

Table S2. Proportion of true transmission routes identified by both maximum likelihood (ML) and genetic similarity. SIR outbreaks with 30 initial susceptibles were simulated and a single genome sample was generated for each infective. For scenarios with bottleneck size >1, it was assumed that transmission events prior to the infection of the source occurred at intervals equal to the mean generation interval. Simulations with a final size <20 were discarded. For each infective, the maximum likelihood source was calculated under the geometric-Poisson approximation, and the genetically closest hosts selected. Simulations for each scenario were repeated 100 times. Baseline parameters: infection rate 0.002, removal rate 0.001, effective population size 5000.

Mutation rate ($\times 10^{-4}$)	1			3			5		
	Bottleneck size	1	5	25	1	5	25	1	5
Prop. routes identified by ML	0.27	0.21	0.21	0.32	0.23	0.22	0.33	0.24	0.21
Prop. routes identified by genetic similarity	0.19	0.17	0.15	0.27	0.20	0.18	0.29	0.22	0.19