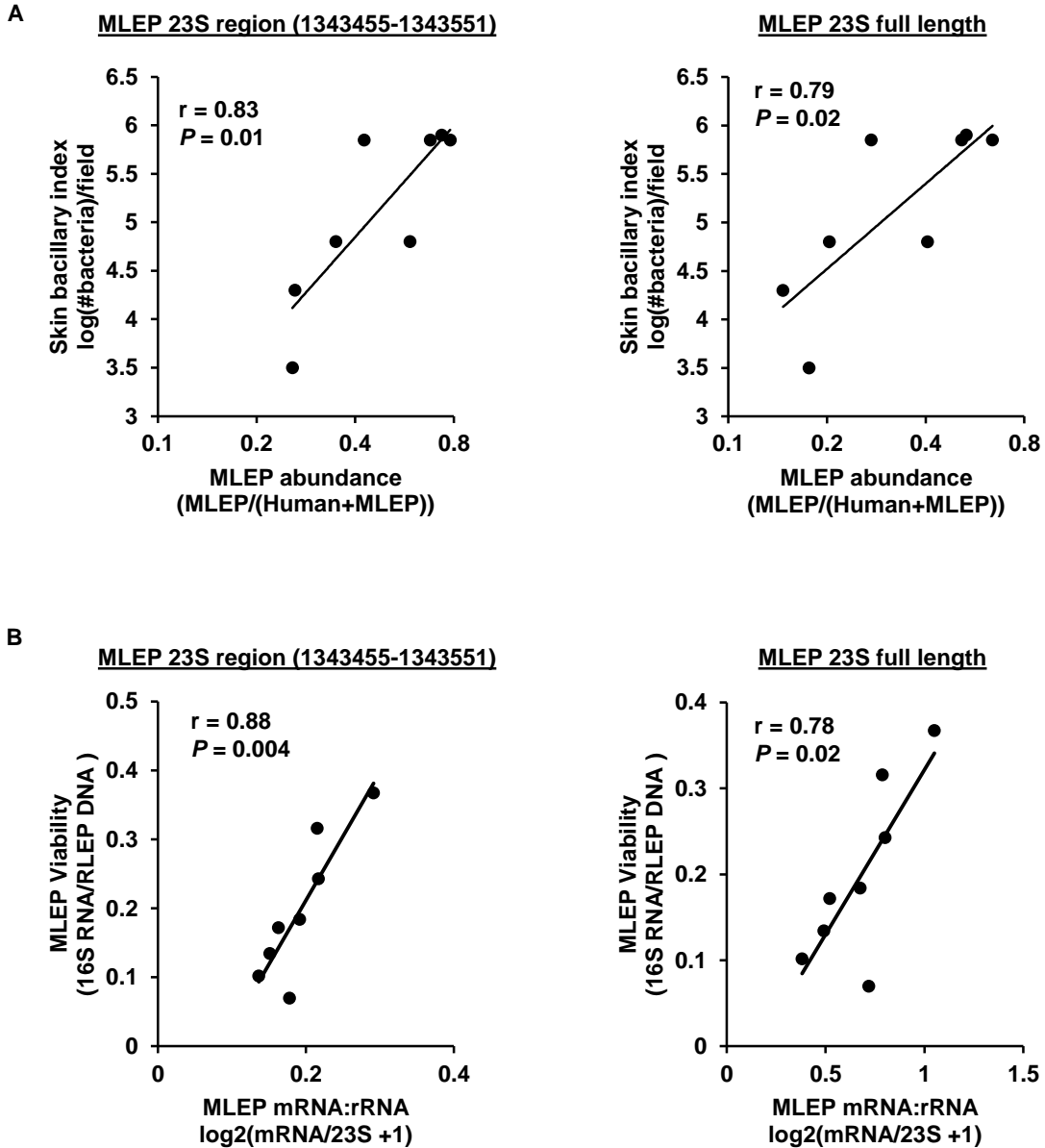
**Figure S1. Related to Figure 1 & Figure 3. Relationship of host macrophage abundance and type I interferon program to *M. leprae* expression.** (A) Correlation plot of *M. leprae* abundance against SaVanT macrophage signature score per L-lep patient. *P*-value by student's t-test ( $n = 9$ ). (B) Correlation plot of the host type I interferon SaVanT signature score versus the percent mRNA expression of *M. leprae* esxA per L-lep patient. *P*-value by student's t-test ( $n = 9$ ).



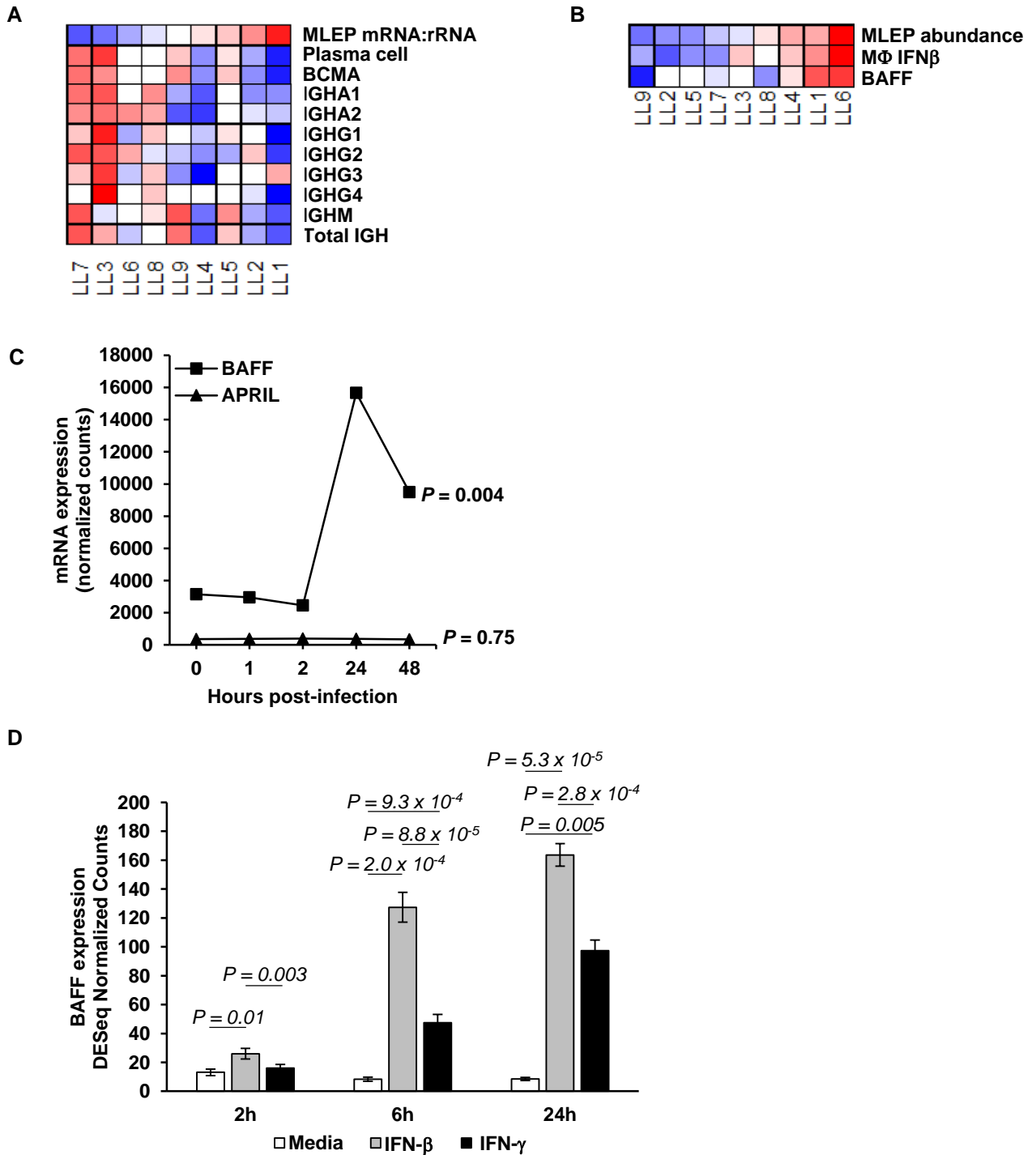
**Figure S2. Related to Figure 2. Read coverage of *M. leprae* ribosomal RNA locus. (A)** Dual RNAseq coverage of reads across 16S and 23S ribosomal RNA loci of *M. leprae*. Each Integrated Genome Browser track (IGV) autoscaled to each sample. Selected region for 23S rRNA quantitation highlighted in red. **(B)** RNA was extracted from two separate leprosy skin lesions, split into two aliquots, one depleted for bacterial rRNA by Ribozero while the other lacked depletion. Read coverage for rRNA locus shown with each IGV track autoscaled to each sample. **(C)** RNA from L-lep samples before rRNA depletion was converted into cDNA libraries. *M. leprae* 23S rRNA or *esxA* mRNA expression of undepleted cDNA assessed via QPCR, normalized to human GAPDH expression (x-axis) and compared to 23S/GAPDH or *esxA*/GAPDH RNAseq expression of rRNA depleted samples. *P*-value by student's *t*-test ( $n = 9$ ).



**Figure S3. Related to Figure 2. Ratio of mRNA:rRNA assessed in *Mycobacterium tuberculosis*.** (A) *M. tuberculosis* H37Rv was grown for 0 or 24h in axenic culture in PBS, without any additional nutrients. RNAseq was performed and mRNA:rRNA ratio of the total abundance of mRNA to the 23S rRNA was measured. Results shown as mean  $\pm$  s.e.m ( $n = 3$ ). (B) Similarly, *M. tuberculosis* H37Rv in log phase versus stationary phase RNAseq mRNA:rRNA ratio was compared. Results shown as mean  $\pm$  s.e.m ( $n = 3$ ).  $P$ -value by student's t-test.



**Figure S4. Related to Figure 1 and Figure 2. Validation of *M. leprae* bacterial measures with current clinical standards.** An additional eight lepromatous patient lesions were sequenced by dual RNA sequencing method, however only human ribosomal was depleted. MLEP abundance (**A**) and MLEP transcriptional activity (**B**) calculated by quantifying the 23S rRNA abundance either by the 23S region (1343455-1343551) or using its full length, were correlated with the skin bacillary or *M. leprae* 16S RNA/RLEP DNA viability per patient, respectively. *P*-value by student's *t*-test ( $n = 8$ ).



**Figure S5. Related to Figure 4. *M. leprae* correlation and direct effect on the host humoral immune response.** Heatmap of the expression human humoral immune components per each of the nine L-lep patients, ordered by *M. leprae* transcriptional activity (A) or *M. leprae* abundance (B). (C) Human M-CSF derived macrophages (MDMs) were infected at a multiplicity of infection of 10 with *M. leprae* derived from the footpad of *nu/nu*. At time indicated, RNA was extracted and RNAseq was performed. Data shown as normalized counts across timecourse. Bonferoni-adjusted *P*-value for change of gene expression one experiment over the five timepoints calculated by Wald test. (D) MDMs stimulated with IFN-β or IFN-γ for 2, 6, or 24h then RNAseq was performed. Data shown as BAFF mRNA normalized counts across timecourse. *P*-value ( $n = 5$ ) by paired student t-test.