

Supporting information for Brosch *et al.* (March 12, 2002) *Proc. Natl. Acad. Sci. USA*,

10.1073/pnas.052548299

**Table 1: RD, RvD and TbD1 regions**

| Region absent from BCG | Gene                | Size kb | Internal primer pair                            | Flanking primers or 2 <sup>nd</sup> internal* primer pair |
|------------------------|---------------------|---------|-------------------------------------------------|-----------------------------------------------------------|
| RD1                    | Rv3871-Rv3879c      | 9.5     | RD1in-Rv3878F<br>GTC AGC CAA GTC AGG CTA CC     | RD1-flank.left<br>GAA ACA GTC CCC AGC AGG T               |
|                        |                     |         | RD1in-Rv3878R<br>CAA CGT TGT GGT TGT TGA GG     | RD1-flank.right<br>TTC AAC GGG TTA CTG CGA AT             |
| RD2                    | Rv1978-Rv1988       | 10.8    | RD2-Rv1979.int.F<br>TAT AGC TCT CGG CAG GTT CC  | RD2-flank.F<br>CTC GAC CGC GAC GAT GTG C                  |
|                        |                     |         | RD2-Rv1979.int.R<br>ATC GGC ATC TAT GTC GGT GT  | RD2-flank.R<br>CCT CGT TGT CAC CGC GTA TG                 |
| RD3*                   | Rv1573-Rv1586c      | 9.2     | RD3-Rv1586.int.F<br>TTA TCT TGG CGT TGA CGA TG  | RD3-int-REP.F<br>CTG ACG TCG TTG TCG AGG TA*              |
|                        |                     |         | RD3-Rv1586.int.R<br>CAT ATA AGG GTG CCC GCT AC  | RD3-int-REP.R<br>GTA CCC CCA GGC GAT CTT*                 |
| RD4                    | Rv1505c-Rv1516c     | 12.7    | RD4-Rv1516.int.F<br>CAA GGG GTA TGA GGT TCA CG  | RD4-flank.F<br>CTC GTC GAA GGC CAC TAA AG                 |
|                        |                     |         | RD4-Rv1516.int.R<br>CGG TGA TTC GTG ATT GAA CA  | RD4-flank.R<br>AAG GCG AAC AGA TTC AGC AT                 |
| RD5*                   | Rv2346c-Rv2353c     | 9.0     | RD5A-Rv2348.int.F<br>AAT CAC GCT GCT GCT ACT CC | RD5B-plcA.int.F<br>CAA GTT GGG TCT GGT CGA AT             |
|                        |                     |         | RD5A-Rv2348.int.R<br>GTG CTT TTG CCT CTT GGT C  | RD5B-plcA.int.R<br>GCT ACC CAA GGT CTC CTG GT             |
| RD6*                   | Rv3425-Rv3428c      | 4.9     | RD6-IS1532F<br>CAG CTG GTG AGT TCA AAT GC       | ND                                                        |
|                        |                     |         | RD6-IS1532R<br>CTC CCG ACA CCT GTT CGT          | ND                                                        |
| RD7                    | Rv1964-Rv1977       | 12.7    | RD7-Rv1976.int.F<br>TGG ATT GTC GAC GGT ATG AA  | RD7-flank.F<br>GGT AAT CGT GGC CGA CAA G                  |
|                        |                     |         | RD7-Rv1976.int.R<br>GGT CGA TAA GGT CAC GGA AC  | RD7-flank.R<br>CAG CTC TTC CCC TCT CGA C                  |
| RD8                    | <i>ephA-lpqG</i>    | 5.9     | RD8-ephA.F<br>GGT GTG ATT TGG TGA GAC GAT G     | RD8-flank.F<br>CAA TCA GGG CTG TGC TAA CC                 |
|                        |                     |         | RD8-ephA.R<br>AGT TCC TCC TGA CTA ATC CAG GC    | RD8-flank.R<br>CGA CAG TTG TGC GTA CTG GT                 |
| RD9                    | <i>cobL</i> -Rv2075 | 2.0     | RD9-intF<br>CGA TGG TCA ACA CCA CTA CG          | RD9-flankF<br>GTG TAG GTC AGC CCC ATC C                   |
|                        |                     |         | RD9-intR<br>CTG GAC CTC GAT GAC CAC TC          | RD9-flankR<br>GCC CAA CAG CTC GAC ATC                     |
| RD10                   | Rv0221-Rv0223       | 1.9     | RD10-intF<br>GTA ACC GCT TCA CCG GAA T          | RD10-flankF<br>CTG CAA CCA TCC GGT ACA C                  |
|                        |                     |         | RD10-intR<br>GTC AAC TCC ACG GAA AGA CC         | RD10-flankR<br>GTC ATG AAC GCC GGA CAG                    |

|                                                         |                              |      |                                                 |                                            |
|---------------------------------------------------------|------------------------------|------|-------------------------------------------------|--------------------------------------------|
| RD11                                                    | Rv2645-Rv2659c               | 11.0 | RD11-Rv2646F<br>CGG CAG CTA GAC GAC CTC         | RD11-fla-F<br>TCA CAT AGG GGC TGC GAT AG   |
|                                                         |                              |      | RD11-Rv2646R<br>AAC GTG CTG CGA TAG GTT TT      | RD11-fla-R<br>AGA GGA ACC TTT CGG TGG TT   |
| RD12                                                    | <i>sseC</i> -Rv3121          | 2.8  | RD12-Rv3120.int.F<br>GAA ATA CGA GTG CGC TGA CC | RD12-flank.F<br>GCC ATC AAC GTC AAG AAC CT |
|                                                         |                              |      | RD12-Rv3120.int.R<br>CTC TGA ACC ATC GGT GTC G  | RD12-flank.R<br>CGG CCA GGT AAC AAG GAG T  |
| RD13                                                    | Rv1255c-Rv1257c              | 3.0  | RD13intF<br>GGA TGT CAC TCG GAA CGG CA          | RD13-flank.F<br>CGA TGG TGT TTC TTG GTG AG |
|                                                         |                              |      | RD13intR<br>CAC CGG GCT GAT CGA GCG A           | RD13-flank.R<br>GGA TCG GCT CAG TGA ATA CC |
| RD14                                                    | Rv1765c-Rv1773c              | 9.0  | RD14-Rv1769.int.F<br>GTG GAG CAC CTT GAC CTG AT | RD14-flankF<br>TTG ATT CGC CAA CAA CTG AA  |
|                                                         |                              |      | RD14-Rv1769.int.R<br>CGT CGA ATA CGA GTC GAA CA | RD14-flankR<br>GGG CTG GTT AGT GTC GAT TC  |
| <b>Region missing from <i>M. tuberculosis</i> H37Rv</b> |                              |      |                                                 |                                            |
| RvD1*                                                   |                              | 5.0  | RvD1-int1F<br>AGC GCG TCG AAC ACC GGC           | RvD1-int2.F<br>GAG CCA CTC CGA TGT TGA CT  |
|                                                         |                              |      | RvD1-int1R<br>CCT GAA TCC GCG CAA TTC CAT       | RvD1-int2.R<br>CAC GCG AAC CCT ACC TAC AT  |
| RvD2*                                                   | <i>plcD</i>                  | 5.1  | RvD2-int1F<br>GTT CTC CTG TCG AAC CTC CA        | RvD2-int2F<br>GGA CGG TGA CGG TAT TTG TC   |
|                                                         |                              |      | RvD2-int1R<br>ACT TCA CCG GTT TCA TCT CG        | RvD2-int2R<br>TCG CCA ACT TCT ATG GAC CT   |
| RvD3                                                    |                              | 1.0  | RvD3-intF<br>ATC GAT CAG GTC GTC AAT GC         | RvD3-flank.F<br>AAA CCA TGC AGC GTC TGC CA |
|                                                         |                              |      | RvD3-intR<br>ACG CCA CCA TCA AGA TCC            | RvD3-flankR<br>GCG TTT CTG CGT CTG GTT GA  |
| RvD4*                                                   | PPE gene                     | 0.8  | RvD4-intF-PPE<br>GGT TGC CAA CGT TAC CGA TGC    | ND                                         |
|                                                         |                              |      | RvD4-intR-PPE<br>CCG GTG GTG GTG GCG GCT        | ND                                         |
| RvD5                                                    | <i>moa</i>                   | 4.0  | RvD5intF<br>GGG TTC ACG TTC ATT ACT GTT C       | RvD5-flankF<br>CCC ATC GTG GTC GTT CAC C   |
|                                                         |                              |      | RvD5intR<br>CCT GCG CTT ATC TCT AGC GG          | RvD5-flankR<br>GTA CCC GCA CCA CCT GCT G   |
| TbD1                                                    | <i>mmpS6</i><br><i>mmpL6</i> | 2.1  | TbD1intS.F<br>CGT TCA ACC CCA AAC AGG TA        | TbD1fla1-F<br>CTA CCT CAT CTT CCG GTC CA   |
|                                                         |                              |      | TbD1intS.R<br>AAT CGA ACT CGT GGA ACA CC        | TbD1fla1-R<br>CAT AGA TCC CGG ACA TGG TG   |

| <i>katG</i> , <i>gyrA</i> , <i>oxyR</i> <sup>'</sup> , <i>pncA</i> and <i>mmpL6</i> | PCR                                                        | and sequencing primers                                     |
|-------------------------------------------------------------------------------------|------------------------------------------------------------|------------------------------------------------------------|
| <i>katG</i> <sup>463</sup>                                                          | <i>katG</i> -2154,225-PCR-F<br>CTA CCA GCA CCG TCA TCT CA  | <i>katG</i> -2154,872-SEQ-R<br>ACA AGC TGA TCC ACC GAG AC  |
|                                                                                     | <i>katG</i> -2155,157-PCR-R<br>AGG TCG TAT GGA CGAACA CC   |                                                            |
| <i>gyrA</i> <sup>95</sup>                                                           | <i>gyrA</i> -7,127-PCR-F<br>GTT CGT GTG TTG CGT CAA GT     | <i>gyrA</i> -7,461F<br>CGG GTG CTC TAT GCA ATG TT          |
|                                                                                     | <i>gyrA</i> - 8,312-PCR-R<br>CAG CTG GGT GTG CTT GTA AA    |                                                            |
| <i>oxyR</i> <sup>285</sup>                                                          | <i>oxyR</i> 2725,559F<br>TAT GCG ATC AGG CGT ACT TG        | <i>oxyR</i> -2726,024-SEQ-R<br>CAA AGC AGT GGT TCA GCA GT  |
|                                                                                     | <i>oxyR</i> -2726,024-PCR-R<br>CAA AGC AGT GGT TCA GCA GT  |                                                            |
| <i>pncA</i> <sup>57</sup>                                                           | <i>pncA</i> -2288,678-PCR-F<br>ATC AGG AGC TGC AAA CCA AC  | <i>pncA</i> - 2289,319-SEQ-R<br>GGC GTC ATG GAC CCT ATA TC |
|                                                                                     | <i>pncA</i> - 2289,319-PCR-R<br>GGC GTC ATG GAC CCT ATA TC |                                                            |
| <i>mmpL6</i> <sup>551</sup>                                                         | <i>mmpL</i> -seq5F<br>GTA TCA GAG GGA CCG AGC AG           | <i>mmpL</i> -seq5F<br>GTA TCA GAG GGA CCG AGC AG           |
|                                                                                     | TBD1fla1-R<br>CAT AGA TCC CGG ACA TGG TG                   |                                                            |

The RD nomenclature used in this table is based on that used by Brosch *et al.* (1) and differs from that proposed by Behr *et al.* (2). Primer sequences are shown in 5' →3' direction.

More information on the exact location of the deleted regions can be obtained at <http://www.pasteur.fr/recherche/unites/Lgmb/>.

\* Regions where a second pair of internal primers was used rather than flanking primers, because of flanking repetitive regions, and/or mobile genetic elements.

## References

1. Brosch, R., Gordon, S. V., Eiglmeier, K., Garnier, T., Tekaia, F, Yeramian, E., & Cole, S.T. (2000) in *Molecular Genetics of Mycobacteria*, eds. Hatfull, G. F., & Jacobs, W. R., Jr. (Am. Soc. Microbiol., Washington, DC), pp. 19-36.
2. Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S. & Small, P. M. (1999) *Science* **284**, 1520-1523.