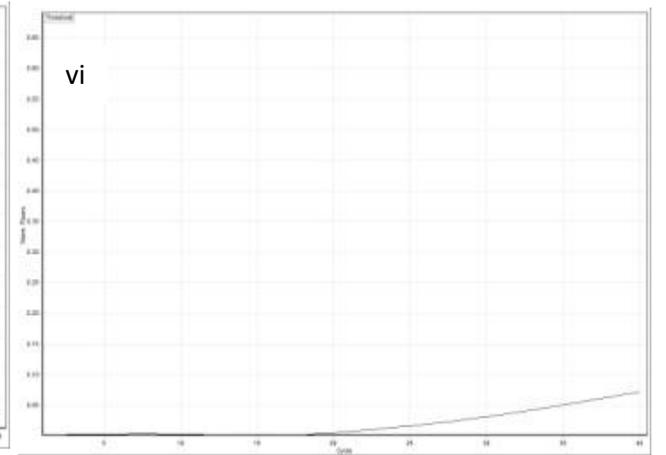
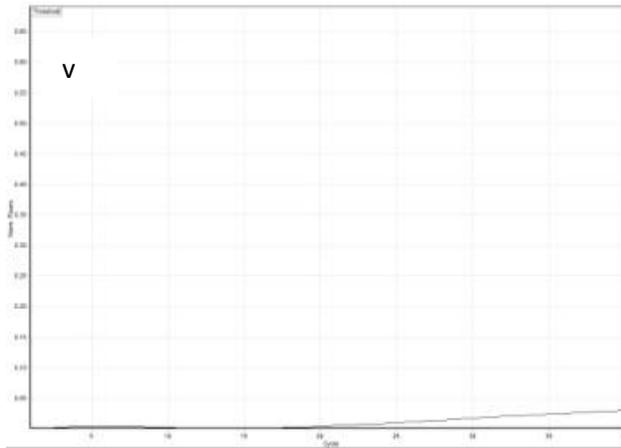
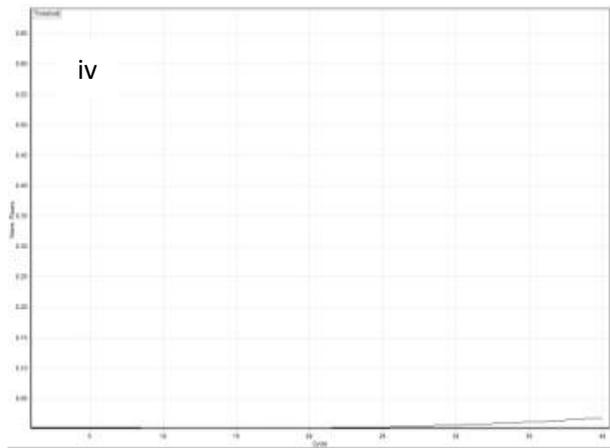
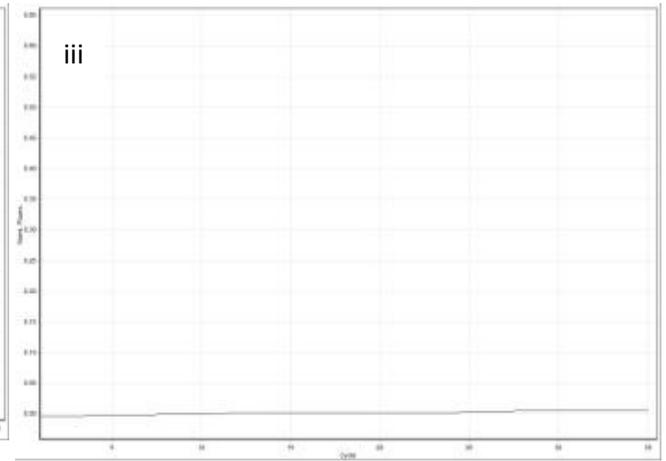
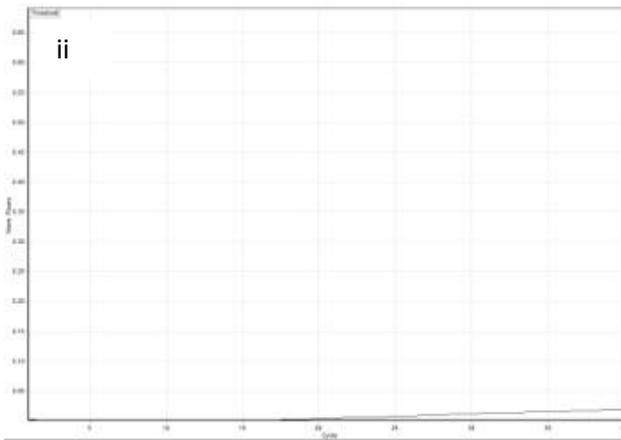
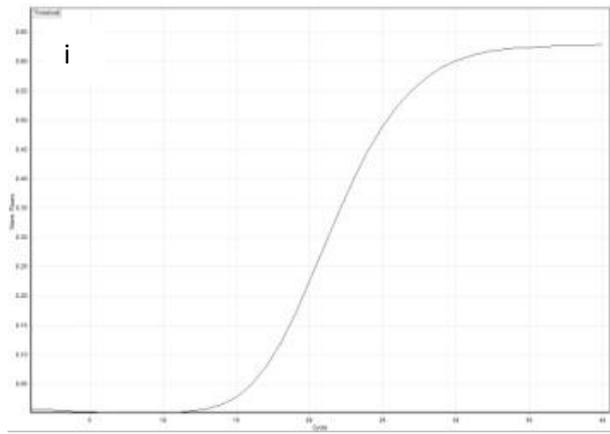
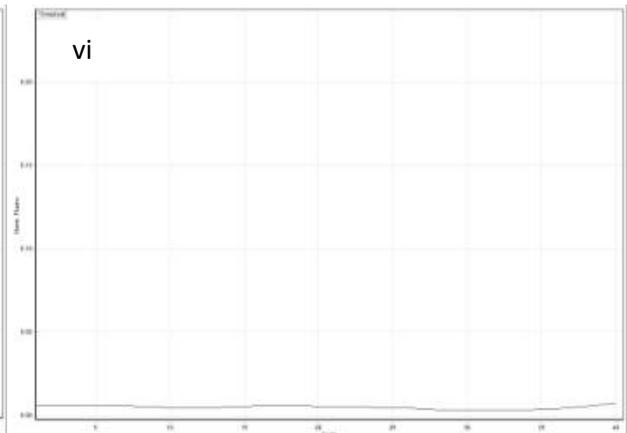
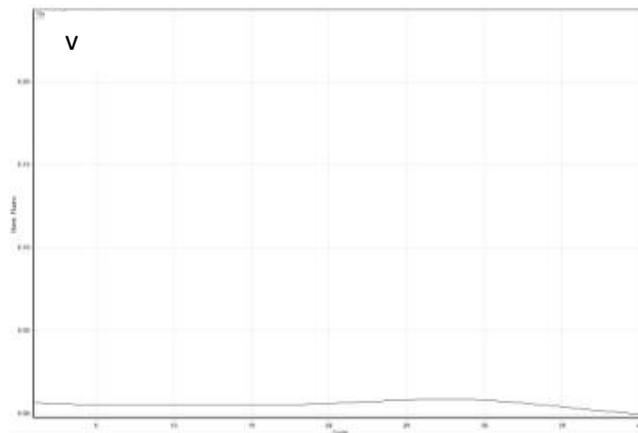
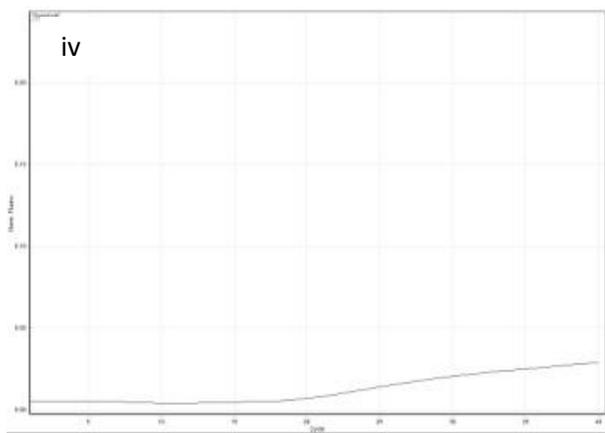
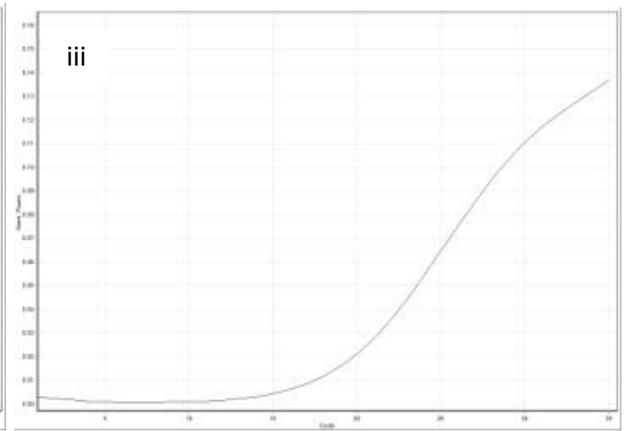
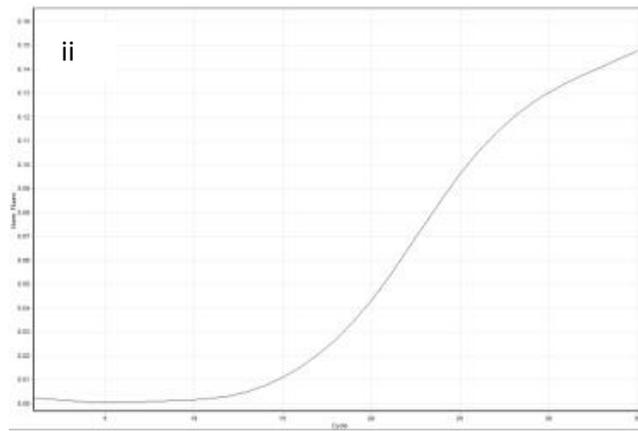
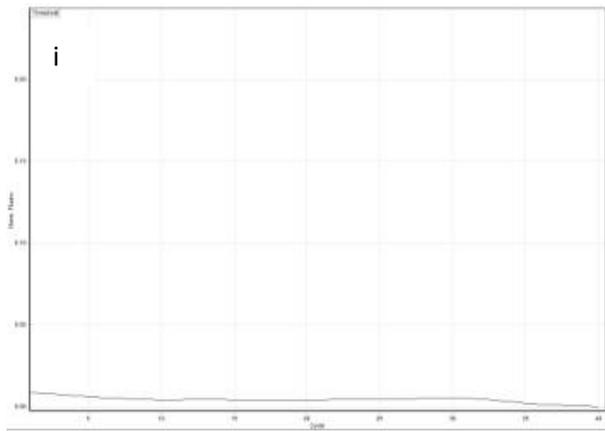


**Supplement Figure 1. The traces from two assumed Mendelian markers with a single BLAST hit against the *Emiliana huxleyi* CCMP 1516 genome sequence (Read et al. 2013) and one non-Mendelian marker. Each arrow indicates a putative microsatellite allele. a) Locus P02F11 did not exhibit stutter peaks occasionally associated with microsatellite trace profiles, but produced reliable single peaks. This trace demonstrates a homozygous strain at P02F11 (i.e., allele 1). b) Locus P01E05 exhibited stutter bands. However, it was possible to score the largest and furthest right peak as the locus profile was reproducible within and among strains (i.e., the same number of stutter peaks with the largest peak). This trace demonstrates a heterozygous strain at P01E05 (i.e., allele 1 and allele 2). c) Locus P02E10 produced multiple peaks for the tested strains which were not reproducible between different PCRs of the same strain. This locus was found five times in the CCMP 1516 genome. Moreover, it was impossible to score the alleles as exhibited in this trace where there were four possible alleles. The fourth allele (indicated with an asterix) does not follow the typical stutter motif pattern as the largest peak is not the furthest right. This problem was also found in different PCRs where the peak heights within the stutter motif were altered in which the largest peak could be the first, second or third. The remaining loci, EHMS15 and P02E11, exhibited similar types of patterns as demonstrated at P02E10.**

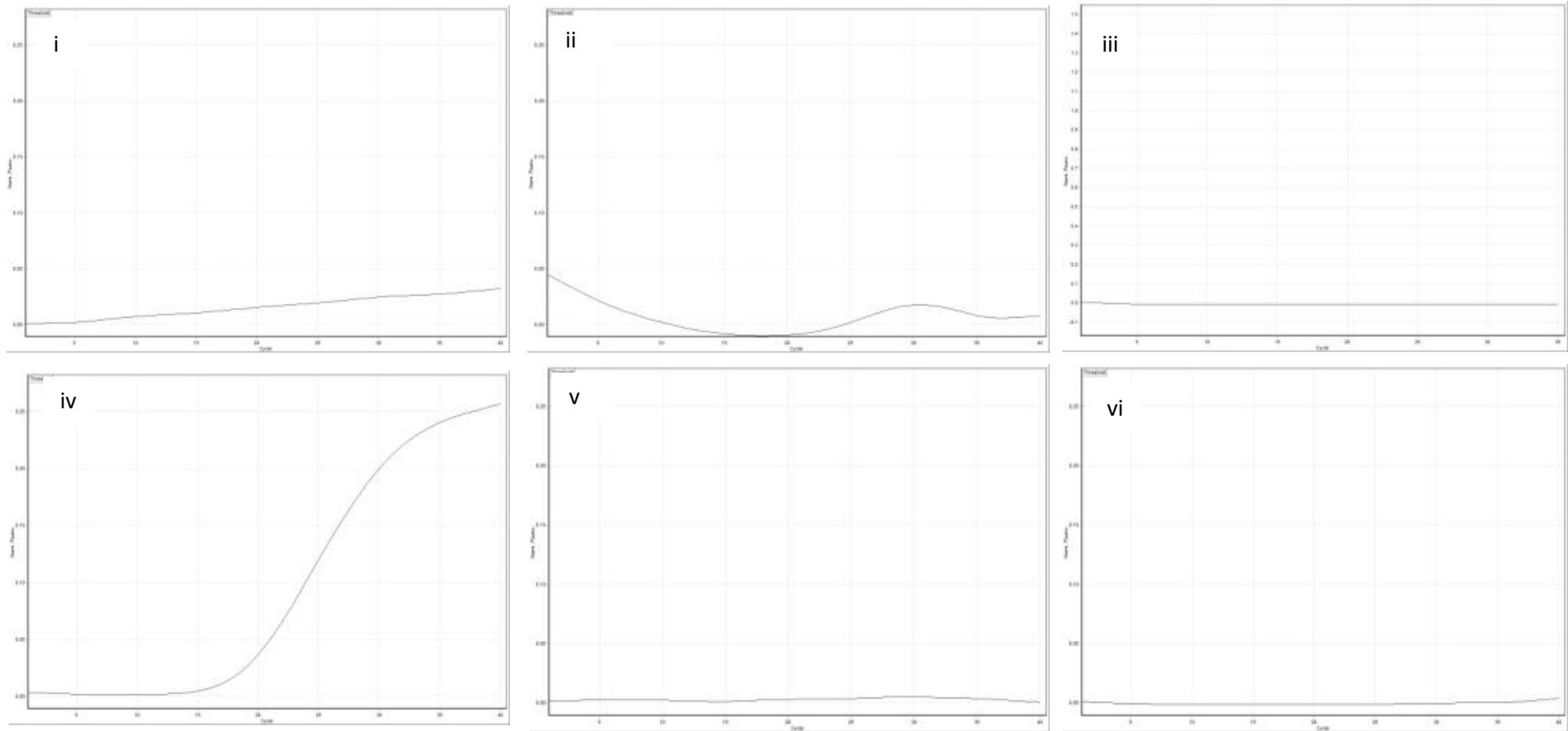
a)



b)

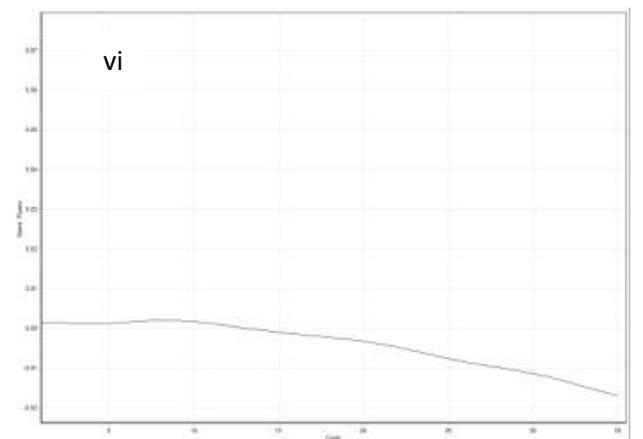
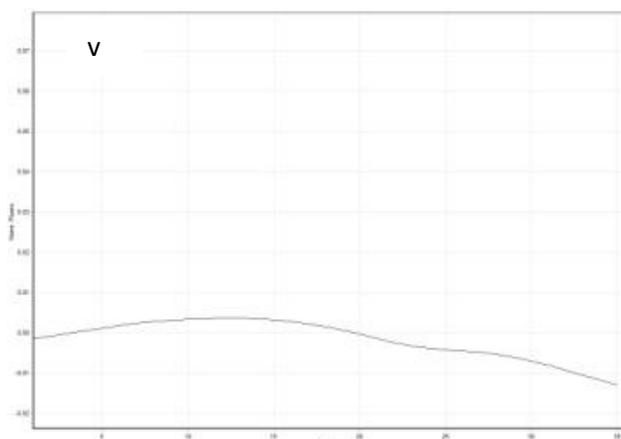
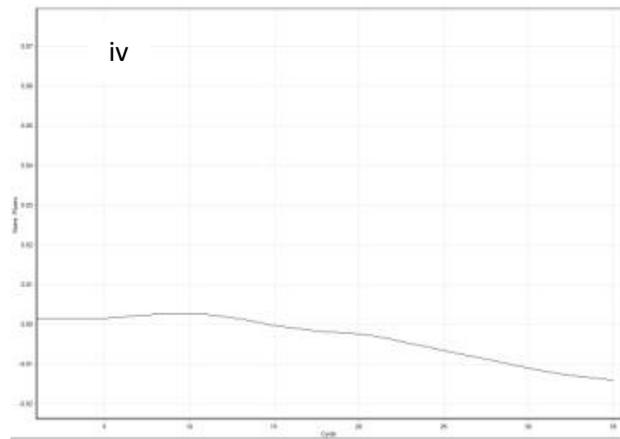
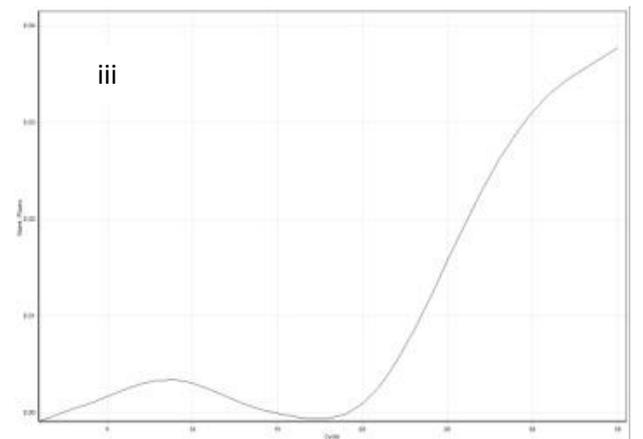
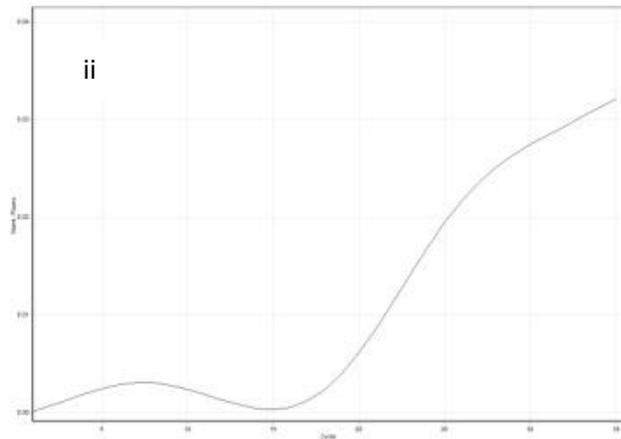
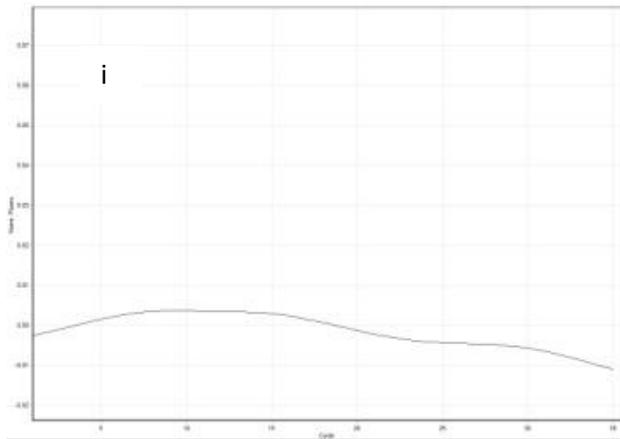


c)

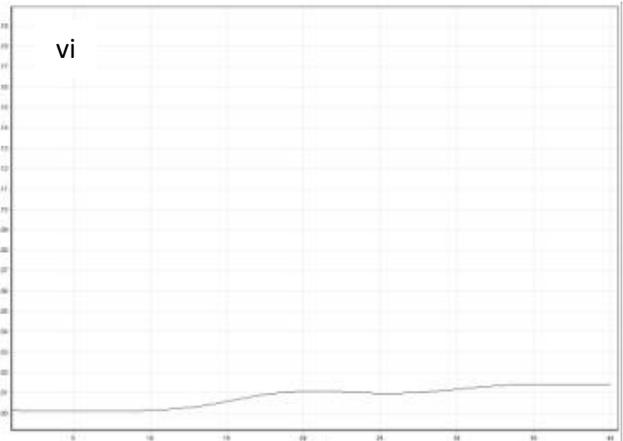
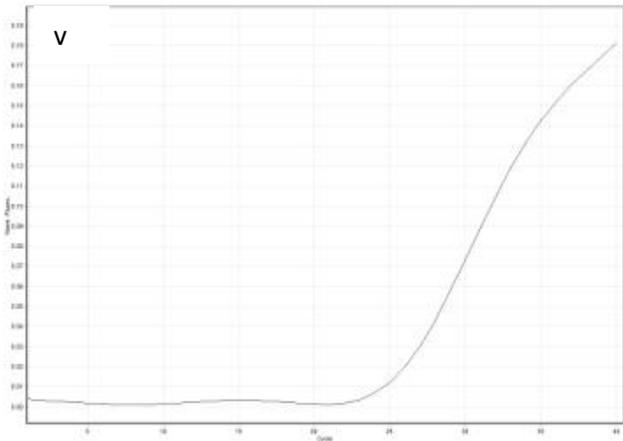
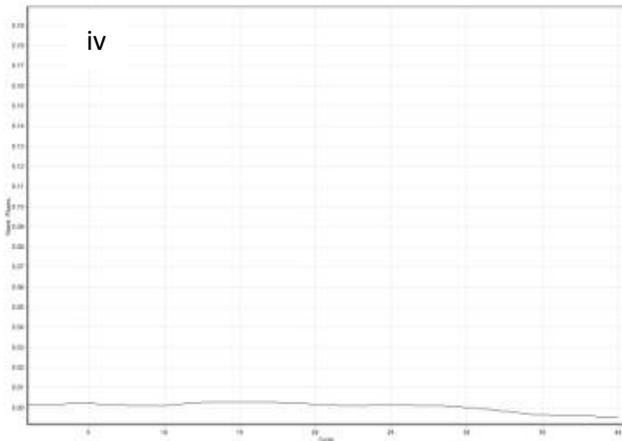
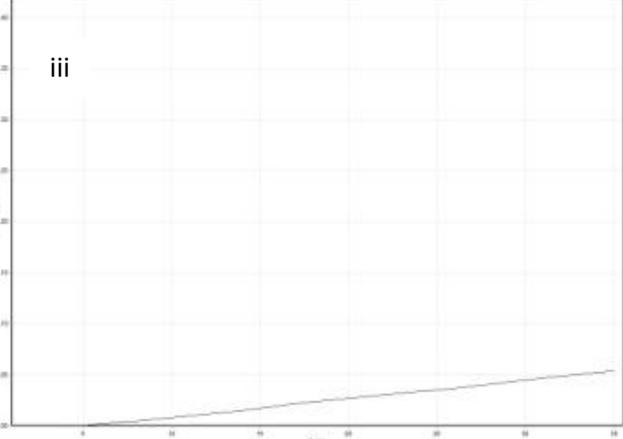
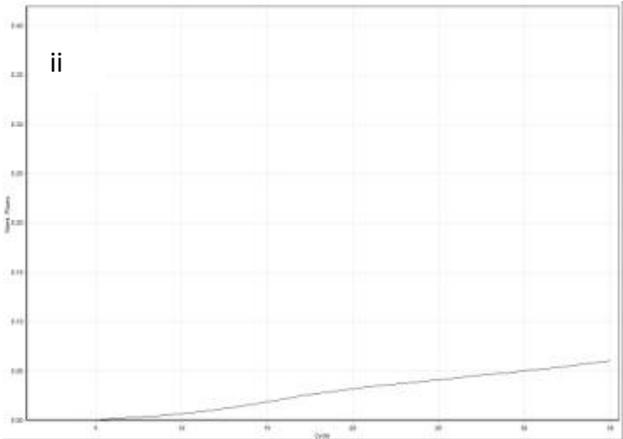
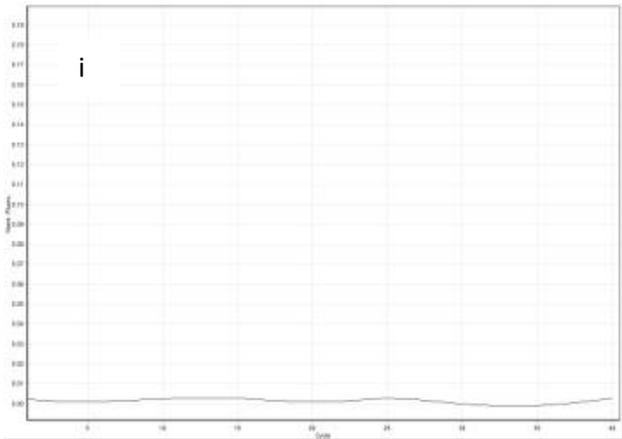


**Supplement Figure 2.** The following panels show the probes-specificity for their respective sequences for Multiplex 1: a) probe I (6-FAM –green channel), b) probe II (ATTO680 – chromson channel) and c) probe III (CY5- red channel) (Table 3). i -vi) CMM I, II, IIb, III, IV and IVb template DNA, respectively.

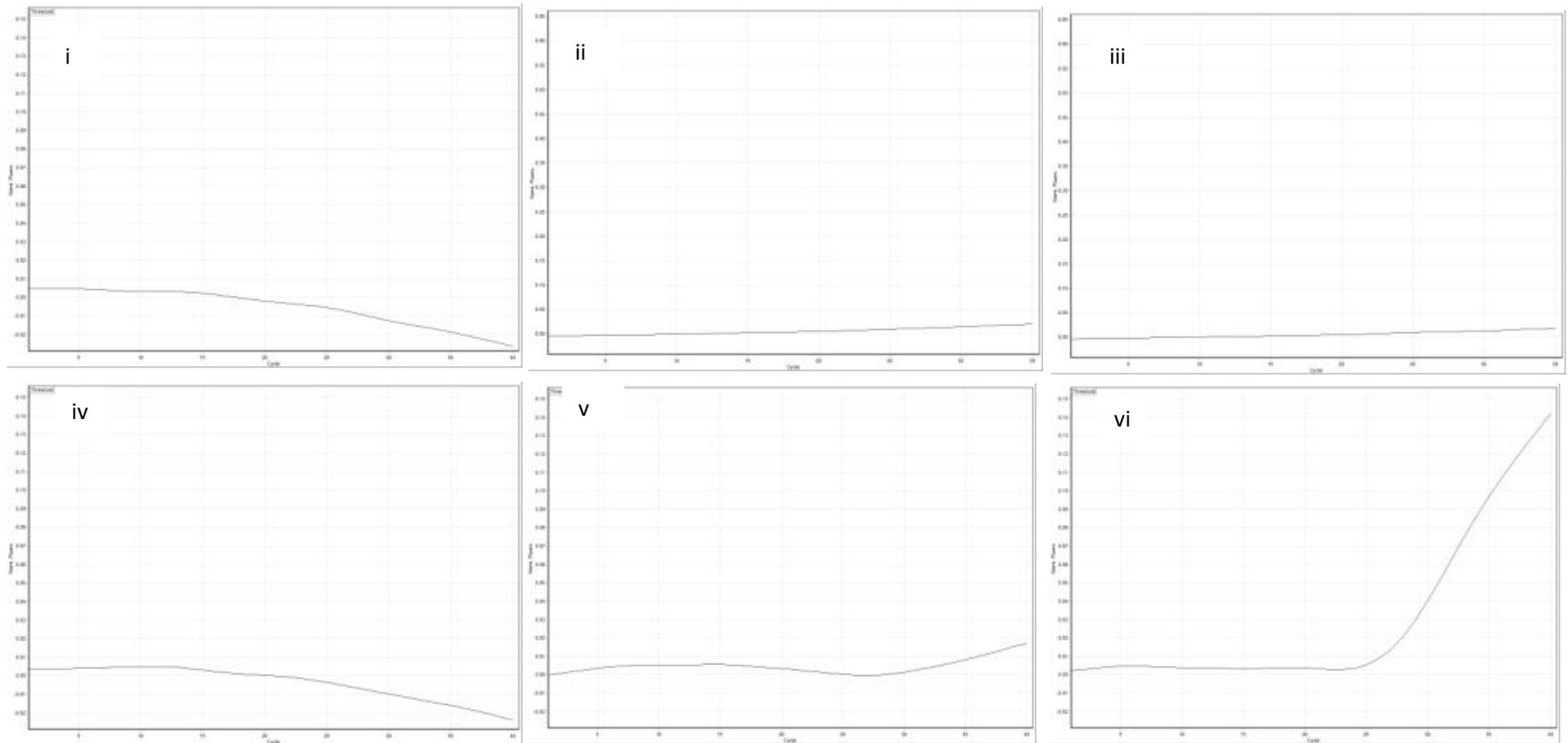
a)



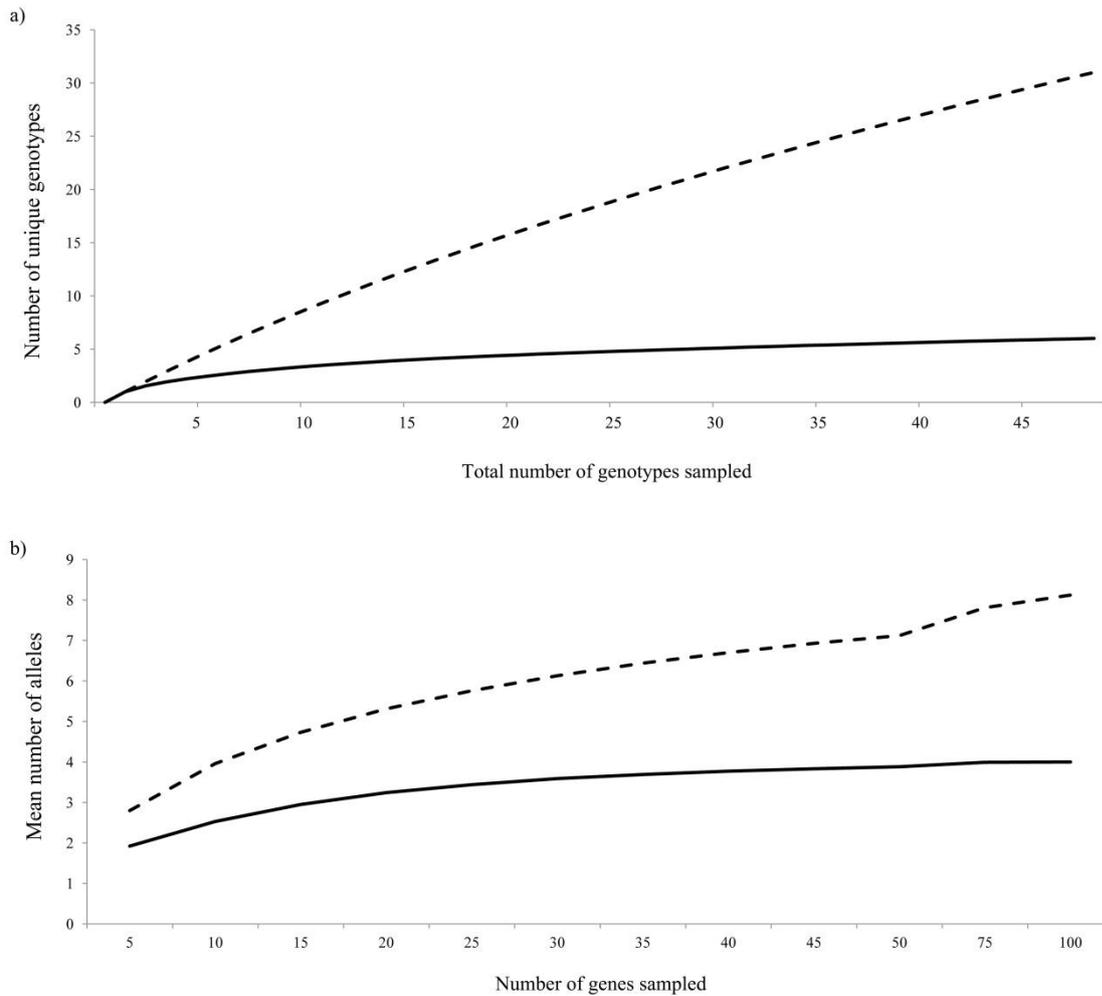
b)



c)



**Supplement Figure 3.** The following panels show the probes-specificity for their respective sequences for Multiplex 2: a) probe IIb (HEX –yellow channel), b) probe IV (ATTO680 – chromson channel) and c) probe IVb (6-FAM –green channel) (Table 3). i -vi) CMM I, II, IIb, III, IV and IVb template DNA, respectively.



**Supplement Figure 4.** a) Genotype rarefaction curve for CMM genotype (solid line) and microsatellites (dashed line) generated using FASTGROUPII (Yu et al., 2006). The curve reaches a plateau for CMM genotypes whereas, there appeared to be microsatellite MLGs which were not sampled. b) The variation in the mean number of alleles (averaged over 5 microsatellite loci) observed with different numbers of genes sampled using HP-RARE (Kalinowski 2005) indicated that after sampling at least 75 genes (or at least 38 diploid isolates), the chances of encountering a rare allele decreased.