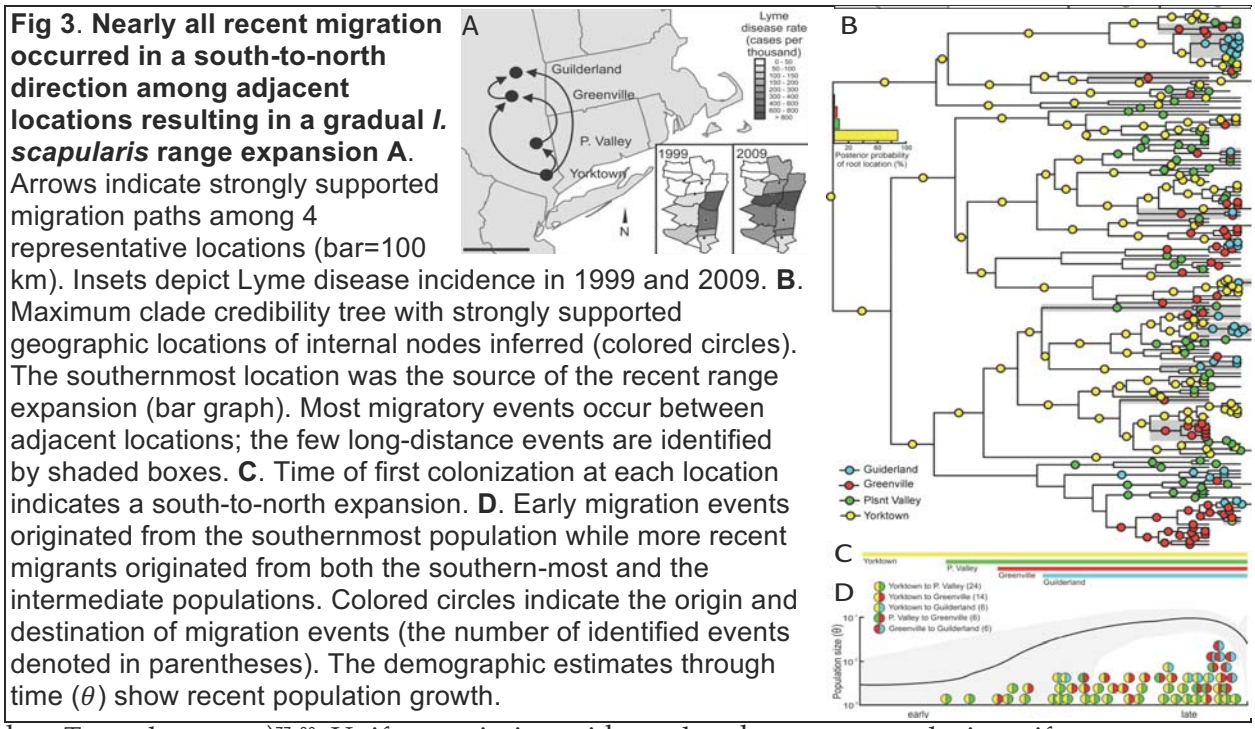

Fig 3. **Nearly all recent migration occurred in a south-to-north direction among adjacent locations resulting in a gradual *I. scapularis* range expansion A**. Arrows indicate strongly supported migration paths among 4 representative locations (bar=100 km). Insets depict Lyme disease incidence in 1999 and 2009. **B**. Maximum clade credibility tree with strongly supported geographic locations of internal nodes inferred (colored circles). The southernmost location was the source of the recent range expansion (bar graph). Most migratory events occur between adjacent locations; the few long-distance events are identified by shaded boxes. **C**. Time of first colonization at each location indicates a south-to-north expansion. **D**. Early migration events originated from the southernmost population while more recent migrants originated from both the southern-most and the intermediate populations. Colored circles indicate the origin and destination of migration events (the number of identified events denoted in parentheses). The demographic estimates through time (θ) show recent population growth.

than *Taq* polymerase)[77–80]. Uniform priming with random hexamers results in uniform amplification of all DNA within a sample. In contrast, SWGA takes advantage of the fact that every species has different mutational biases that create over- and under-represented sequence motifs within their genomes[49]. Motifs that are over-represented in the target microbial genome, but under-represented in the tick genome, are used to create primer sets (2-20 primers each 6-12bp in length) for ϕ29 reactions that amplify only the target microbial genome[50]. We have used these technologies to develop successful SWGA protocols to sequence a variety of microbial species without prior culturing including: *Wolbachia pipientis* (fruit flies)[49], multiple *Plasmodium* species (primate blood)[51], *Trypanosoma cruzi* (triatomine bug), *Mycobacterium tuberculosis* (human) and *B. burgdorferi* (ticks). Here we describe three representative cases. *W. pipientis* - a maternally-inherited bacterium in fruit flies - accounts for ~2% of MiSeq reads from genomic extracts of infected fruit flies[49]. Thus, 98% of reads are wasted on non-target DNA if *W. pipientis* is sequenced directly from fly genomic extracts. However, sequencing *W. pipientis* from the same fly extract after SWG amplification results in ~70% of sequencing reads mapping to *W.*

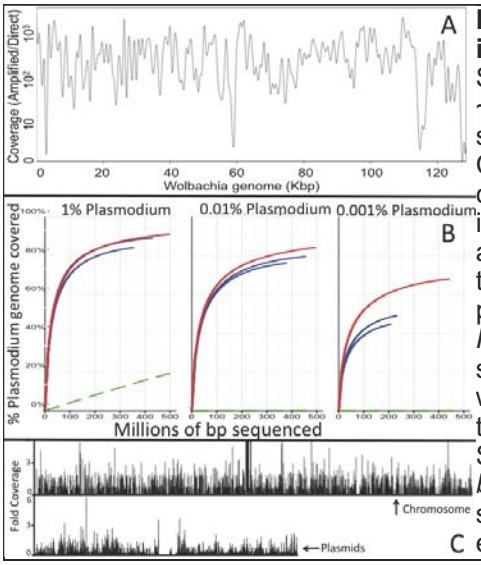

Fig 4. Selective whole genome amplification dramatically improves sequencing coverage of the target genome. A. Sequencing coverage (normalized by sequencing effort) was ~100X greater in **SWG amplified** samples than directly sequenced samples across the *Wolbachia* genome. Coverage of amplified samples at each bp was divided by coverage from directly sequenced samples. Only 4 locations in the *Wolbachia* genome had equivalent coverage in the amplified and non-amplified samples (no reads mapped to those areas for either sample). **B**. SWGA performed on primate genomic DNA containing known quantities of *Plasmodium* DNA (1% to 0.001%). The sequencing effort is shown in relation to the percent of the *Plasmodium* genome with sequencing coverage. Red lines show the combination of two independent amplifications (blue lines). Coverage without SWGA is shown in green. **C**. Although no reads map to *B. burgdorferi* from tick extracts prior to SWGA, post-SWGA samples with equivalent sequencing effort have relatively even coverage.