



Figure S2. Elevation in LD values in Zambian and Raleigh data after down sampling to 50 strains. Zambian and Raleigh strains were down-sampled to 50 strains 10 times and the resulting LD distributions were plotted (grey). In contrast to expectations under any neutral demographic model tested with a sample size of 50, all samples of 50 strains have elevated long range LD. This indicates that the elevation in LD is not driven by any sub-population. The LD calculations were conditioned on minor allele frequency (MAF) classes as follows: (A) Zambian (ZI) data and demographic models with MAFs between 0.3 and 0.5. (B) Zambian data and demographic models with MAFs between 0.05 and 0.5. (C) Zambian data and demographic models with MAFs between 0.0 and 0.05. (D) Raleigh (RA) data and demographic models with MAFs between 0.3 and 0.5. (E) Raleigh data and demographic models with MAFs between 0.05 and 0.5. (F) Raleigh data and demographic models with MAFs between 0.0 and 0.05.