

# Transcriptome Analysis of *Mycobacterium leprae*

By: Isabella Fregoso

# What is a transcriptome analysis?



---

Analysis of mRNA transcripts to see the gene expression patterns

*How are the genes expressed?*



---

Quantifies the entire mRNA transcript to tell us about the cellular activity



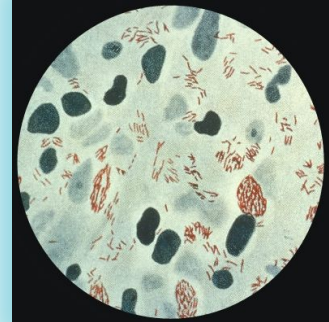
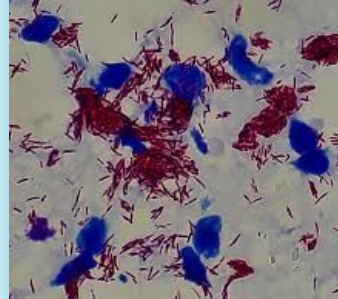
---

RNA-seq used for non coding regions and to identify splice variants

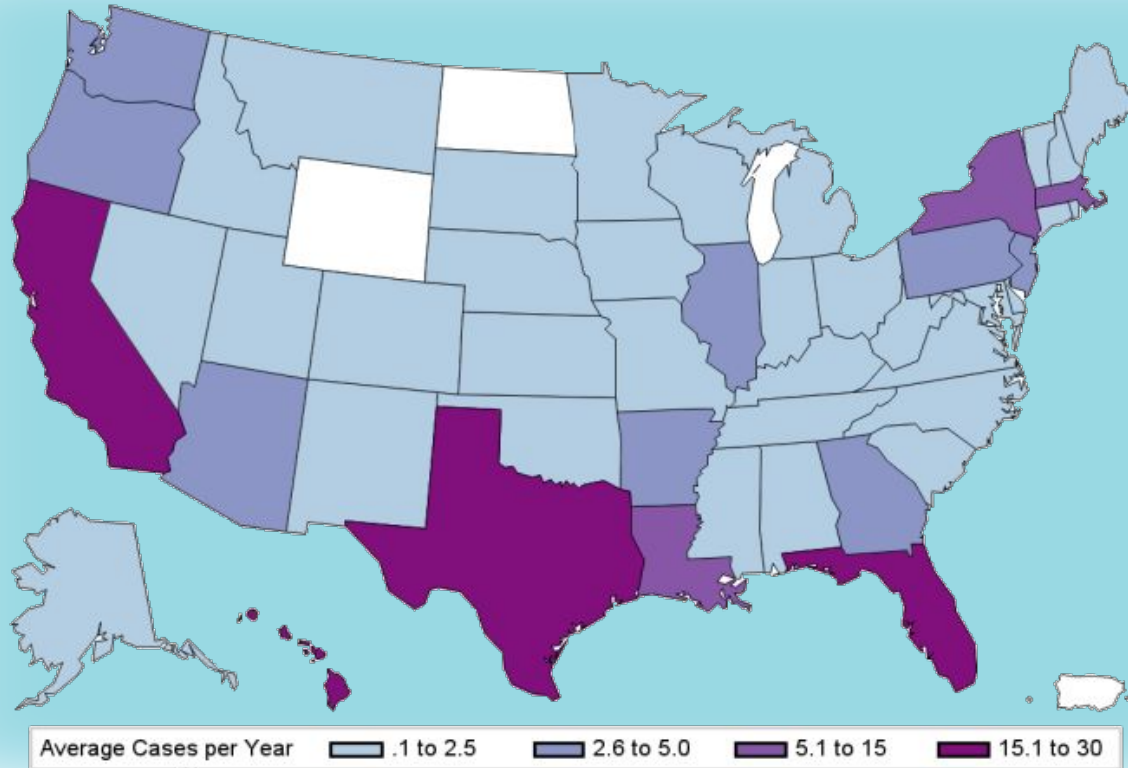
**Splice variant**= different variant of mRNA from same gene, but spliced in a different way

# Introduction

- Acid-fast, rod shaped, gram-positive, obligate intracellular bacterial parasite
- Targets peripheral nerves, epithelial mucosa, eyes, and skin
- Can we map the genes, that code for leprosy?
- Start with gff, fna, and fasta and 2 fastq.gz for each sample



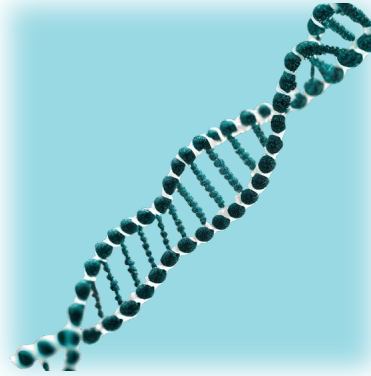
### 10-Year Case Averages by Jurisdiction



Average number of cases of leprosy per year in US (2004–2013) as reported by US HHS

# Methods

- Create index files  
Creates a structured table of contents
- Fastq to sam to bam
- Raw seq data>seq alignment map>  
compressed binary format
- Samtools sort  
Organizes chronologically
- Stringtie assemblies  
RNA seq reads into transcripts

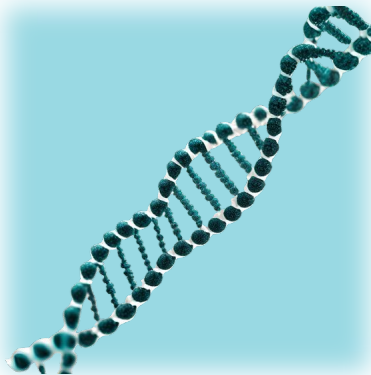


```
.fna  
-----  
.fastq  
.bam  
.sam
```



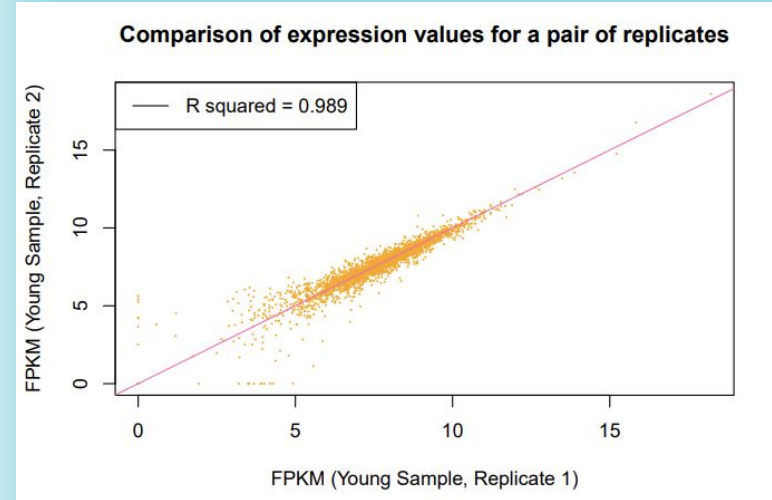
# Methods Cont...

- Stringtie --merge  
Merge transcripts together in single string
- Cat stringtie\_merged.gff  
Counting the transcripts
- Gffcompare -r mleprae.gff  
Compares annotated to known
- Cat merged.status  
Checking the quality



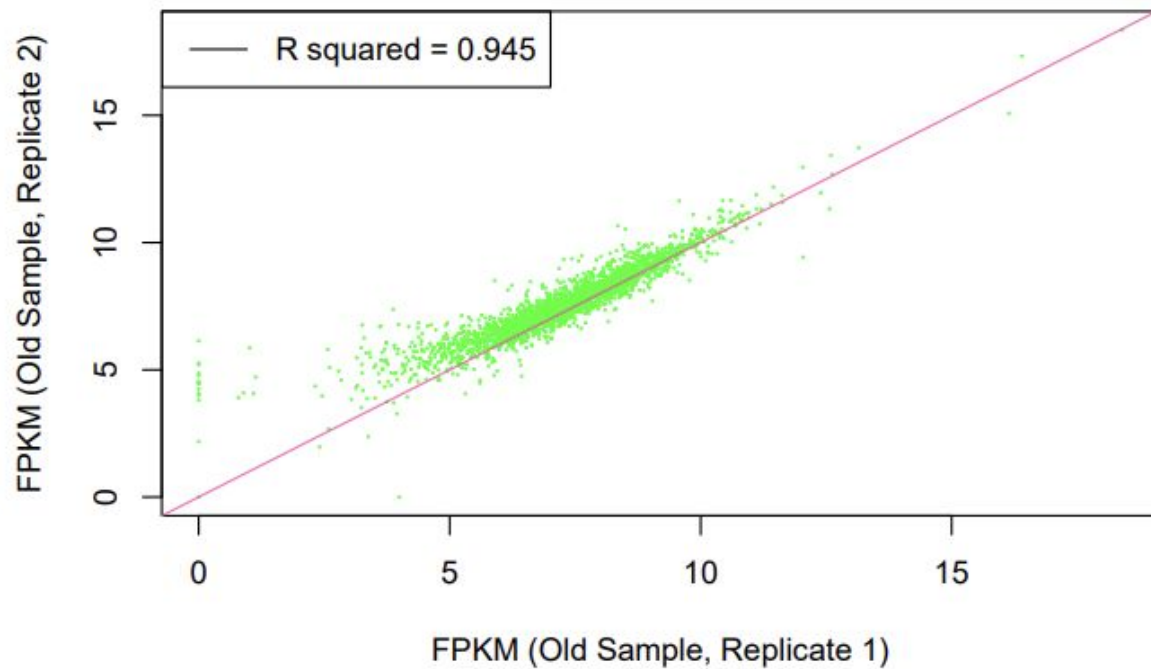
# Results

- $R^2$  closer to one indicates accuracy between predicted and actual expression
- FPKM is Fragment Per Kilobase of transcript per Million mapped reads
- Strong agreement between the replicates; highly consistent and reliable





### Comparison of expression values for a pair of replicates

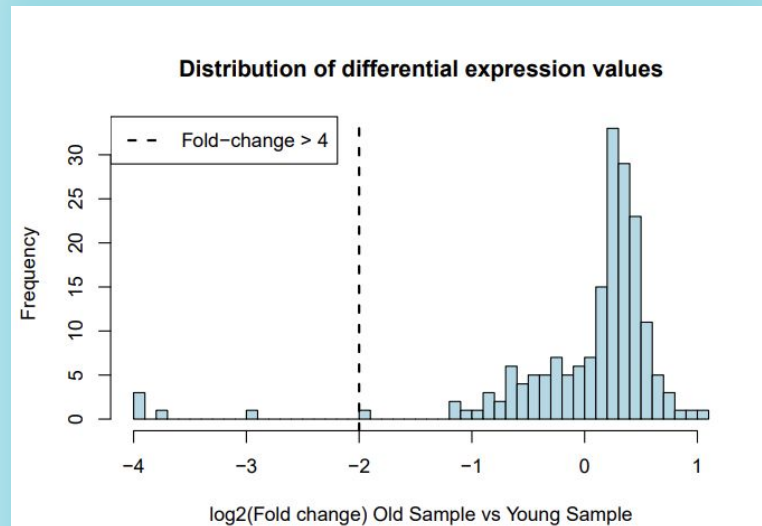




# Results

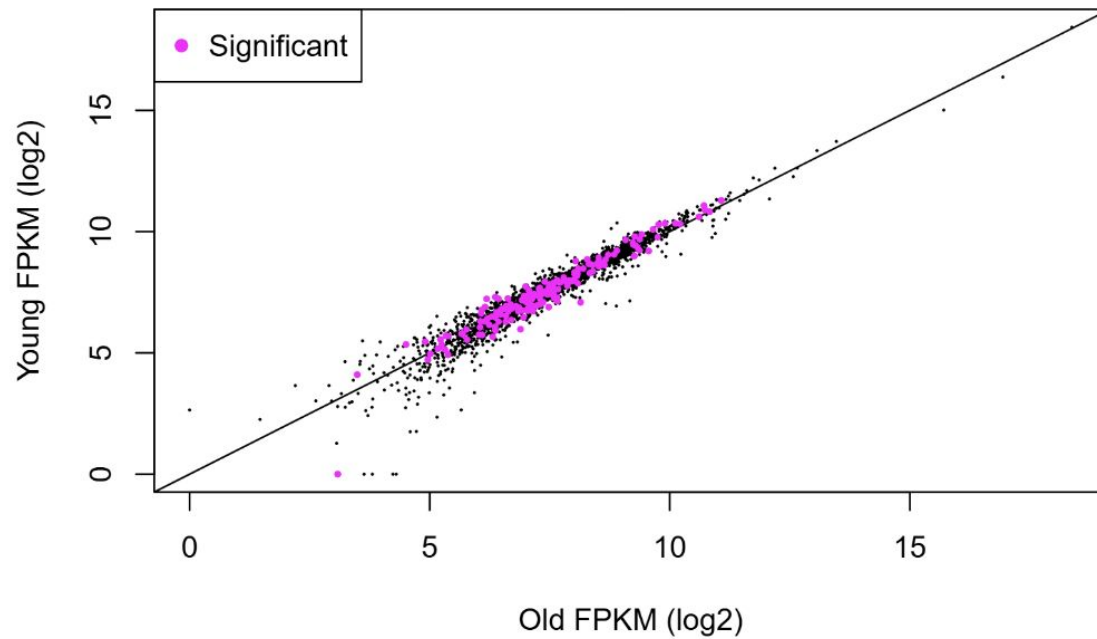
- This depicts the change in expression between two samples
- Pos fold change means inc in expression, neg fold change means dec in expression

- 
- Black is not significantly differentially expressed
  - Genes above the line are higher expressed in the young, below the line are higher expressed in the old



- Few outliers which indicates significant up/downregulation

## Old vs Young FPKMs



| geneNames | transcriptNames    | id   |
|-----------|--------------------|------|
| .         | gene-DIJ64_RS18560 | 1840 |
| .         | gene-DIJ64_RS21400 | 1343 |
| .         | gene-DIJ64_RS04330 | 887  |
| .         | gene-DIJ64_RS21230 | 929  |
| .         | gene-DIJ64_RS01285 | 249  |
| .         | gene-DIJ64_RS01500 | 289  |
| .         | gene-DIJ64_RS21165 | 831  |
| .         | gene-DIJ64_RS10190 | 2115 |
| .         | gene-DIJ64_RS22250 | 2934 |
| .         | gene-DIJ64_RS18390 | 1644 |
| .         | gene-DIJ64_RS15665 | 2375 |
| .         | gene-DIJ64_RS06005 | 1234 |
| .         | gene-DIJ64_RS00940 | 178  |
| .         | gene-DIJ64_RS07865 | 1619 |
| .         | gene-DIJ64_RS21750 | 1989 |
| .         | gene-DIJ64_RS02640 | 534  |
| .         | gene-DIJ64_RS11485 | 2385 |
| .         | gene-DIJ64_RS14390 | 3001 |
| .         | gene-DIJ64_RS07750 | 1597 |
| .         | gene-DIJ64_RS11330 | 2352 |
| .         | gene-DIJ64_RS08195 | 1689 |
| .         | gene-DIJ64_RS18345 | 1612 |
| .         | gene-DIJ64_RS15070 | 311  |
| .         | gene-DIJ64_RS00270 | 50   |

#### List of genes in the transcripts

|        |  |
|--------|--|
| eccCa  | secretes virulence factors into environment            |
| coaA   | mycolic acid synthesis (general development)           |
| sucB   | synthesize peptidoglycan                               |
| eccA   | part of bacterial secretion pathway                    |
| Pgi    | sugar metabolism and virulence factor                  |
| rpmG   | encodes ribosomal protein S16                          |
| rpmE   | involved in bacterial gene reg and chromatin structure |
| eccCa  | bacterial protein secretion                            |
| uvrC   | nucleotide excision repair (mut prevention)            |
| Fgd    | G nucleotide exchange factor/ structure                |
| kasB   | mycolic acid synthesis                                 |
| ppc    | replenishes kreb's cycle                               |
| fadD26 | binding fatty acids                                    |
| coaE   | aids in CoA production                                 |
| prcB   | encodes parts of the proteasome core                   |
| thrC   | encodes for threonine (amino acid) synthase            |
| era    | encodes GTPase protein for cell cycle                  |
| typA   | regulates virulence and pathogenicity                  |
| regX   | two-component sensory transduction protein             |
| mhuD   | encodes heme oxygenase enzyme to degrade heme          |
| hisD   | encodes for histidine                                  |
| rfaA   | encodes RmlA helpful for polysaccharides production    |
| sucC   | encodes beta subunit for succinyl-CoA                  |