

Figure S1. The diagram of genetic variation of PPE39 among *M. tb* Beijing/K, *M. tb* H37Rv and *M. bovis* BCG Pasteur 1173. The sequence alignment of PPE39 among the *M. tb* Beijing/K (A), *M. tb* H37Rv (B) and *M. bovis* BCG (C) was represented. The dotted square indicates the N-terminus region of PPE39 in the Beijing/K strain which is disrupted by IS6110 insertion in H37Rv. PPE39 (MTBK_24820) identified in the *M. tb* Beijing/K strain includes additional 259 amino acids in N-terminus compared with the PPE39 in *M. tb* H37Rv. BCG pasteur 1173P2 strain has a PPE39 homologue but about 150 amino acids at C-terminus are fused with a part of PPE40. The black boxes within the *M. tb* H37Rv and *M. bovis* BCG genomes represent variant regions of amino acids compared with Beijing/K strain.

