LEPROSY RESEARCH INITIATIVE

Turing Foundation

Genomic markers for pathological variants and transmission of leprosy bacilli

Pushpendra Singh[1], Itu Singh[2] Purna Dwivedi[1], Rahul Sharma[3,4], Richard W Truman[3,4]

[1] **ICMR-National Institute of Research in Tribal Health,** Jabalpur, MP (India)

[2] **Stanley Browne Laboratory (SBL),** New Delhi-India

[3] **Hansen's Disease Program (NHDP),** Baton Rouge LA – USA

[4] **LSU, School of Veterinary Medicine (LSU-SVM),** Baton Rouge LA-70803 USA

r2STOP
Research to STOP neglected tropical disease transmission.

Hypothesis and Objectives

- It was hypothesized that comprehensive genomic information will identify new informative markers for local transmission and bacterial genomic variations associated with pathological differences of the disease.
- In this project, we aim to develop a cost-effective bait system for the enrichment of M. leprae DNA and apply it for sequencing of clinical specimens representing broad spectrum of the disease and geographical origin. This will help in identifying the genomic markers suitable to monitor local transmission of infection and pathological variations of leprosy bacilli.
- The conventional genotyping methods use SNP identification by conducting PCR followed by Sanger sequencing. In the present study we report a successful method that bypasses Sanger sequencing for *M. leprae* genotyping to identify 1D and 3I, the most prevalent genotypes in the world.

RESEARCH OBJECTIVES / OUTCOMES:

A. **To develop cost effective bait system for enrichment of M. leprae DNA from patient samples and optimize WGS technologies for use across the leprosy spectrum.**

B. **Increase the number of *M. leprae* genomes in public domain to represent the complete spectrum of the disease with diverse geographical distribution.**

C. **To identify the genomic markers for local transmission and pathological variations in leprosy.**

Research pursued in the direction of:

Identification of unique SNPs and development of simple tools for their detection

- Sanger sequencing based SNP confirmation
- Development of Restriction enzyme based Genotype detection and drug resistance detection

Preparation of NGS libraries for bait based selection metgod for whole genome sequencing of pathogen

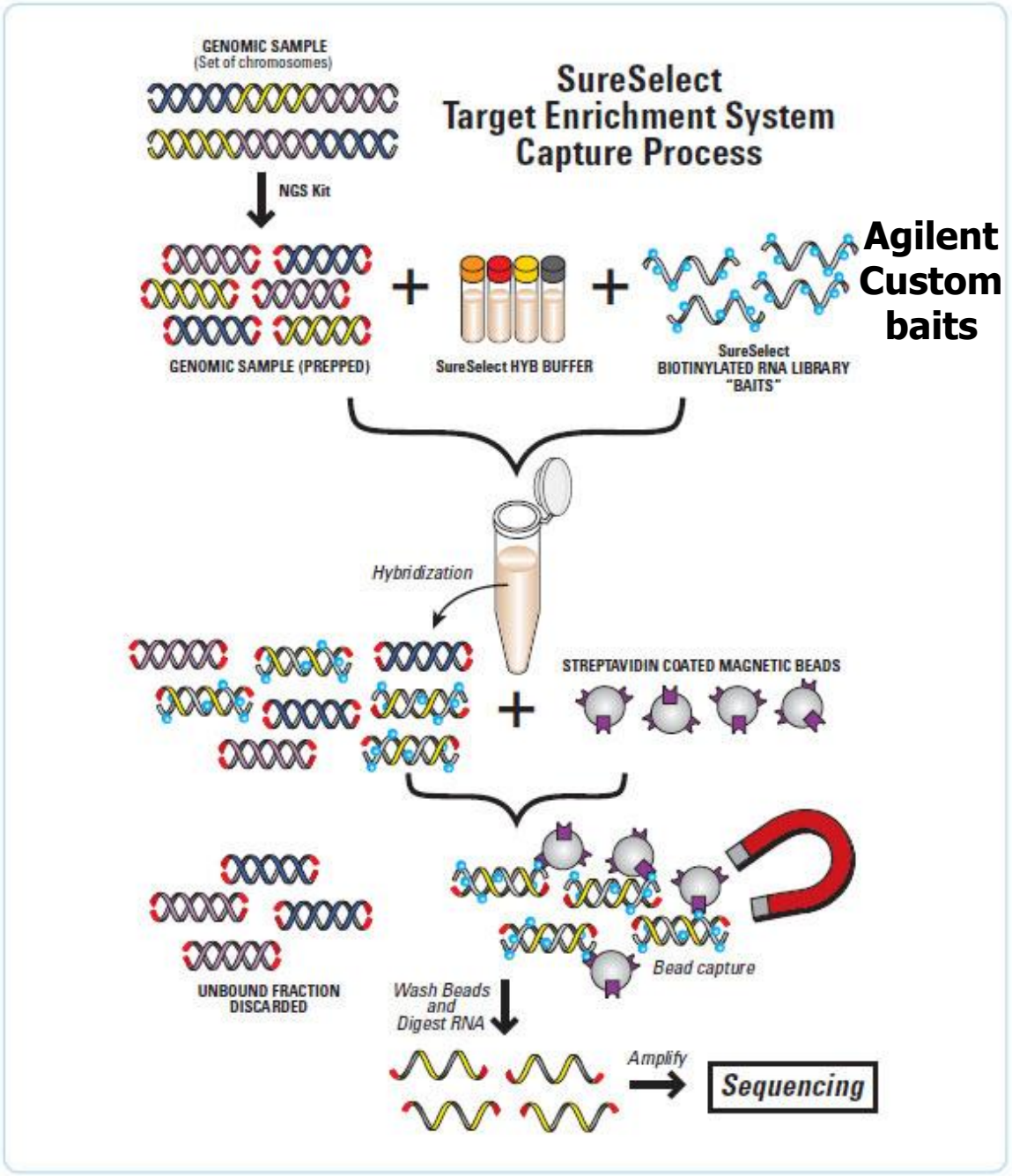
r2STOP
Research to STOP neglected tropical disease transmission.

Enrichment using Commercial bait :

Agilent baits for Enrichment of *M. leprae* DNA from patient biopsies

Expensive method (as each time new baits have to be custom-synthesized)

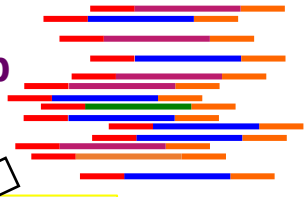
Cost=$11K/16 rxn v!



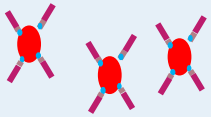
New methods for Genomics w/o cultivation: Enrichment of pathogen DNA from patient biopsies

Step 1.
Genomic DNA
(host + *M. leprae*)

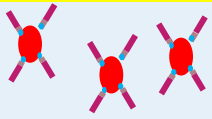
Step 2:
Library prep



Biotinylated *M. leprae* baits on magnetic streptavidin beads

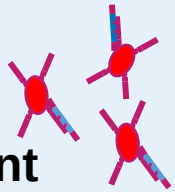


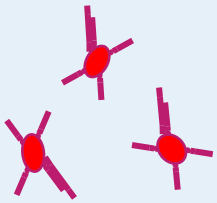
Step 3:
Hybridization capture



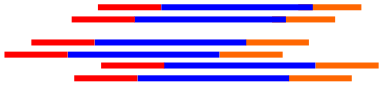
Biotinylated H.A.M. DNA on magnetic streptavidin baits

Positive Enrichment





Negative Enrichment



Step 4:
Enriched target for NGS

For metagenomic studies, for targeted sequencing, exome capture

r2STOP
Research to STOP neglected tropical disease transmission.

Preparation of baits

***M. leprae* DNA**

- From MFP derived *M. leprae*

Preparation of Bait library

- Adapter ligation (**Ion Proton**)
- Amplification with biotinylated primers

Heat Denaturation

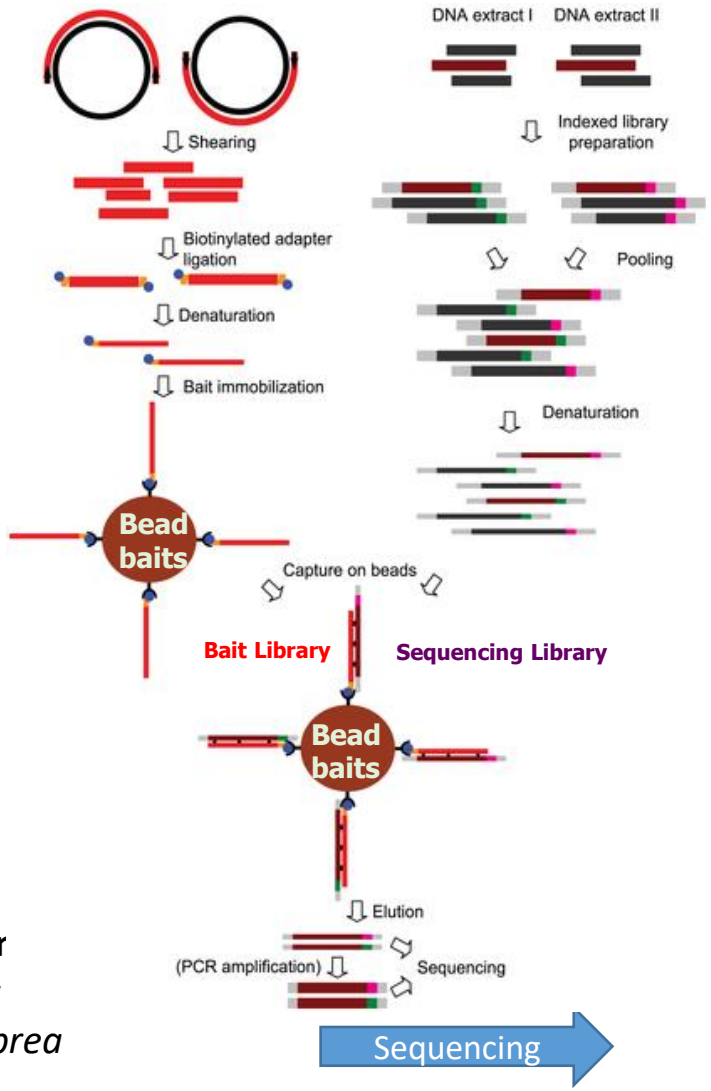
Bait Immobilization

- On Streptavidin bead (Magnetic beads)

Hybridization Capture (Enrichmer

- Bait Library + Sequencing Library
- Positive Selection (Collect *M. leprea* DNA)
- Negative selection (Remove host DNA)

Elution and amplification



Processing of sample for sequencing

- **DNA Extraction**
- **Sequencing Library Preparation**
- adaptor Ligation **(Illumina)**
- Amplification
- Profile and normalization
- sequencing

Multiplex sequencing
MiSeq (illumina)

r2STOP
Research to STOP neglected tropical disease transmission.

Majority of the *M. leprae* strain in India belong to ID genotype in the Indian sample subset.

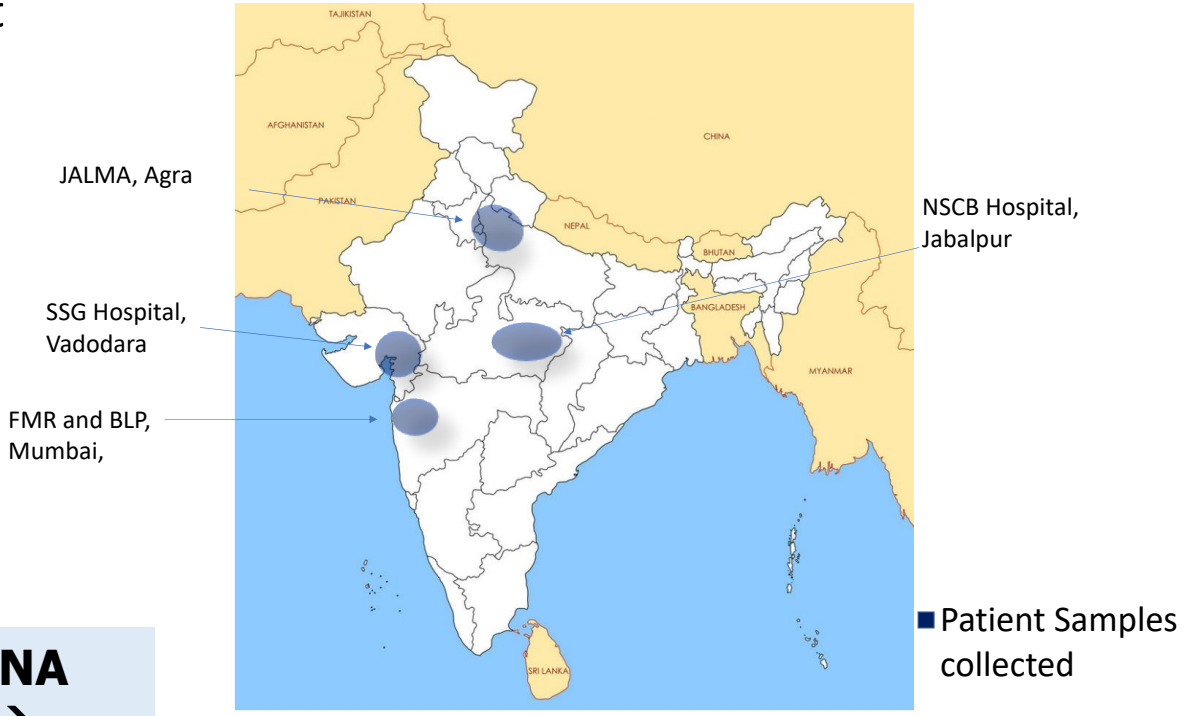
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	# SNPs and InDels identified in 154 *M. leprae* genomes.												
2	SNP type ->												
3	POS (AL450	REF (TN)	ALT	SNP or	EFF	Feature t	Gene na	ML code	Feature length	Codon_change	A.acid_change	Comment	Number
806	639580	A	G	SNP	intergenic_region	intergenic	ML0526c-ML	ML0526c-ML0527c	698	639580A>G		1D	9
2949	3016895	C	A	SNP	missense_variant	CDS	eccC3	ML2535c	3990	3541G>T	Val1181Leu	1D	9
3256	3262657	C	T	SNP	synonymous_variant	CDS	parB	ML2706c	1008	459G>A	Leu153Leu	1D	9
3279													
3280													

(Benjak A, Avanzi C, Singh P *et al*, *Nature Communications* 2018)

r2STOP
Research to STOP neglected tropical disease transmission.

Obj III: Pathogen Genomics for identifying phylogeographic markers for genotyping

- n=350 samples from different areas.
- Genotyping and drug resistance surveillance study.
  - SNP-type 1D >90%
  - Drug resistance
    - is <1% for Rif (1st line)
    - 10% for Oflo (2nd line)



Sites of Leprosy clinical samples collected from India

Enrichment of *M. leprae* DNA by host depletion method→ WGS → Bioinformatics analysis

M. leprae Genotypes reported from India using PCR-based studies

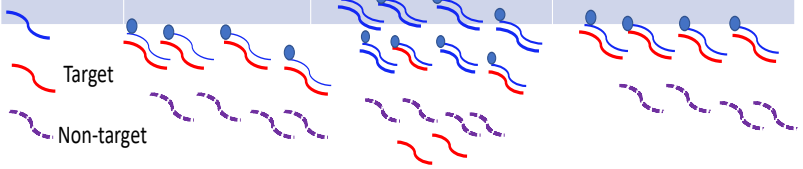
Genotype	Kuruwa et al. 2011 (Maharashtra, n=109)	Lavania et al. 2013 (Delhi, UP,WB, n=180)	Das et al. 2016 (South India, n=160)	Combined data (n=389)
1A	-	-	3.12%	1.28%
1B	6.25%	-	-	0.77%
1C	-	-	15.62%	6.42%
1D	93.75%,	92.2%	55%	**76.86%**
2E	-	5%	-	2.31%
2G	-	3.33%	20.62%	10.2%
2H	-	-	5.62%	2.31%

Development of *M. leprae* DNA Enrichment methods

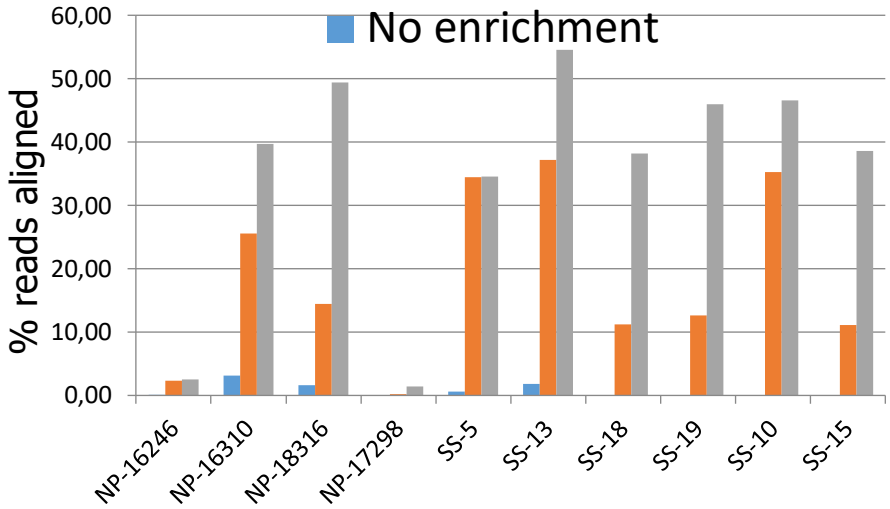
Preparation of PCR-amplifiable biotinylated bait-Library:

-For capture of target DNA (*M. leprae* DNA)
-reduction of human DNA

Baits	Agilent Sureselect baits	PCR-biotin baits (biotin on both primers) LepBaits-v1	PCR-biotin (on one strand only) LepBaits-v2
Nature of baits	**Synthetic single-stranded RNA** (120 bp probes that bind to target DNA	Library of purified *M. leprae* DNA (renewable by PCR), **Double stranded, propensity of self-binding**	v1 baits enriched using synthetic baits **to remove host DNA & the self-complementary strand**
Cost	Very expensive ($700/rxn)	Inexpensive (PCR amplifiable)	Inexpensive (PCR-amplifiable)
Efficiency	★★★★	★★	★★★★



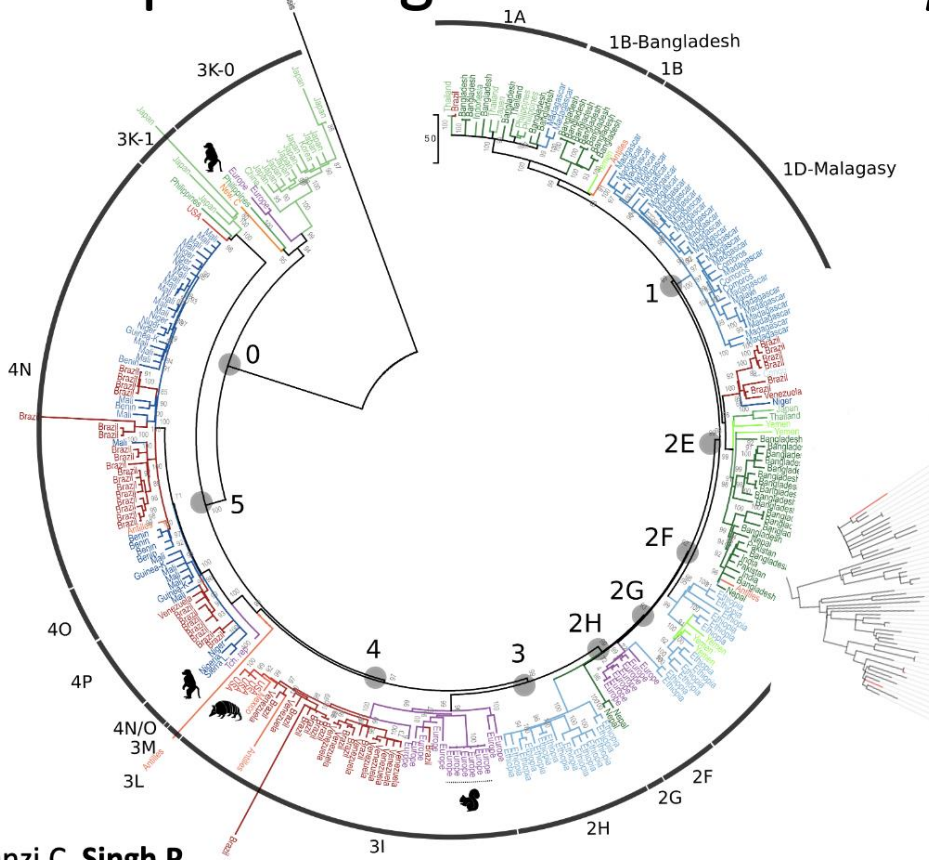
After enrichment, more reads align to *M. leprae*

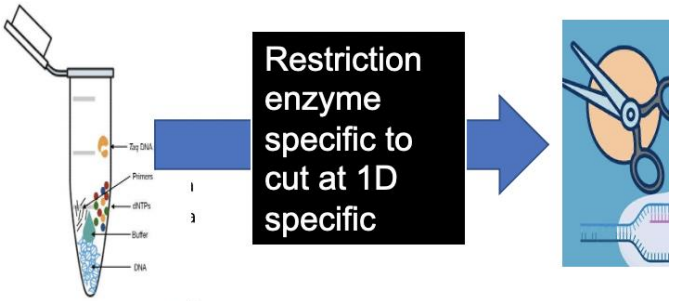


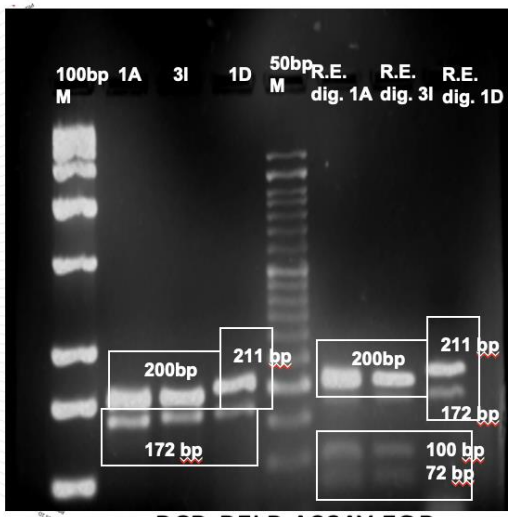
Results

- Nearly 20 NGS libraries were prepared for whole genome sequencing.
- The whole genome sequencing of *M. leprae* by bait enrichment method is being carried out and the WGS data is awaited for the bioinformatic analysis and subsequent comparative genome analysis.
- The PCR-RFLP experiment garnered encouraging results. Due to difference of 11 bp between 1D and 3I genotypes at the SNP 17915 and restriction digestion separate bands are observed on gel visualization. By following this method, we would be able to identify and differentiate 1D, 3I genotypes at the same time by conducting just one PCR followed by restriction digestion of the PCR amplicons.

Comparative genomics of *M. leprae* strains







PCR-RFLP ASSAY FOR GENOTYPING

Avanzi C, **Singh P,** Truman, Suffys PN. **Inf Gen Evol 2020**

Summary of the *M. leprae* Genomics work

- PCR-amplifiable baits are **cost-effective methods for WGS of uncultivated** pathogens.
- Developed a PCR-RFLP based assay for **simultaneous identification and genotyping** of *M. leprae* strains prevalent in Indian subcontinent.
- Sequencing independent differentiation of *M. leprae* and *M. lepromatosis.*
-

Acknowledgments

Ramalingaswami Re-entry Fellowship
Department of of Biotechnology, Govt of India
-Dr Marivick Balagon , Mr Nirmal Jha & all other team members

Thanks for your attention and feedback.

ICMR-NIRTH Jabalpur
MSU Baroda
SBL, TLMTI, New Delhi
NJILOMD Agra
NHDP USA.

For clinical samples:
Foundation of Medical Research, Mumbai.
Bombay Leprosy Programme, Mumbai.
ICMR-JALMA, Agra.
SSG Hospital, Vadodara.
NSCB Hospital, Jabalpur.

Funding agencies:
LRI & Turing Foundation
R2STOP Canada
Seeding Labs Boston, USA

LRI
Netherlands

Publications

1. Sharma R*, **Singh P***, McCoy RC, Lenz SM, Donovan K, Estrada-Garcia OI, Silva-Miranda M, Jurado F, Cruz S, Balagon MF, Stryjewska B, Scollard DM, Pena MT, Lahiri R, Williams DL, Truman RW, Adams LB. Isolation of *Mycobacterium lepromatosis* and development of molecular diagnostic assays to distinguish *M. leprae* and *M. lepromatosis*. *Clinical Infectious Diseases* 2020, **(**https://doi.org/10.1093/cid/ciz1121**)** (*_**Joint first author**_) **(IF=9.055).**
2. Sharma R, **Singh P**, Pena M, R Subramanian, V Chouljenko, J Kim, N Kim et al; Differential growth of *M. leprae* strains (SNP genotypes) in armadillos. Infection, Genetics and Evolution 2018; 62:20-26. (IF= 2.545)
3. **Singh P***, Tufariello J, Wattam AR, Gillis TP, and Jacobs WR. Genomics insights into the biology and evolution of leprosy bacilli. In: International Textbook of Leprosy 2017, Eds: Thomas Gillis and David Scollard. (Corresponding author and first author).
4. Avanzi C, **Singh P**, Truman RW and Suffys PN. Molecular epidemiology of leprosy: an update. Infection, Genetics and Evolution (in press): MEEGID-D-20-00679).
5. Sharma M, **Singh P***. Use of Artificial Intelligence in Research and Clinical Decision Making for combating Mycobacterial Diseases. In: Artificial Intelligence and Machine Learning in Healthcare Perspective. 2020. Ed(s) Saxena A, Chandra S. Springer Publishers.
6. Benjak A, Avanzi C, **Singh P**, .., Cole ST. “Phylogenomics and antimicrobial resistance of the leprosy bacillus Mycobacterium leprae" Nature Communications 2018; 9(1):352-62. (IF 12.124).
7. Sharma M, Gupta Y, Dwivedi P, K Prakasha, **Singh P***. Mis-annotation of the Genes Encoding coproporphyrinogen III oxidases in Mycobacteria.
8. Chowdhury S, Pena M, Sharma R, **Singh P***. Identification of biomarkers in leprosy affected host signaling pathways using Systems Biology approach
9. Sharma M, Dwivedi P, **Singh P***. Current situation of leprosy in tribal areas of India in the post-elimination era. (in press).