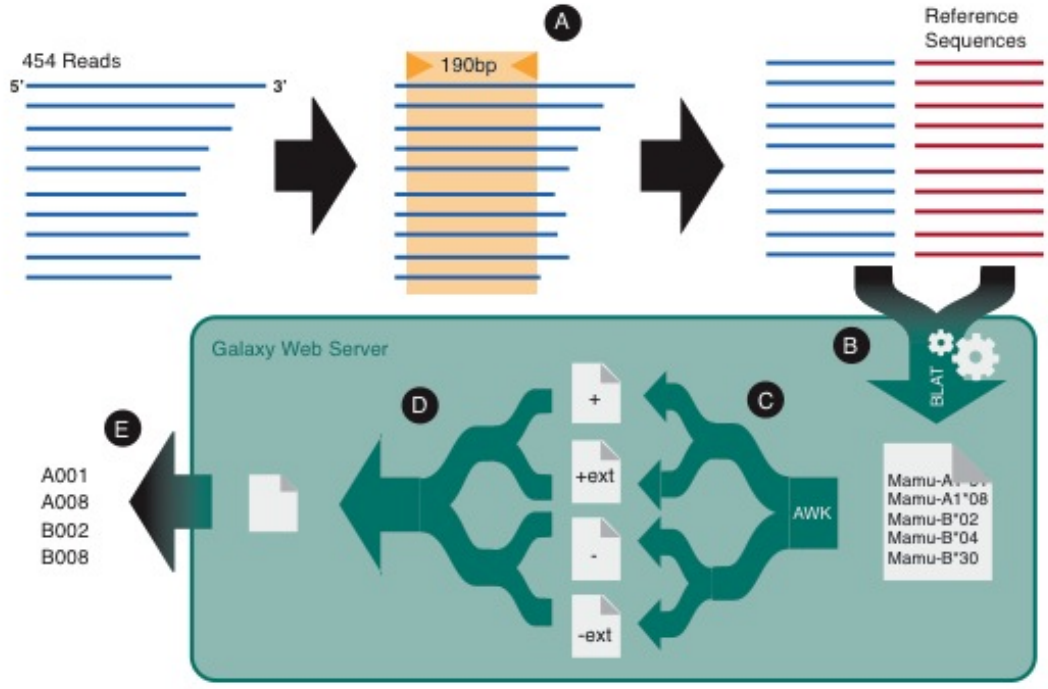

Figure S1 Galaxy-based data analysis workflow. (A) Roche/454 reads were uniformly trimmed down to 190 bp, removing the 5' multiplex identifier (MID) sequence and trimming the remaining length from the 3' end. (B) Reads and reference sequences were uploaded into our Galaxy web server and processed through the BLAST-like analysis tool (BLAT). (C) High-quality, full-length matches were extracted from the BLAT output and binned by direction. (D) Matches were grouped by sample, counted and merged into one line per allele. (E) The report file was then processed manually to generate the haplotype calls.