

Population genomics of *Macrophomina* spp. reveals cryptic host specialization and evidence for meiotic recombination

K.K. Pennerman<sup>1</sup>, P. Goldman<sup>1</sup>, C.J. Dilla-Ermita<sup>1,2</sup>, G. Ramos<sup>1</sup>, J.H. Jaime<sup>1</sup>, J. Lopez-Hernandez<sup>1</sup>, J. Ramos<sup>1</sup>, M. Aviles<sup>3</sup>, C. Borrero<sup>3</sup>, A.O. Gomez<sup>4</sup>, J.M. Neal<sup>4</sup>, M. Chilvers<sup>5</sup>, V. Ortiz<sup>5</sup>, E.H. Stukenbrock<sup>6,7</sup>, G.H. Goldman<sup>8,9</sup>, A. Mengitsu<sup>10</sup>, H.D. Lopez-Nicora<sup>11</sup>, G.O. Sacher<sup>12</sup>, N. Vaghefi<sup>13,14</sup>, L. Kiss<sup>14</sup>, J.P. Benz<sup>9,15</sup>, A.R. Machado<sup>16</sup>, T.E. Seijo<sup>17</sup>, N.A. Peres<sup>17</sup>, F.N. Martin<sup>1</sup>, J.C. Broome<sup>18</sup>, K. Ivors<sup>19</sup>, G.S. Cole<sup>2</sup>, S.J. Knapp<sup>2</sup>, D.J. McFarlane<sup>20,21</sup>, S.W. Mattner<sup>20,21</sup>, M. Gambardella<sup>22</sup>, E. Gluck-Thaler<sup>23,24</sup>, P.M. Henry<sup>1,\*</sup>

<sup>1</sup> United States Department of Agriculture, Agricultural Research Service, Salinas, CA, USA

<sup>2</sup> University of California Davis, Department of Plant Sciences, Davis, CA, USA

<sup>3</sup> Dept. Agronomía, ETSIA, Universidad de Sevilla, Sevilla, Spain

<sup>4</sup> Queensland Department of Agriculture and Fisheries, Nambour, Queensland, Australia

<sup>5</sup> Michigan State University, Department of Plant, Soil and Microbial Sciences, East Lansing, MI, USA

<sup>6</sup> Christian-Albrechts University of Kiel, Kiel, Germany

<sup>7</sup> Max Planck Institute for Evolutionary Biology, Plön, Germany

<sup>8</sup> University of São Paulo, São Paulo, Brazil

<sup>9</sup> Technical University of Munich, Institute for Advanced Study, Garching, Germany

<sup>10</sup> United States Department of Agriculture, Agricultural Research Service, Jackson, TN, USA

<sup>11</sup> The Ohio State University, Department of Plant Pathology, OH, USA

<sup>12</sup> Pennsylvania State University, Department of Plant Pathology and Environmental Microbiology, PA, USA

<sup>13</sup> University of Melbourne, School of Agriculture, Food, and Ecosystem Sciences, Melbourne, Australia

<sup>14</sup> Centre for Crop Health, University of Southern Queensland, Toowoomba, Australia

<sup>15</sup> Technical University of Munich, TUM School of Life Sciences, Freising, Germany

<sup>16</sup> Universidade Federal de Pernambuco, Recife, Brazil

<sup>17</sup> University of Florida, IFAS, Gulf Coast Research and Extension Center, Wimauma, FL, USA

<sup>18</sup> University of California Santa Cruz, Department of Environmental Sciences, CA, USA

<sup>19</sup> Driscoll's, Inc., Watsonville, CA, USA

<sup>20</sup> VSICA Research, Toolangi, Australia

<sup>21</sup> La Trobe University, School of Agriculture, Biomedicine and Environment, Melbourne, Australia

<sup>22</sup> Pontificia Universidad Católica de Chile, Santiago, Chile

<sup>23</sup> University of Wisconsin-Madison, Department of Plant Pathology, Madison, WI, USA

<sup>24</sup> Wisconsin Institute for Discovery, Madison, WI, USA

**Corresponding Author:** Peter Henry  
Research Plant Pathologist  
USDA Agricultural Research Service  
peter.henry@usda.gov

## Abstract

Knowledge of the factors structuring populations of pathogenic fungi is fundamental to disease management efforts and basic biology. However, this crucial information is missing for many important pathogens, including broad host range and drought-associated pathogens from the globally distributed *Macrophomina* genus. The objectives of this work were to evaluate the evidence for host specialization, geographic adaptation, and recombination using a global survey of *Macrophomina* isolates from diverse geographic, temporal, and host sources. We obtained high-quality short-read sequence data for 463 *Macrophomina* spp. isolates, representing four putative species, collected from 91 host plant species and soil in 23 countries. Analysis of bi-allelic, single nucleotide polymorphisms revealed high diversity, admixture, and equal mating type ratios suggesting on-going recombination. Although most tested isolates asymptotically colonized strawberry, only strawberry-derived isolates caused disease on this host. These isolates were all in a single lineage, suggesting the ability to cause disease on strawberry is not widespread among *M. phaseolina*. Significant associations were also found between isolation from soybean plants and specific population clusters, suggesting that specialization for virulence or reproduction has also occurred for soybean. Geography  $\times$  isolate genotype associations were weak, suggesting *Macrophomina* spp. were frequently trafficked between regions. Reference free whole genome comparisons support current boundaries among four *Macrophomina* species, and new molecular markers were designed to specifically identify each species. Contrary to expectations, *M. phaseolina* should be considered a single species with both specialist and generalist populations in which meiosis can maintain genetic diversity.

## Introduction

Effective management of plant disease requires knowledge of the factors structuring pathogen genomic diversity, such as host and geographic specialization, and the mechanisms driving the evolution of new pathogenic strains. For example, host specialization, which refers to a pathogen's adaptation to cause more severe disease on a specific host, can occur at the sub-specific level (Edel-Hermann and Lecompte 2019; Gladieux et al. 2018), and it must be understood to develop accurate diagnostic methods and evaluate host resistance. Associations between geographic regions and pathogen genotype can illuminate routes of dispersal that could be disrupted by phytosanitary regulations (Ali et al. 2014; Feurtey et al. 2023; Sotiropoulos et al. 2022). Additionally, the frequency of meiotic or parasexual recombination can indicate the likelihood for new, virulent strains to emerge (McDonald and Linde 2002). A durable understanding of these factors can only be achieved with globally representative sampling, as international trade continues to move pests and pathogens to new areas (Schertler et al. 2024; Sikes et al. 2018). These factors are only characterized for a few well-studied fungi, and expanding research to novel systems is necessary to broaden knowledge of fungal evolution.

The genus *Macrophomina* contains economically important plant pathogens, such as the cosmopolitan type species, *M. phaseolina*. This species was first identified in the early 1900s, has been reported worldwide (Pennerman et al. 2024), and can cause total crop losses in soybean and strawberry (Bradley et al. 2021; Steele et al. 2023). As a species, *M. phaseolina* has a broad host range of at least 97 plant species in addition to humans with compromised immune systems (Pennerman et al. 2024). It is widely speculated that abiotic stresses induced by climate change will lead to increased damage by *M. phaseolina* (Cohen et al. 2022), as this fungus causes greater disease severity in warmer and drier conditions (Baggio et al. 2021; Chilakala et al. 2022; de

Sousa Linhares et al. 2020; Diourte et al. 1995). Four other *Macrophomina* species were recently described: *M. pseudophaseolina* (Sarr et al. 2014), *M. euphorbiicola* (Machado et al. 2019), *M. tecta* (Poudel et al. 2021), and *M. vaccinii* (Zhao et al. 2019). These species have only subtle differences in morphology and were probably mis-identified as *M. phaseolina* before the widespread adoption of molecular phylogenetics (Poudel and Vaghefi 2023). Thus, their global distribution and host range are unknown and could be much wider than is reported.

Pathogens are typically classified along a continuum ranging from host specific, infecting a single host, to generalists that infect many. Between these extremes are host specialists that predominantly infect a single host but may retain virulence on others (Gladieux et al. 2018; Mercier et al. 2019; Resende et al. 1994). Efforts to describe host specialization in *M. phaseolina* have usually relied on comparing isolates from different hosts without knowledge of their genomic differences, and this approach has not yielded clear results (Gupta et al. 2012; Manici et al. 1995; Reznikov et al. 2019; Su et al. 2001). An exception to this trend is for strawberry-associated strains of *M. phaseolina*, which were found to cause severe strawberry disease in Spain, Argentina, and California, USA (Burkhardt et al. 2019a; Koike et al. 2016; Martin et al. 2015; Viejobueno et al. 2022). However, the broad applicability of these findings is unknown, as no association between host isolation source and virulence in strawberry was found in Israel (Zveibil et al. 2012), and the populations of *M. phaseolina* from strawberry in Chile and Spain appear to be genetically distinct (Sánchez et al. 2017). High resolution genotypic information is needed to resolve these discrepancies and illuminate the status of host specialization in *M. phaseolina*.

While isolate genotype  $\times$  geography associations have been reported for specific populations of *M. phaseolina* (Ortiz et al. 2023; Sánchez et al. 2017), it remains unknown if the

global population of *Macrophomina* spp. is structured by geography, or if specific genotypes are found in disparate locations. A geographic influence on population structure seems plausible, given the apparent absence of aerial dispersal and the fact that specific environmental conditions play a large role in disease development. However, there is also some evidence for the ability of *Macrophomina* spp. to asymptotically infect plants (Fuhlbom et al. 2012; González and Tello 2011; Jamal et al. 2022; Koike et al. 2016; Moslemi et al. 2023), and be seedborne (Dell'Olmo et al. 2023; Fuhlbom et al. 2013; Mengitsu et al. 2012; Reuveni et al. 1983; Singh and Singh 1982) so they could have a high potential for dissemination between geographic regions.

Continuous adaptation to host and environment could be accelerated by on-going meiotic recombination. Many fungi that were historically considered asexual (Crous et al. 2006; Sarr et al. 2014) are now known to undergo meiosis at least periodically (Taylor et al. 2016). Both mating types exist in *M. phaseolina* (Nagel et al. 2018), but their ratios and the extent of admixture are unknown in extant populations. Gene flow could also occur through horizontal transfer of giant *Starship* transposable elements, which carry accessory genes to recipients (Gluck-Thaler et al. 2022; Urquhart et al. 2024). Parasexual recombination can happen during anastomosis of vegetatively compatible hyphae (Pereira et al. 2018; Sybuia et al. 2022), but allo-recognition systems usually prevent anastomosis between genetically distinct strains (Gonçalves and Glass 2020; Pereira et al. 2018; Sybuia et al. 2022). Because of this, parasexual recombination would be less likely to produce novel genotypes than meiotic recombination, which could occur between diverse genotypes (Bennett and Turgeon 2016). Exploring the potential for meiosis in *Macrophomina* spp. is key to understanding the evolutionary potential of this genus.

To fill these knowledge gaps, the objectives of this work were to evaluate the evidence for host specialization, geographic adaptation, and recombination using a global survey of *Macrophomina* isolates from diverse geographic, temporal, and host sources. This investigation is critical to the global efforts to control diseases caused by *Macrophomina* spp., which are increasing along with abiotic stresses induced by climate change, and for gaining a broad understanding of the factors shaping fungal evolution at a global scale. To our knowledge, this is the first global genomic diversity study for any genus or species in the Botryosphaeriaceae.

## Results

### **A global genomic survey revealed fourteen genotypic clusters of *Macrophomina* spp.**

Whole genome shotgun sequences were obtained for 463 *Macrophomina* spp. isolates (435 newly sequenced and 28 retrieved from NCBI GenBank; Supplementary Table S1). These isolates were derived from 23 countries, 91 plant hosts and soil. Collection years ranged between 1927 and 2021, and therefore included many isolates described as *M. phaseolina* by morphology alone and/or before 2014, when additional species were described within the genus *Macrophomina*. The most common hosts in our study reflected the most frequently researched hosts of *Macrophomina* spp. identified by Pennerman et al. (2024) (Supplementary Figure S1).

High-quality, bi-allelic single nucleotide polymorphisms (SNPs) with presence in at least 90% of isolates were identified using reference genomes of strains: Mp11-12 (from strawberry (Burkhardt et al. 2019a); 578,310 SNPs; average of 8.07 SNPs per kb for scaffolds with SNP loci), AL-1 (from alfalfa (Burkhardt et al. 2019a); 588,334 SNPs; 9.39 SNPs per kb), and mp117 (from soybean field soil (Gluck-Thaler et al. 2022); 584,184 SNPs; 10.32 SNPs per kb). Using these SNPs, fourteen population clusters were identified by a combination of admixture models,

reticulate phylogenetic trees and principal component analyses (PCAs; Figure 1; Supplementary Figures S2 and S3). The cluster assignments were largely congruent among references. The only discrepancies were related to cluster assignment of isolates F285 (at least 10% membership by Q-value in clusters VIII, IX, XIII and XIV) and F385 (clusters II, III, IX and XI). Cluster assignments in figures throughout this work are shown using the Mp11-12 reference, unless otherwise noted. Two clusters contained isolates that were not described as *M. phaseolina*: cluster I contained *M. euphorbiicola* and *M. pseudophaseolina* (12 isolates total), and cluster II contained eight *M. tecta*, one putative *M. vaccinii* isolate, and two *M. phaseolina* isolates (identified by multi-gene phylogenies shown below; cluster II = 11 isolates total). The remaining twelve clusters contained isolates described as *M. phaseolina*.

### **Pangenomic gene content was structured by the 14 population clusters.**

A pangenomic analysis of gene content was conducted to evaluate the strength of the cluster designations and the genetic diversity within *Macrophomina* spp. A total of 113,290 homologs were identified across all genome assemblies. These homologs comprised 10,984 "super-core" homologs (present in all 463 *Macrophomina* spp. isolates), 18,866 "core" homologs (present in at least 95% of the isolates), and 45,233 "accessory" homologs (present in at least two isolates but fewer than 95% of all isolates; Figures 2A, 2B, 2C). Hierarchical clustering based on core and accessory homolog presence/absence variation yielded groups that were congruent with those identified by admixture and PCA analyses (Figure 2D). There were 38,207 singletons, and the pangenome appeared slightly open when these genes were included. When singletons were excluded, the pangenome structure was clearly closed (Figure 2E). The rarefaction observed for



highly sampled clusters (e.g. V, VII, X, XII, XIII and XIV), suggested that continued sampling would have corroborated our finding of a closed pangenome structure (Figure 2F).

**Strawberry-derived isolates from *M. phaseolina* cluster XIII were specialized for virulence in strawberry and associated with recent outbreaks of strawberry crown rot disease**

Significantly more isolates from strawberry (85 out of 106) were in cluster XIII than would be expected by random chance ( $\chi^2 P < 0.001$ ). We hypothesized that cluster XIII isolates were specialized for virulence in strawberry and tested this with virulence assays of 41 isolates representing eight *M. phaseolina* clusters. Five of the six tested isolates from the strawberry-associated cluster XIII consistently caused disease in the susceptible strawberry cultivar, 'Monterey', as indicated by area under the disease progress curve (AUDPC) and measured plant and fruit biomass (Figure 3). No isolate from another cluster caused disease on strawberry plants, but almost all isolates could be recovered from asymptomatic plants at the end of the seven-week growth period (Figure 3D). This indicates that the potential for asymptomatic, endophytic growth on strawberry is common in *M. phaseolina*. Isolates from clusters other than XIII were not derived from strawberry, and there may potential for some strawberry derived isolates in other clades to cause disease on this host. Nevertheless, these data indicate that not all strains of *M. phaseolina* are specialized for virulence on strawberry.

The cluster XIII host-specialized strains appeared to have emerged recently and were responsible for most new reports of *M. phaseolina* causing disease in strawberry. Consistent with recent emergence, isolates in cluster XIII had low genetic diversity, as evidenced by their high fraction of gene conservation, and low anchor distance divergence (Figures 2D and 2F, and Supplementary Figure S4). Our dataset included only three isolates of XIII before it was first

recovered from strawberry: chrysanthemum in 1934 from Missouri (F103), slash pine in 1983 from Paraguay (F287), and pepper in 2000 from Australia (F322). From 2001 onward, cluster XIII isolates were associated with the nearly simultaneous first reports of highly destructive disease in strawberry in Florida (Mertely et al. 2005), Spain (Avilés et al. 2008), California (Koike 2008), Mexico, and Australia (Golzar et al. 2007; Hutton et al. 2013) (Supplementary Figure S5).

#### **Clusters IV, V, VI, VII, and VIII were associated with the soybean host**

Most isolates from soybean or soil from soybean fields were assigned to five clusters: IV (7 soybean and soybean field isolates/9 total), V (34 isolates/41 total), VI (13 isolates/15 total), VII (31 isolates/38 total), and VIII (5 isolates/7 total; collectively 90/110). These distributions were significantly different from what would be expected under the null hypothesis of no association between the soybean host and isolate genotype ( $\chi^2 P < 0.001$ ). The 110 soybean isolates were derived from the USA (56 isolates), Paraguay (42 isolates), Australia (5 isolates), and Denmark (1 isolate), indicating this trend appears to be a general phenomenon not limited to specific geographic areas. We hypothesize that the genotype  $\times$  soybean host association results from increased virulence of these population clusters on soybean, as it was for strawberry.

#### **Evidence of isolate genotype $\times$ host associations was not observed for other clusters, suggesting some populations contain generalists.**

Strawberry and soybean were the only hosts that were sampled sufficiently to enable robust associations between isolate genotype and host origin. However, the available data did not reveal additional isolate genotype  $\times$  host associations and suggest that some clusters have

generalist populations. For example, common bean ( $n = 31$ ) and muskmelon ( $n = 15$ ) were among the most highly sampled hosts after strawberry and soybean, and their isolates were derived from nine and five clusters, respectively. Some clusters contained isolates from a wide diversity of hosts, such as cluster III with 12 hosts out of 17 isolates sampled, or cluster X with 26 hosts out of 57 isolates sampled. These data suggest that true generalist populations exist alongside specialist populations in *Macrophomina* spp.

**Lack of genotypic association with geographic location suggests *Macrophomina* spp. were frequently trafficked between countries**

Overall, SNP-based genotype clustering was not associated with the geographic origins of isolates ( $P = 0.616$  after 10,000 permutations), as indicated by a Mantel test with the latitudes, longitudes and Q-values of 415 isolates with known sub-country regions. Similarly, analysis of molecular variance (AMOVA) identified more variation of Mp11-12 SNP alleles within countries (90.44%) than between countries (9.56%;  $P < 0.001$ ) or between host isolation sources (15.14%;  $P < 0.001$ ). The average calculated fixation index (Hendrick's corrected  $G_{ST}$ ,  $G'_{ST}$ ) among the genotype clusters ( $G'_{ST} = 0.439$ ) was higher than when grouping isolates by country of origin ( $G'_{ST} = 0.155$ ; Supplementary Figure S6), indicating that genetic diversity was lower when isolates were grouped by genotype than by geographic origin.

Similarly, most clusters were found in each of the three well-sampled continents. All SNP-based clusters were found in North America, from which most isolates were derived from USA ( $n = 220$  isolates from the USA out of 230 total from North America). Clusters VI and VIII were exclusively found in the USA; these were the only clusters to be found only in one country or continent. The second-most sampled continent was South America ( $n = 86$  isolates), which

contained all clusters except VI, VIII and XIV. Australia was the third-most extensively sampled continent ( $n = 83$ ), and it collectively contained all clusters except IV, VI, and VIII (Figure 4A). This indicates that isolates from each genotypic cluster are widely dispersed across continents rather than found specifically within certain regions. The weak influence of geography on population structure was visualized with minimum spanning trees generated from the Mp11-12 SNP data, which show isolates were more clearly defined by genotype than by host country (Figures 4B and 4C).

**Recombination appears to be ongoing in *Macrophomina* spp., including between *M. phaseolina* and *M. tecta* and, and could be the result of meiosis**

Evidence of apparently recent and on-going recombination was revealed by fractional ancestries (quantified by Q-values) and LD decay. Most isolates (361 out of 463) had high (>95%) fractional ancestries, meaning that most of their SNP genotype calls corresponded to their assigned cluster (Supplementary Table S1). However, there were also 57 isolates with less than 75% of their fractional ancestry attributed to their assigned cluster, and 30 isolates with less than 50% of their fractional ancestry represented by their assigned cluster. In LD decay plots, the  $r^2$  values declined to around 0.15, after which they tended to remain level, similar to what has been shown for populations with on-going recombination (Supplementary Figure S7) (Feurtey et al. 2023). There was evidence suggesting hybridization occurred between *M. tecta* (cluster II) and *M. phaseolina*, as several cluster II isolates had high ancestry fractions for *M. phaseolina* clusters (Figure 1A; Table 1). Evidence suggesting hybridization between cluster I (*M. pseudophaseolina* and *M. euphorbiicola*) and *M. phaseolina* clusters was less apparent, with Q-values for the assigned cluster among cluster I isolates being at least 90.4% for the Mp11-12

reference. This suggests that recombination could be occurring between *M. tecta* and *M. phaseolina*, but may not be occurring between species in cluster I and *M. phaseolina*.

A heterothallic mode of sexual reproduction for *M. phaseolina* was previously reported by Nagel et al. (2018). This was confirmed here for *M. phaseolina* and newly determined for *M. pseudophaseolina*, *M. euphorbiicola*, and *M. tecta*. Across all isolates there were nearly equal numbers of each mating type (mating type 1 = 238, mating type 2 = 225), but mating type distribution was significantly different from the expected 1:1 ratio for nine clusters ( $\chi^2 P = 0.0442$ ; Table 1). Five clusters had equal mating type ratios ( $\chi^2 P = 0.427$ ; Table 1). While admixed isolates overall had partial memberships in all clusters (Supplementary Table S2), four of these clusters, II, III, VII, and IX, also had relatively low average Q-values corresponding to the main genotype for all references (Supplementary Table S4). This suggests that isolates in these populations were more prone to "outcrossing" with isolates from other clusters. These data indicate that the genetic potential for meiosis exists in *Macrophomina* spp. and is likely to contribute to recombination in these species.

### **Genomic evidence supports the delineation of currently described *Macrophomina* spp., which can be differentiated by new species-specific primers**

Reference-independent, whole genome comparisons of anchor nucleotide distances (ANDI) supported the circumscription of *M. phaseolina*, *M. pseudophaseolina*, *M. euphorbiicola*, and *M. tecta*. The minimum ANDI for any two isolates of different species (= 0.0113) was greater than the maximum ANDI for any two isolates of the same species (= 0.009; Supplementary Figure S4). There was just one isolate (F289) that did not fall into any of these species, as its pairwise ANDI values were all  $\geq 0.0113$ ; above the threshold observed for within-

species comparisons. A multi-locus tree based on a concatenation of four commonly-used phylogenetic markers showed that F289 was most similar to the type strain of *M. vaccinii* (Supplementary Figures S8 and S9). Morphological comparisons between F289 and selected *M. phaseolina* strains showed pycnidia or conidia sizes overlapped with those reported for both *M. phaseolina* and *M. vaccinii* (Zhao et al. 2019) (data not shown), indicating that these characteristics could not clearly differentiate these species. We considered F289 to be a putative *M. vaccinii* isolate but could not confirm this with genomic comparison to known strains of this species.

A series of species-specific qPCR assays were designed and validated to enable rapid identification of the species analyzed in this study. The qPCR assays were first validated *in silico* to target an identified locus that was present in each tested species (*M. phaseolina*, *M. pseudophaseolina*, *M. euphorbiicola*, and *M. tecta*) but absent in off-target organisms in Botryosphaeriaceae and the nr/nt database of Genbank. Each target was tested *in vitro* on a panel of 97 isolates and found to generate the predicted results for all assays, with one exception: the *M. tecta* assay also amplified F289, the one isolate that could not be classified to species. Additionally, the *M. phaseolina* assay did not amplify an individual *M. phaseolina* isolate that was highly admixed with *M. tecta*, as expected (Supplementary Tables S1, S2 and S3). These results highlight the limitations of species boundaries in the genus *Macrophomina*, which may be blurred by inter-specific hybridization.

We also designed a conventional PCR assay that generates a 507 bp amplicon, here called the "mspp locus", that can be sequenced and aligned to known references. This locus is within a predicted fatty acid synthase beta subunit, annotated as gene "M11\_12\_v1\_07558" in Burkhardt et al. (2019b). The mspp PCR assay was validated with representatives of each

*Macrophomina* spp.; the amplicons were sequenced and yielded the expected result (Supplementary Figure S10). Alignment of sequences at the mspp locus can differentiate all *Macrophomina* species and the population cluster XIII with a high degree of accuracy (Table 2; Figure 5). The mspp locus provided a similar amount of phylogenetic information as the standard approach of sequencing and concatenating alignments of ITS, *TEFI- $\alpha$* , *actin*, and *beta-tubulin* sequences, and it is more informative than any of these standard reference genes on their own (Figure 5; Supplementary Figure S9). Notably, the ITS sequence was not sufficient to delineate *Macrophomina* spp., as multiple species are represented by single ITS haplotypes (Supplementary Figure S8).

## Discussion

This detailed look at the factors structuring genetic diversity in *Macrophomina* spp. has highlighted new avenues for research that were previously overlooked. Long touted as being a broad host range generalist, this work shows that host specialization exists in *M. phaseolina* and therefore not all isolates are virulent on all hosts. Research into the diversity of symbiotic relationships between *Macrophomina* spp. and host plants is necessary to fully elucidate these organisms' potential for asymptomatic infection and its influence on geographic distribution. Lastly, the new methods developed to discriminate *Macrophomina* spp. enable broader searches for the rare species in this genus and should help to elucidate their distributions and host ranges.

The hypothesis that *M. phaseolina* is a generalist, broad host range pathogen (reviewed in Pennerman et al. (2024)) implies that the sub-specific genotype is irrelevant, as any isolate could cause disease in any host. Our data provide a contradiction to this by demonstrating that a specific lineage, cluster XIII, is specialized for virulence on strawberry, and that specific clusters

are associated with soybean (Gladieux et al. 2018; Mercier et al. 2019; Resende et al. 1994).

*Macrophomina phaseolina* is best characterized as both a poly-specialist, with multiple populations specialized for different hosts, and a generalist, with populations containing isolates derived from many hosts, no obvious specialization and a capacity for asymptomatic infection.

Although specialized for virulence on strawberry, these isolates are likely to retain the ability to cause disease in other hosts as has been shown for other host specialized pathogens, such as *Verticillium dahliae* (Resende et al. 1994), *Botrytis cinerea* (Mercier et al. 2019; Mercier et al. 2021) and *Magnaporthe oryzae* (Gladieux et al. 2018). Lineages of *M. oryzae* are specialized for pathogenicity in specific grasses; when sampling for *M. oryzae* from natural blast infections, the host-specialized strain is almost always recovered (Ascari et al. 2024; Gladieux et al. 2018), similar to what we observed for strawberry and soybean. However, *M. oryzae* isolates can cause disease in grasses other than the one they are specialized in and can sometimes be recovered from other species (Gladieux et al. 2018; Rahnema et al. 2023). We also observed a small percentage of strawberry- and soybean-derived isolates in other clusters and speculate that isolates in the strawberry- and soybean-associated clusters retain virulence in other hosts. Evaluating the host range of specialized genotypes will be an important avenue of future research.

The emergence of specialized virulence in strawberry appears to be a major factor in the increased frequency and severity of strawberry crown rot disease caused by *M. phaseolina* worldwide. The emergence of this host-specialized strain also coincided with the withdrawal of methyl bromide as a highly effective soil fumigant. Several authors have postulated that *M. phaseolina* emerged in strawberries due to the discontinuation of methyl bromide as a highly effective soil fumigant, partly because epidemics mostly arose in fields that had adopted



substitute fumigants (Hutton et al. 2013; Koike 2008; Zveibil et al. 2012). This hypothesis is not universally supported by the data: side-by-side comparisons often do not show methyl bromide controlled *M. phaseolina* significantly more than replacement fumigants (Chamorro et al. 2016; Hutton et al. 2013; Koike et al. 2013; Zveibil et al. 2012). Furthermore, virulent isolates have been recovered from strawberry nurseries, where methyl bromide continues to be used (Burkhardt et al. 2019a). Our data show that the emergence of a host-specialized strain was a crucial factor for explaining new reports and the surging destructiveness of this pathogen in strawberry in multiple countries.

The strawberry-derived isolates from clusters XI, XII, and XIV were not tested. Some isolates in these three clusters were found to be virulent in previous studies and are associated with strawberry crown rot disease in Argentina (Baino et al. 2011), Australia (Herrington 2018), Chile (Sánchez et al. 2013), Israel (Zveibil and Freeman 2007), and Italy (Gerin et al. 2018). Thus, it is likely that other isolates of *M. phaseolina* could cause disease on strawberry, potentially under different environmental conditions and/or other cultivars than were tested in our study.

The lack of geographic influence on population structure was surprising, given that many pathogens have population structures influenced by geography, even aerially dispersed fungi (Ali et al. 2014; Feurtey et al. 2023; Sotiropoulos et al. 2022). However, many Botryosphaeriaceae pathogens are known to asymptotically colonize plant tissues (Slippers and Wingfield 2007), and this niche has been considered a reason why populations of *Botryosphaeria dothidea* are not structured by geography (Marsberg et al. 2017). The potential for asymptomatic growth on plant tissues or seeds by *M. phaseolina* is known for some hosts (Fuhlborn et al. 2012, 2013; González and Tello 2011; Jamal et al. 2022; Koike et al. 2016; Moslemi et al. 2023), but is a

relatively understudied phenomenon. In support of this hypothesis, our data showed that most *M. phaseolina* isolates could initiate asymptomatic infections in strawberry (Figure 3D). Similarly, we found that cluster XIII isolates of *M. phaseolina* could cause asymptomatic infections in a broad range of Fabaceae and Brassicaceae cover crops (Qin et al. 2025). Asymptomatic infections on plants or seeds are likely to play a much larger role in *M. phaseolina*'s lifecycle than was previously appreciated and warrant greater attention in future research.

The currently described *Macrophomina* species were supported by reference-free whole genome comparisons (Supplementary Figure S4) and phylogenies of core genes (Supplementary Figure S9). However, admixture between species indicate that the boundaries in *Macrophomina* spp. are fluid and could complicate future efforts at specific detection. This led to challenges designing qPCR assays with 100% analytical specificity, as an admixed isolate with *M. phaseolina* and *M. tecta* ancestry did not amplify with either assay. However, our species-specific assays performed well overall and were the first of their kind for *M. pseudophaseolina*, *euphorbiicola*, and *tecta*. The *M. phaseolina* assay provides a slight improvement on the existing qPCR assay (Babu et al. 2011) by including a TaqMan probe in the design, which enables multiplexing with other targets. The existing *M. phaseolina* assay by Babu et al. (2011) was *in silico* predicted to correctly amplify all *M. phaseolina* isolates and not other *Macrophomina* species. Our new *M. phaseolina* assay also has an intentionally broader target than the Burkhardt et al. (2018) assay, which is only predicted to amplify all isolates in cluster XIII except F219.

## Methods

### Global collection and sequencing of *Macrophomina* spp. isolates

Effort was made to collect isolates from diverse hosts, geographic locations, and times. Isolates were drawn from public culture collections (e.g. the Queensland Plant Pathology Herbarium, Westerdijk Institute, NARO Genebank, and the National Reference Repository Library), private collections of individual investigators, and recently published works (Burkhardt et al. 2019a; Weiland 2022). A standardized naming system was applied to 463 isolates received, ranging from F001 to F518. Metadata for each isolate were largely provided by the collectors and are listed in Supplementary Table S1. Longitudes and latitudes were assigned to isolates based on the midpoint of known states, provinces, or localities.

### **DNA extraction and sequencing**

All isolates were single-hyphal-tipped to ensure genetic uniformity. The tissue used for DNA extraction was grown on potato dextrose agar, potato dextrose broth, or 10% V8 juice broth, and was sometimes lyophilized. DNA extractions were performed with the OmniPrep for Fungus kit (G-Biosciences, USA) per manufacturer's instructions, with an added step of treatment by Invitrogen RNase A/T1 Cocktail (Fisher Scientific, Massachusetts, USA) at the final step and incubating for 60 minutes at 45°C followed by 30 minutes at room temperature. Most DNA extracts were then further purified by AMPure XP beads (Beckman Coulter, California, USA) to improve the quality of DNA extracts before library preparation. Libraries were prepared with the DNA *KAPA Hyper Prep Kit* (Kapa Biosystems-Roche, Basel, Switzerland) and sequenced on an Illumina NovaSeq6000 with 150 bp paired-end reads. For isolates from Brazil, DNA was extracted by the method described in Kjærboelling et al. (2018) and sequenced on an Illumina NextSeq2000 at the Max Planck Genome Center in Cologne, Germany. Additional reads were retrieved from NCBI GenBank for 28 isolates; accession

numbers are available in Supplementary Table S1. Estimated genomic sequencing coverage was typically between 72× and 108×, while a small number of isolates had coverage between 26× and 64× (Supplementary Table S1).

### Sequence read processing and mapping

Unless specified, libraries, modules and codes were used at their default parameters. Sequencing reads were preprocessed by removing contaminants, unique molecular identifiers, adaptors, ambiguous and low-quality nucleotides using HTStream version 1.3.3 (<https://github.com/s4hts/HTStream>). BBDMap version 35.85 (<https://sourceforge.net/projects/bbmap/>) was used to limit the total number of bases to a maximum of 5.5 Gb (approximately 100× coverage) with a minimum read length of 100 bases for each isolate. Cleaned reads were mapped to the assembled genomes of *M. phaseolina* strains Mp11-12 (F001), AL-1 (F003) and mp117 (F509) (Burkhardt et al. 2019b; Burkhardt et al. 2018; Gluck-Thaler et al. 2022) using the Burrows-Wheeler Aligner version 0.7.17 (Li and Durbin 2009) and alignments were subsequently processed with SAMtools version 1.16.1 (Danecek et al. 2021). BEDtools version 2.30.0 (Quinlan and Hall 2010) converted the resulting BAM files into BED files. SAMtools flagstat evaluated mapping efficiency. The BED files of mapped read coverages were used to determine if the strawberry *M. phaseolina* detection locus (Burkhardt et al. 2018) and the general *M. phaseolina* detection locus (Babu et al. 2011) were present among the isolates by counting at least five sequencing reads mapped to a total of at least 95% of the length of the reference locus.

### Genome assembly and annotation

Cleaned reads were *de novo* assembled into genomes with SPAdes version 3.15.5 (Prjibelski et al. 2020), and genome assembly quality was evaluated using QUAST version 5.2.0 (Gurevich et al. 2013) and BUSCO version 5.4.5 (Manni et al. 2021) with the dothiideomycetes\_odb10 database (version 2019-11-20), which contained 3,786 benchmarking universal single-copy homologs. SPAdes assemblies were processed through the NCBI Foreign Contamination Screen to remove sequencing adapters and prokaryotic contaminants before submission to NCBI WGS (Astashyn et al. 2024).

Protein-coding genes were annotated from these assemblies with GeneMark-ES version 4.71 (Ter-Hovhannisyan et al. 2008) with the fungus parameter. Proteins with fewer than 100 amino acids were excluded from analyses. These putative proteins were consolidated into a pan-proteome with CD-Hit version 4.8.1 (Fu et al. 2012) with the reciprocal sequence identity threshold and length difference thresholds at 0.8. “Super-core” homologs were represented by every isolate included in the current study. “Core” homologs included super-core homologs and those that were present in at least 95% of isolates. “Accessory” homologs were present in at least 2 isolates and up to 95% of isolates. “Singleton” sequences were only identified in one isolate.

### **Variant calling and population genomic analyses**

Biallelic SNPs with a minimum coverage of five reads for each isolate, a minimum base quality of 20 for each isolate, and a minimum allele frequency of 5% overall among all *Macrophomina* spp. isolates were identified with Freebayes version 1.3.6 (Garrison and Marth 2012). SNP loci missing from more than 10% of used *Macrophomina* spp. isolates were excluded. Admixture was analyzed with the unpruned loci with sNMF version 1.2 (Frichot et al. 2014) with K ranging 2 to 20. A final K of 14 was chosen in accordance with the elbow method

for clustering algorithms. SNP intra-scaffold LD correlations were calculated with VCFTools (Danecek et al. 2011) and the unpruned loci. Fixation indices were computed with R package vcfR (Knaus and Grünwald 2017). The R package poppr was used to generate minimum spanning trees and to conduct AMOVAs (Kamvar et al. 2014).

Sequences corresponding to the *MAT1-1-1* and *MAT1-2-1* genes were extracted from full mating type loci published by Nagel et al. (2018) and then BLASTn searches of these queries against each reference assembly were used to determine which gene was present in each isolate. Taxonomic assignments were verified by retrieval of reference nucleic sequences of ITS (reference NCBI accessions KF951717.1, KF951791.1, OR052068.1, MK687450.1 and NR\_175162.1), *TEFI- $\alpha$*  (KF952088.1, KF952153.1, OR030468.1, MK687426.1 and MW592271.1), *actin* (KF951857.1, KF951918.1, MH712508.1, MK687442.1 and MW592058.1), *beta-tubulin* (KF952178.1, KF952233.1, MF457658.1, MK687434.1 and MW592300.1) genes or genomic regions. These reference sequences represented phylogenetic markers from *M. phaseolina* CPC 21420, *M. pseudophaseolina* CPC 21417, *M. tecta* BRIP:70781, *M. vaccinii* CGMCC3.19503, and *M. euphorbiicola* strains CMM4134, McBr87 and MFLU 23-0004. The reference sequences were used in BLAST searches (Camacho et al. 2009) against our assembled genomes to identify the corresponding sequence in each isolate. The ITS, *TEFI- $\alpha$* , *actin*, and *beta-tubulin* sequences were aligned by MUSCLE (Edgar 2004) and RAxML-NG version 1.2.1 (Kozlov et al. 2019) was used to generate phylogenetic trees with the GTR+G evolution model and 1,000 bootstrap replicates. Phylograms were visualized with the R package: ggtree (Xu et al. 2022). The software *andi* (Haubold et al. 2015) was also used to estimate isolate relatedness with the assembled genomes. Reticulate phylogenetic trees were generated using SplitsTree version 4.19.1 (Huson and Bryant 2005) (NeighborNet function).

Other figures were made with Inkscape 1.3, Python 3.11.4 libraries matplotlib and sklearn, and R 2023.03.1. Heatmaps were generated using Ward's method parameter in the Python Seaborn library and using data that was one-hot encoded. GIS shapefiles were retrieved from the IPUMS database (Ruggles et al. 2024).

### **Development of species-specific qPCR assays and a PCR protocol for a phylogenetically-informative locus**

Phylogenetically-informative sequences commonly shared among all isolates were identified using the BEDTools suite (Quinlan and Hall 2010) to compare BED files of read coverages to the Mp11-12 genomic reference. Species-specific target candidates for qPCR assays were identified by searching the CD-HIT determined homologs for those that were specific to the target *Macrophomina* spp. Putative species-specific homologs were further verified by BLAST searches of the locus against all genome assemblies and NCBI nr/nt and wgs sequence databases.

Species-specific Taqman qPCR assays were designed for *M. phaseolina*, *M. pseudophaseolina*, *M. euphorbiicola*, and *M. tecta*. All assays used a 20 µL reaction mixture containing 1× PerfeCTa qPCR ToughMix (Quantabio, Massachusetts, USA), 0.25 µM of forward and reverse primers (Table 2), 0.15 µM of Taqman probe (Table 2), and approximately 15 ng of DNA. DNA for these reactions was isolated by the OmniPrep kit as described above or by the Quick-DNA Miniprep kit (Zymo Research, California, USA). Reactions were run in a LightCycler II (Roche, USA) with an initial denaturation step of 95°C for 2 min followed by 40 cycles of 15 sec at 95°C and 30 sec at the annealing temperatures provided in Table 2. These qPCR assays were tested on 97 isolates that included at least two representatives of each species and at least five representatives of each *M. phaseolina* population cluster (Supplementary Table

S3). All qPCR assays were predicted to yield exactly zero or one amplicon depending on the isolate (Supplementary Table S1).

An additional protocol was developed for PCR amplification of a single, phylogenetically-informative locus present in all *Macrophomina* spp. that can be sequenced to determine species assignment. This assay was run with a 50 µl reaction mixture containing 0.5 µM of each primer (Table 2), 1× Promega Go Taq Master Mix (#M7123, Promega, Wisconsin, USA), and 2 µL of DNA isolated using the Quick-DNA Fecal-Soil Microbe kit or the Quick-DNA Miniprep kit (Zymo Research, Irvine, CA). Reactions were run in a BioRad C1000 Touch (Bio-Rad Laboratories, California, USA) with an initial denaturation step of 95°C for 5 min followed by 40 cycles of 95°C for 30 sec, 62°C for 30 sec, and 72°C for 1 min, then a final 72°C extension for 5 min. PCR products were purified using the GeneJet PCR Purification Kit (Thermo Fisher Scientific, New Hampshire, USA) and the 507 bp bands were visualized with GelRed Nucleic Acid Stain (Biotum, California, USA) on a 1.5% agarose gel for 40 minutes at 110V. Purified amplicons were sequenced with forward and reverse primers by Eton Biosciences (California, USA).

### **Virulence assays in strawberry**

USA-derived isolates were selected at random from most *M. phaseolina* clusters: IV, V, VIII, X, XI, XII, XIII, and XIV. Isolates were grown in the dark for one week on potato dextrose agar amended with 50 mg/mL kanamycin, 100 mg/mL streptomycin and 12.5 mg/mL chlortetracycline. One 5-mm diameter plug of each culture was transferred to 200 mL of sterile 10% V8 broth and incubated at 25°C and 200 rpm in the dark for three days. The liquid cultures were then blended in a Waring blender for seven seconds and allowed to settle for five minutes



in a 50 mL conical tube before the volumes of settled mycelia were quantified. Suspensions were diluted to 8% mycelia by volume with sterile deionized water.

Bare root, frozen strawberry crowns (cultivar Monterey) from a low-elevation nursery were planted in potting soil (Sunshine Mix #1, Nutrien Ag Solutions, Colorado, USA) and grown at 25°C/20°C day/night temperatures in a greenhouse with no supplemental lighting for one week to break dormancy. After dormancy was broken, the roots were rinsed with tap water and submerged in fungal inoculum for ten minutes. The negative control was submerged in 10% V8 broth for ten minutes. Plants were then repotted in 10.8 cm x 10.8 cm x 12.4 cm pots and watered with the respective leftover inoculum. Inoculated plants were then grown at approximately 28°C/22°C in a greenhouse with day lengths between 14 and 14.6 hours in Salinas, California. For the course of the experiments, plants were watered as needed and fertilized weekly using Peter's Professional 20-20-20 fertilizer (ICL Fertilizers, Missouri, USA). Ten biological replicates for each fungal isolate were established per experiment, and the experiment was conducted twice.

Ordinal disease ratings were conducted weekly for seven weeks on a 1-5 scale, where 1 = asymptomatic, 2 = stunting, 3 = wilting of less than 50% of the foliage, 4 = wilting of more than 50% of foliage, and 5 = plant death. All peduncles bearing fruit and flowers were removed and weighed at five- and seven- weeks post-inoculation. After removing peduncles, fruit, and flowers for separate measurements at seven weeks post-inoculation, whole plants were removed from the soil, roots were trimmed flush with the crown, and above-ground fresh weight was recorded. From each plant, two cross-sections of the crown were excised and their cortical tissues removed (as per Henry et al. (2019)). These crown stele tissues were surface sterilized in 1% sodium hypochlorite solution for two minutes, then plated on potato dextrose agar amended with

kanamycin, streptomycin, and tetracycline as described above. Plates were incubated at 30°C for 5-7 days, and the presence of *M. phaseolina* determined by the morphology of fungal growth emanating from these tissues. Data were analyzed by checking for normality with the Shapiro-Wilk and Levene's tests. The Wilcoxon test identified statistical significance at  $\alpha = 0.05$ .

## References

- Ali, S., Gladieux, P., Leconte, M., Gautier, A., Justesen, A. F., Hovmøller, M. S., Enjalbert, J., and de Vallavieille-Pope, C. 2014. Origin, Migration Routes and Worldwide Population Genetic Structure of the Wheat Yellow Rust Pathogen *Puccinia striiformis* f.sp. *tritici*. *PloS Pathogens* 10:e1003903.
- Ascari, J. P., Cazón, L. I., Rahnama, M., Lamour, K., Fernandez, J. M. C., Farman, M., and Del Ponte, E. M. 2024. *Pyricularia* are mostly host-specialized with limited reciprocal cross-infection between wheat and endemic grasses in Minas Gerais, Brazil. *Phytopathology* 114:226-240.
- Astashyn, A., Tvedte, E. S., Sweeney, D., Sapojnikov, V., Bouk, N., Joukov, V., Mozes, E., Strobe, P. K., Sylla, P. M., Wagner, L., Bidwell, S. L., Brown, L. C., Clark, K., Davis, E. W., Smith-White, B., Hlavina, W., Pruitt, K. D., Schneider, V. A., and Murphy, T. D. 2024. Rapid and sensitive detection of genome contamination at scale with FCS-GX. *Genome Biology* 25:60.
- Avilés, M., Castillo, S., Bascon, J., Zea-Bonilla, T., Martín-Sánchez, P. M., and Pérez-Jiménez, R. M. 2008. First report of *Macrophomina phaseolina* causing crown and root rot of strawberry in Spain. *Plant Pathology* 57:382.
- Babu, B. K., Mesapogu, S., Sharma, A., Somasani, S. R., and Arora, D. K. 2011. Quantitative real-time PCR assay for rapid detection of plant and human pathogenic *Macrophomina phaseolina* from field and environmental samples. *Mycologia* 103:466-473.
- Baggio, J. S., Ruschel, R. G., Noling, J. W., and Peres, N. A. 2021. Physical, cultural, and chemical alternatives for integrated management of charcoal rot of strawberry. *Plant Disease* 105:295-304.
- Baino, O. M., Salazar, S. M., Ramallo, A. C., and Kirschbaum, D. S. 2011. First report of *Macrophomina phaseolina* causing strawberry crown and root rot in northwestern Argentina. *Plant Disease* 95:1477.
- Bennett, R. J., and Turgeon, B. G. 2016. Fungal sex: the *Ascomycota*. *Microbiology Spectrum* 4:FUNK-0005-2016.
- Bradley, C. A., Allen, T. W., Sisson, A. J., Bergstrom, G. C., Bissonnette, K. M., Bond, J., Byamukama, E., Chilvers, M. I., Collins, A. A., Damicone, J. P., Dorance, A. E., Dufault, N. S., Esker, P. D., Faske, T. R., Fiorellino, N. M., Giesler, L. J., Hartman, G. L., Hollier, C. A., Isakeit, T., Jackson-Ziems, T. A., Jardine, D. J., Kelly, H. M., Kemerait, R. C., Kleczewski, N. M., Koehler, A. M., Kratochvil, R. J., Kurle, J. E., Malvick, D. K.,

- Markell, S. G., Mathew, F. M., Mehl, H. L., Mehl, K. M., Mueller, D. S., Mueller, J. D., Nelson, B. D., Overstreet, C., Padgett, G. B., Price, P. P., Sikora, E. J., Small, I., Smith, D. L., T.N., S., Tande, C. A., Telenko, D. E. P., Tenuta, A. U., Thiessen, L. D., Warner, F., Wiebold, W. J., and Wise, K. A. 2021. Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2015 to 2019. *Plant Health Progress* 22:483-495.
- Burkhardt, A., Ramon, M. L., Smith, B., Koike, S. T., and Martin, F. 2018. Development of molecular markers to detect *Macrophomina phaseolina* from strawberry plants and soil. *Phytopathology* 108:1386-1394.
- Burkhardt, A., Childs, K. L., Wang, J., Ramon, M. L., and Martin, F. N. 2019a. Assembly, annotation, and comparison of *Macrophomina phaseolina* isolates from strawberry and other hosts. *BMC Genomics* 20:1-18.
- Burkhardt, A., Henry, P. M., Koike, S. T., Gordon, T. R., and Martin, F. 2019b. Detection of *Fusarium oxysporum* f. sp. *fragariae* from infected strawberry plants. *Plant Disease* 103:1006-1013.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.
- Chamorro, M., Seijo, T. E., Noling, J. C., de los Santos, B., and Peres, N. A. 2016. Efficacy of fumigant treatments and inoculum placement on control of *Macrophomina phaseolina* in strawberry beds. *Crop Protection* 90:163-169.
- Chilakala, A. R., Mali, K. V., Irulappan, V., Patil, B. S., Pandey, P., Rangappa, K., Ramegowda, V., Kumar, M. N., Puli, C. O. R., Mohan-Raju, B., and Senthil-Kumar, M. 2022. Combined drought and heat stress influences the root water relation and determine the dry root rot disease development under field conditions: A study using chickpea genotypes. *Frontiers in Plant Science* 13:890551.
- Cohen, R., Elkabetz, M., Paris, H. S., Gur, A., Dai, N., Rabinovitz, O., and Freeman, S. 2022. Occurrence of *Macrophomina phaseolina* in Israel: Challenges for disease management and crop germplasm enhancement. *Plant Disease* 106:15-25.
- Crous, P. W., Slippers, B., Wingfield, M. J., Rheeder, J., Marasas, W. F. O., Philips, A. J. L., Alves, A., Burgess, T., Barber, P., and Groenewald, J. Z. 2006. Phylogenetic lineages in the *Botryosphaeriaceae*. *Studies in Mycology* 55:235-253.
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., and Li, H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10.
- Danecek, P., Auton, A., Abecasis, G., Albers, A., Banks, E., DePristo, M. A., Handsaker, R., Lunter, G., Marth, G., Sherry, S. T., McVean, G., Durbin, R., and Group, T. G. P. A. 2011. The variant call format and VCFtools. *Bioinformatics* 15:2156-2158.
- de Sousa Linhares, C. M., Ambrósio, M. M. Q., Castro, G., Torres, S. B., Esteras, C., de Sousa Nunes, G. H., and Picó, B. 2020. Effect of temperature on disease severity of charcoal rot of melons caused by *Macrophomina phaseolina*: implications for selection of resistance sources. *European Journal of Plant Pathology* 158:431-441.
- Dell'Olmo, E., Zaccardelli, M., Basile, B., Corrado, G., and Sigillo, L. 2023. Identification and Characterization of New Seedborne Pathogens in *Phaseolus vulgaris* Landraces of Southern Italy. *Pathogens* 12:108.

- Diourte, M., Starr, J. L., Jeger, M. J., Stack, J. P., and Rosenow, D. T. 1995. Charcoal rot (*Macrophomina phaseolina*) resistance and the effects of water stress on disease development in sorghum. *Plant Pathology* 44:196-202.
- Edel-Hermann, V., and Lecompte, C. 2019. Current status of *Fusarium oxysporum formae speciales* and races. *Phytopathology* 109:512-530.
- Edgar, R. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32:1792-1797.
- Feurtey, A., Lorrain, C., McDonald, M. C., Milgate, A., Solomon, P. S., Warren, R., Puccetti, G., Scalliet, G., Torriani, S. F. F., Gout, L., Marcel, T. C., Suffert, F., Alassimone, J., Lipzen, A., Yoshinaga, Y., Daum, C., Barry, K., Grigoriev, I. V., Goodwin, S. B., Genissel, A., Seidl, M. F., Stukenbrock, E. H., Lebrun, M.-H., Kema, G. H. J., McDonald, B. A., and Croll, D. 2023. A thousand-genome panel retraces the global spread and adaptation of a major fungal crop pathogen. *Nature Communications* 14:1059.
- Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., and François, O. 2014. Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196:973-983.
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. 2012. CD-HIT: accelerated for clustering the next generation sequencing data. *Bioinformatics* 28:3150-3152.
- Fuhlborn, M. J., Ryley, M. J., and Aitken, E. A. B. 2012. New weed hosts of *Macrophomina phaseolina* in Australia. *Australasian Plant Disease Notes* 7:193-195.
- Fuhlborn, M. J., Ryley, M. J., and Aitken, E. A. B. 2013. Infection of mungbean seed by *Macrophomina phaseolina* is more likely to result from localized pod infection than from systemic plant infection. *Plant Pathology* 62:1271-1284.
- Garrison, E., and Marth, G. 2012. Haplotype-based variant detection from short-read sequencing. *arXiv preprint 1207.3907v2:1-9*.
- Gerin, D., Dongiovanni, C., De Miccolis Angelini, R. M., Pollastro, S., and Faretra, F. 2018. First report of *Macrophomina phaseolina* causing crown and root rot on strawberry in Italy. *Plant Disease* 102:1857.
- Gladieux, P., Condon, B., Ravel, S., Soanes, D., Maciel, J. L. N., Nhani, A., Chen, L., Terauchi, R., Lebrun, M.-H., Tharreau, D., Mitchell, T., Pedley, K. F., Valent, B., Talbot, N., Farman, M., and Fournier, E. 2018. Gene flow between divergent cereal- and grass-specific lineages of the rice blast fungus *Magnaporthe oryzae*. *mBio* 9.
- Gluck-Thaler, E., Ralston, T., Konkel, Z., Ocampos, C. G., Ganeshan, V. D., Dorrance, A. E., Niblack, T. L., Wood, C. W., Slot, J. C., Lopez-Nicora, H. D., and Vogan, A. A. 2022. Giant *Starship* elements mobilize accessory genes in fungal genomes. *Molecular Biology and Evolution* 39:msac109.
- Golzar, H., Phillips, D., and Mack, S. 2007. Occurrence of strawberry root and crown rot in Western Australia. *Australasian Plant Disease Notes* 2:145-147.
- Gonçalves, A. P., and Glass, N. L. 2020. Fungal social barriers: to fuse, or not to fuse, that is the question. *Communicative & Integrative Biology* 13:39-42.
- González, V., and Tello, M. L. 2011. The endophytic mycota associated with *Vitis vinifera* in central Spain. *Fungal Diversity* 47:29-42.
- Gupta, G. K., Sharma, S. K., and Ramteke, R. 2012. Biology, epidemiology and management of the pathogenic fungus *Macrophomina phaseolina* (Tassi) Goid with special reference to charcoal rot of soybean (*Glycine max* (L.) Merrill). *Journal of Phytopathology* 160:167-180.

- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. 2013. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072-1075.
- Haubold, B., Klötzl, F., and Pfaffelhuber, P. 2015. andi: fast and accurate estimation of evolutionary distances between closely related genomes. *Bioinformatics* 15:1169-1175.
- Henry, P. M., Pastrana, A. M., Leveau, J. H. J., and Gordon, T. R. 2019. Persistence of *Fusarium oxysporum* f. sp. *fragariae* in soil through asymptomatic colonization of rotation crops. *Phytopathology* 109:770-779.
- Herrington, M. E. 2018. National Strawberry Varietal Improvement Program. Hort Innovation, Level 8, 1 Chifley Square, Sydney, New South Wales 2000.
- Huson, D. H., and Bryant, D. 2005. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23:254-267.
- Hutton, D. G., Gomez, A. O., and Mattner, S. W. 2013. *Macrophomina phaseolina* and its association with strawberry crown rot in Australia. *International Journal of Fruit Science* 13:149-155.
- Jamal, H. A. A., Husaini, A., Sing, N. N., Roslan, H. A., Zulkharnain, A., and Akinkunmi, W. A. 2022. Characterization of bioactive compounds produced by endophytic fungi isolated from *Gynura procumbens* (*Sambung Nyawa*). *Brazilian Journal of Microbiology* 53:1857-1870.
- Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281.
- Kjærboelling, I., Vesth, T. C., Frisvad, J. C., Nybo, J. L., Theobald, S., Kuo, A., Bowyer, P., Matsuda, Y., Mondo, S., Lyhne, E. K., Kogle, M. E., Clum, A., Lipzen, A., Salamov, A., Ngan, C. Y., Daum, C., Chiniquy, J., Barry, K., LaButti, K., Haridas, S., Simmons, B. A., Magnusson, A., Mortensen, U. H., Larsen, T. O., Grigoriev, I., Baker, S. E., and Andersen, M. R. 2018. Linking secondary metabolites to gene clusters through genome sequencing of six diverse *Aspergillus* species. *Proceedings of the National Academy of Sciences* 115:E753-E761.
- Knaus, B. J., and Grünwald, N. J. 2017. vcfr: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources* 17:44-53.
- Koike, S. T. 2008. Crown rot of strawberry caused by *Macrophomina phaseolina* in California. *Plant Disease* 92:1253.
- Koike, S. T., Arias, R. S., Hogan, C. S., Martin, F. N., and Gordon, T. R. 2016. Status of *Macrophomina phaseolina* on strawberry in California and preliminary characterization of the pathogen. *International Journal of Fruit Science* 16:148-159.
- Koike, S. T., Gordon, T. R., Daugovish, O., Ajwa, H., Bolda, M. P., and Subbarao, K. 2013. Recent developments on strawberry plant collapse problems in California caused by *Fusarium* and *Macrophomina*. *International Journal of Fruit Science* 13:76-83.
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35:4453-4455.
- Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760.
- Machado, A. R., Pinho, D. B., Soares, D. J., Gomes, A. A. M., and Pereira, O. L. 2019. Bayesian analyses of five gene regions reveal a new phylogenetic species of *Macrophomina*

- associated with charcoal rot on oilseed crops in Brazil. *European Journal of Plant Pathology* 153:89-100.
- Manici, L. M., Caputo, F., and Cerato, C. 1995. Temperature responses of isolates of *Macrophomina phaseolina* from different climatic regions of sunflower production in Italy. *Plant Disease* 79:834-838.
- Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and Zdobnov, E. M. 2021. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution* 38:4647-4654.
- Marsberg, A., Kemler, M., Jami, F., Nagel, J. H., Potsma-Smidt, A., Naidoo, S., Wingfield, M. J., Crous, P. W., Spatafora, J. W., Hesse, C. N., Robbertse, B., and Slippers, B. 2017. *Botryosphaeria dothidea*: a latent pathogen of global importance to woody plant health. *Molecular Plant Pathology* 18:477-488.
- Martin, F., Koike, S. T., Arias, R. S., and Ramon, M. L. 2015. Population structure and host preference of *Macrophomina phaseolina* on strawberry. *Phytopathology* 105:89.
- McDonald, B. A., and Linde, C. 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* 40:349-379.
- Mengitsu, A., Arelli, P. A., Bellaloui, N., Bond, J. P., Shannon, G. J., Wrather, A. J., Rupe, J. C., Chen, P., Little, C. R., Canaday, C. H., Newman, M. A., and Pantalone, V. R. 2012. Evaluation of soybean genotypes for resistance to three seed-borne diseases. *Plant Health Progress*.
- Mercier, A., Carpentier, F., Duplaix, C., Auger, A., Pradier, J.-M., Viaud, M., Gladieux, P., and Walker, A.-S. 2019. The polyphagous plant pathogenic fungus *Botrytis cinerea* encompasses host-specialized and generalist populations. *Environmental Microbiology* 21:4808-4821.
- Mercier, A., Simon, A., Lapalu, N., Giraud, T., Bardin, M., Walker, A.-S., Viaud, M., and Gladieux, P. 2021. Population genomics reveals molecular determinants of specialization to tomato in the polyphagous fungal pathogen *Botrytis cinerea* in France. *Phytopathology* 111:2355-2366.
- Mertely, J., Seijo, T., and Peres, N. 2005. First report of *Macrophomina phaseolina* causing a crown rot of strawberry in Florida. *Plant Disease* 89:434.
- Moslemi, M., Ghaderi, F., Karami, S., and Charehgani, H. 2023. Identification and pathogenicity of *Macrophomina* species in sesame fields from Iran. *Tropical Plant Pathology*.
- Nagel, J. H., Wingfield, M. J., and Slippers, B. 2018. Evolution of the mating types and mating strategies in prominent genera in the *Botryosphaeriaceae*. *Fungal Genetics and Biology* 114:24-33.
- Ortiz, V., Chang, H.-X., Sang, H., Jacobs, J., Malvick, D. K., Baird, R., Mathew, F. M., de Jensen, C. E., Wise, K. A., Mosquera, G. M., and Chilvers, M. I. 2023. Population genomic analysis reveals geographic structure and climatic diversification for *Macrophomina phaseolina* isolated from soybean and dry bean across the United States, Puerto Rico, and Colombia. *Frontiers in Genetics* 14:1103969.
- Pennerman, K. K., Dilla-Ermita, C. J., and Henry, P. M. 2024. Exaggerated plurivory of *Macrophomina phaseolina*: Accounting for the large host range claim and the shifting of scientific language. *Phytopathology* 114:119-125.
- Pereira, T. S., Machado Esquissato, G. N., da Silva Franco, C. C., de Freitas Mathias, P. C., Bock, C. H., de Carvalho Nunes, W. M., Zanutto, C. A., and de Castro-Prado, M. A. A.

2018. Heterokaryosis and diploid formation among Brazilian isolates of *Macrophomina phaseolina*. *Plant Pathology* 67:1857-1864.
- Poudel, B., and Vaghefi, N. 2023. Taxonomy of *Macrophomina*-traditional to molecular approaches. Page 6, P. Kumar and R. C. Dubey, eds. Academic Press.
- Poudel, B., Shivas, R. C., Adorada, D. L., Barbetti, M. J., Bithell, S. L., Kelly, L. A., Moore, N., Sparks, A. H., Tan, Y. P., Thomas, G., Van Leur, J., and Vaghefi, N. 2021. Hidden diversity of *Macrophomina* associated with broadacre and horticultural crops in Australia. *European Journal of Plant Pathology* 161:1-23.
- Prijbelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. 2020. Using SPAdes de novo assembler. *Current Protocols in Bioinformatics* 70:e102.
- Qin, C., Goldman, P. H., Leap, J., and Henry, P. M. 2025. Cover cropping attenuates population growth of *Macrophomina phaseolina* by limiting weed biomass, despite asymptomatic colonization of cover crops. *Plant Disease* 109:480-490.
- Quinlan, A. R., and Hall, I. M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841-842.
- Rahnama, M., Condon, B., Ascari, J. P., Dupuis, J. R., Del Ponte, E. M., Pedley, K. F., Martinez, S., Valent, B., and Farman, M. 2023. Recent co-evolution of two pandemic plant diseases in a multi-hybrid swarm. *Nature Ecology & Evolution* 7:2055-2066.
- Resende, M. L. V., Flood, J., and Cooper, R. M. 1994. Host specialization of *Verticillium dahliae*, with emphasis on isolates from cocoa (*Theobroma cacao*). *Plant Pathology* 43:104-111.
- Reuveni, R., Nachmias, A., and Krikun, J. 1983. The role of seedborne inoculum on the development of *Macrophomina phaseolina* on melon. *Plant Disease* 67:280-281.
- Reznikov, S., Chiesa, M. A., Pardo, E. M., De Lisi, V., Bogado, N., González, V., Ledesma, F., Morandi, E. N., Ploper, L. D., and Castagnaro, A. P. 2019. Soybean-*Macrophomina phaseolina*-specific interactions and identification of a novel source of resistance. *Phytopathology* 109:63-73.
- Ruggles, S., Cleveland, L., Lovaton, R., Sarkar, S., Sobek, M., Burk, D., Ehrlich, D., Heimann, Q., and Lee, J. 2024. Integrated Public Use Microdata Series, International: Version 7.5. Minneapolis, MN.
- Sánchez, S., Gambardella, M., Henríquez, J. L., and Díaz, I. 2013. First Report of Crown Rot of Strawberry Caused by *Macrophomina phaseolina* in Chile. *Plant Disease* 97:996.
- Sánchez, S., Chamorro, M., Henríquez, J. L., Grez, J., Díaz, I., de los Santos, B., and Gambardella, M. 2017. Genetic and biological characterization of *Macrophomina phaseolina* (Tassi) Goid. causing crown and root rot of strawberry. 77:325-331.
- Sarr, M. P., Ndiaye, M., Groenewald, J. Z., and Crous, P. W. 2014. Genetic diversity in *Macrophomina phaseolina*, the causal agent of charcoal rot. *Phytopathologia Mediterranea* 53:250-268.
- Schertler, A., Lenzner, B., Dullinger, S., Moser, D., Bufford, J. L., Ghelardini, L., Santini, A., Capinha, C., Monteiro, M., Reino, L., Wingfield, M. J., Seebens, H., Thines, M., Dawson, W., van Kleunen, M., Kreft, H., Pergl, J., Pyšek, P., Weigelt, P., Winter, M., and Essl, F. 2024. Biogeography and global flows of 100 major alien fungal and fungus-like oomycete pathogens. *Journal of Biogeography* 51:599-617.
- Sikes, B. A., Bufford, J. L., Hulme, P. E., Cooper, J. A., Johnston, P. R., and Duncan, R. P. 2018. Import volumes and biosecurity interventions shape the arrival rate of fungal pathogens. *PLoS Biology* 16:e2006025.

- Singh, T., and Singh, D. 1982. Transmission of seed-borne inoculum of *Macrophomina phaseolina* from seed to plant. *Proceedings of the Indian Academy of Sciences* 91:357-370.
- Slippers, B., and Wingfield, M. J. 2007. Botryosphaeriaceae as endophytes and latent pathogens of woody plants: diversity, ecology and impact. *Fungal Biology Reviews* 21:90-106.
- Sotiropoulos, A. G., Arango-Isaza, E., Ban, T., Barbieri, C., Bourras, S., Cowger, C., Czembor, P. C., Ben-David, R., Dinoor, A., Ellwood, S. R., Graf, J., Hatta, K., Helguera, M., Sánchez-Martín, J., McDonald, B. A., Morgounov, A. I., Müller, M. C., Shamanin, V., Shimizu, K. K., Yoshihira, T., Zbinden, H., Keller, B., and Wicker, T. 2022. Global genomic analyses of wheat powdery mildew reveal association of pathogen spread with historical human migration and trade. *Nature Communications* 13:4315.
- Steele, M. E., Mendez, M., Hewavitharana, S. S., and Holmes, G. J. 2023. Survey of soilborne pathogens infecting strawberry in Santa Maria, California. *International Journal of Fruit Science* 23:256-266.
- Su, G., Suh, S.-O., Schneider, R. W., and Russin, J. S. 2001. Host specialization in the charcoal rot fungus, *Macrophomina phaseolina*. *Phytopathology* 91:120-126.
- Sybuia, P. A., de Castro-Prado, G., de Carvalho Nunes, W. M., Zanutto, C. A., Kaneshima, E. N., Soares, D. J., da Silva Franco, C. C., de Freitas Mathias, P. C., and de Castro-Prado, M. A. A. 2022. Parasexual recombination in *Macrophomina pseudophaseolina* and vegetative compatibility reactions in *M. euphorbiicola*. *European Journal of Plant Pathology* 163:937-950.
- Taylor, D. L., Walters, W. A., Lennon, N. J., Bochicchio, J., Krohn, A., Caporaso, J. G., and Pennanen, T. 2016. Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing. *Applied and Environmental Microbiology* 82:7217-7226.
- Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O., and Borodovsky, M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research* 18:1979-1990.
- Urquhart, A. S., Gluck-Thaler, E., and Vogan, A. A. 2024. Gene acquisition by giant transposons primes eukaryotes for rapid evolution via horizontal gene transfer. *Science Advances* 10:eadp8738.
- Viejobueno, J., de los Santos, B., Camacho-Sanchez, M., Aguado, A., Camacho, M., and Salazar, S. M. 2022. Phenotypic variability and genetic diversity of the pathogenic fungus *Macrophomina phaseolina* from several hosts and host specialization in strawberry. *Current Microbiology* 79:189.
- Weiland, J. E. 2022. First Report of *Macrophomina phaseolina* Causing Charcoal Root Rot of Hebe (*Veronica cupressoides*, *V. ochracea*, and *V. pinguifolia*) in Oregon, U.S.A. *Plant Disease* 106:1984.
- Xu, S., Li, L., Luo, X., Chen, M., Tang, W., Zhan, L., Dai, Z., Lam, T. T., Guan, Y., and Yu, G. 2022. ggtree: A serialized data object for visualization of a phylogenetic tree and annotation data. *iMeta* 1:e56.
- Zhao, L., Cai, J., He, W., and Zhang, Y. 2019. *Macrophomina vaccinii* sp. nov. causing blueberry stem blight in China. *MycKeys* 55:1-14.
- Zveibil, A., and Freeman, S. 2007. First report of crown and root rot in strawberry caused by *Macrophomina phaseolina* in Israel. *Plant Disease* 89:1014.



Zveibil, A., Mor, N., Gnayem, N., and Freeman, S. 2012. Survival, host-pathogen interaction, and management of *Macrophomina phaseolina* on strawberry in Israel. Plant Disease 96:265-272.

### **Author contributions**

P.M.H. conceptualized the study. P.M.H., P.G., M.A., C.B., A.G., J.N., M.C., V.O., E.H.S, D.P., G.H.G., A.M., H.D.L.-N., N.V., L.K., J.P.B., A.R.M., T.S., N.A.P., J.B., K.I., G.C., S.K., D.M., S.M., and M.G. obtained and curated isolates for sequencing. P.M.H., G.H.G., J.P.B., A.R.M., D.P., E.H.S., and M.G. conducted whole genome sequencing. K.K.P, P.G., C.J.D and P.M.H performed data curation. K.K.P., P.G., C.J.D, G.R., J.H.J., J.L.-H., J.R., G.O.S., and P.M.H. conducted lab work including DNA extractions and pathogenicity study designs and/or executions. K.K.P. and P.M.H. analyzed data and wrote the manuscript. All authors edited and approved the manuscript.

### **Competing interests**

There are no competing interests to report.

### **Data availability**

Raw sequencing reads were deposited in the NCBI SRA database with BioProject accession PRJNA953043. Individual isolate metadata and SRA accessions are included as part of Supplementary Table S1. Code associated with the analyses reported here are available at the Zenodo repository <https://zenodo.org/records/13367157> with DOI: 10.5281/zenodo.13367156.

### **Acknowledgements**

We thank our funding sources, including the United States Department of Agriculture (USDA) Agricultural Research Service (ARS), the National Institute of Food and Agriculture Specialty Crops Research Initiative (#2017-51181-26833 and #2022-51181-38328) and the Technical University of Munich - TUM Institute for Advanced Study, funded by the German Excellence Initiative for a Hans Fischer fellowship awarded to G.H.G. This research was supported in part by an appointment to the ARS Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and the USDA. ORISE is managed by ORAU under DOE contract number DE-SC0014664. All opinions expressed in this paper are the authors' and do not necessarily reflect the policies and views of USDA, DOE, or ORAU/ORISE.

Tables and Supplementary Information Legends

**Table 1. Mating type ratios of *Macrophomina* spp. genotype clusters.**

Cluster <sup>a</sup>	<i>MAT1-1-I</i> <sup>b</sup>	<i>MAT1-2-I</i> <sup>c</sup>	$\chi^2$ P-value <sup>d</sup>
I	5	7	0.564
II	7	4	0.545
III	15	2	0.0036
IV	8	1	0.0442
V	40	1	< 0.0001
VI	0	15	0.0003
VII	19	19	1
VIII	10	0	0.0016
IX	8	9	1
X	25	32	0.427
XI	16	1	0.0007
XII	34	0	< 0.0001
XIII	0	134	< 0.0001
XIV	51	0	< 0.0001

<sup>a</sup> The population cluster as defined by the highest percent assignment using the Mp11-12 reference.

<sup>b</sup> The number of isolates in each cluster with the *MAT1-1-I* gene, indicating a *MAT1-1* locus.

<sup>c</sup> The number of isolates in each cluster with the *MAT1-2-I* gene, indicating a *MAT1-2* locus.

<sup>d</sup> The *P*-value for a chi-square ( $\chi^2$ ) test with the null hypothesis being equal counts of the mating types for each cluster.

**Table 2. Primers and probes designed for detection and differentiation of *Macrophomina* spp.**

Primer ID <sup>a</sup>	Sequence (5' to 3') <sup>b</sup>	Targeted species	Annealing temp (°C)	Amplicon size (bp)
mp_F	TAGCAGCTGCACTTCCCAA	<i>M. phaseolina</i>	61	78
mp_R	CTTCTCCTATTTCTGGTGCC			
mp_P	/6-FAM/CTACGGGCG/ZEN/CCGATGAGGTC/IABkFQ/			
me_F	CACATTTTCATGCGCTACGTC	<i>M. euphorbiicola</i>	64	192
me_R	CAACGCTGGTCGAGATGAT			
me_P	/6-FAM/ACGTGCTGC/ZEN/TGGTACGCGCC/IABkFQ/			
ms_F	TTCCAGAAGTTTGCTCTGGC	<i>M. pseudophaseolina</i>	66	175
ms_R	CCTCCGTAGATGAGCACTTT			
ms_P	/6-FAM/CCCGCGTCGGTTGCTCTGGA/BHQ-1/			
mt_F	ACTCTACAACCCAACGCCAA	<i>M. tecta</i>	62	203
mt_R	CATCGTCATCAATGGGCTC			
mt_P	/6-FAM/AAGCGCGAA/ZEN/ACCCGCGCGAA/IABkFQ/			
mspp_F	GCCACATTGATGCCTTGAAG	<i>Macrophomina</i>	62	507
mspp_R	GGTAGCGAACTTGAGCAAG	spp. (to sequence)		

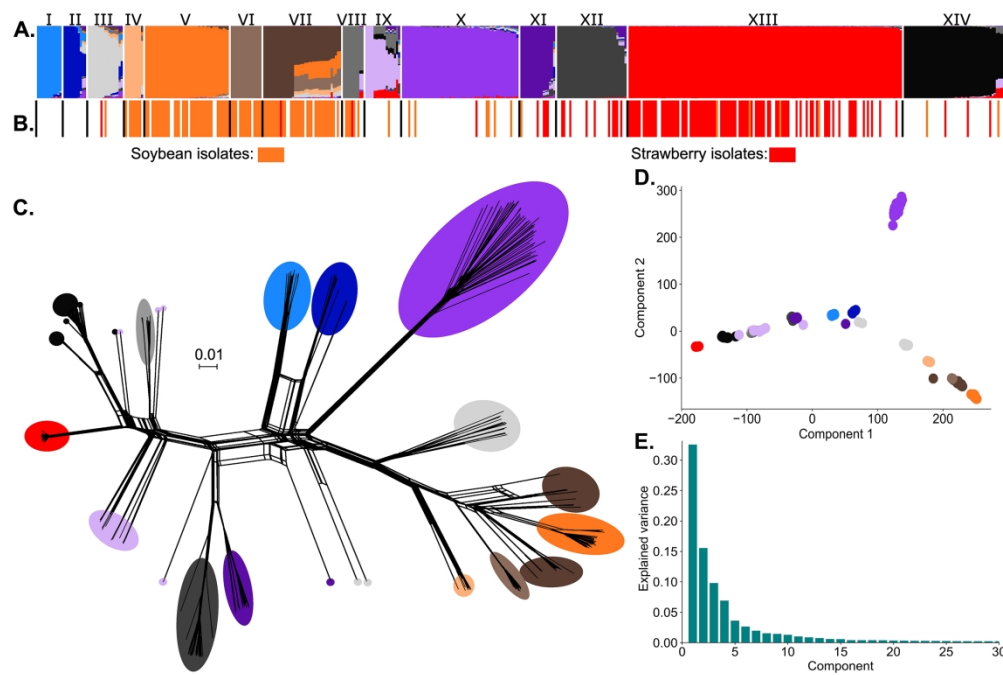
<sup>a</sup> Primer IDs with an “\_F” denote a forward primer; “\_R”, reverse; “\_P”, probe.

<sup>b</sup> The primer or probe sequence used in each assay. Fluorophores and quenchers in probes are delineated with “/” and contain non-IUPAC characters.

**Supplementary Information S1. Multiple sequence alignment of the ITS regions of *Macrophomina* spp. isolates and reference genomes.** Sequences were de-duplicated, and headers contain the IDs of represented isolate(s) and/or reference(s).

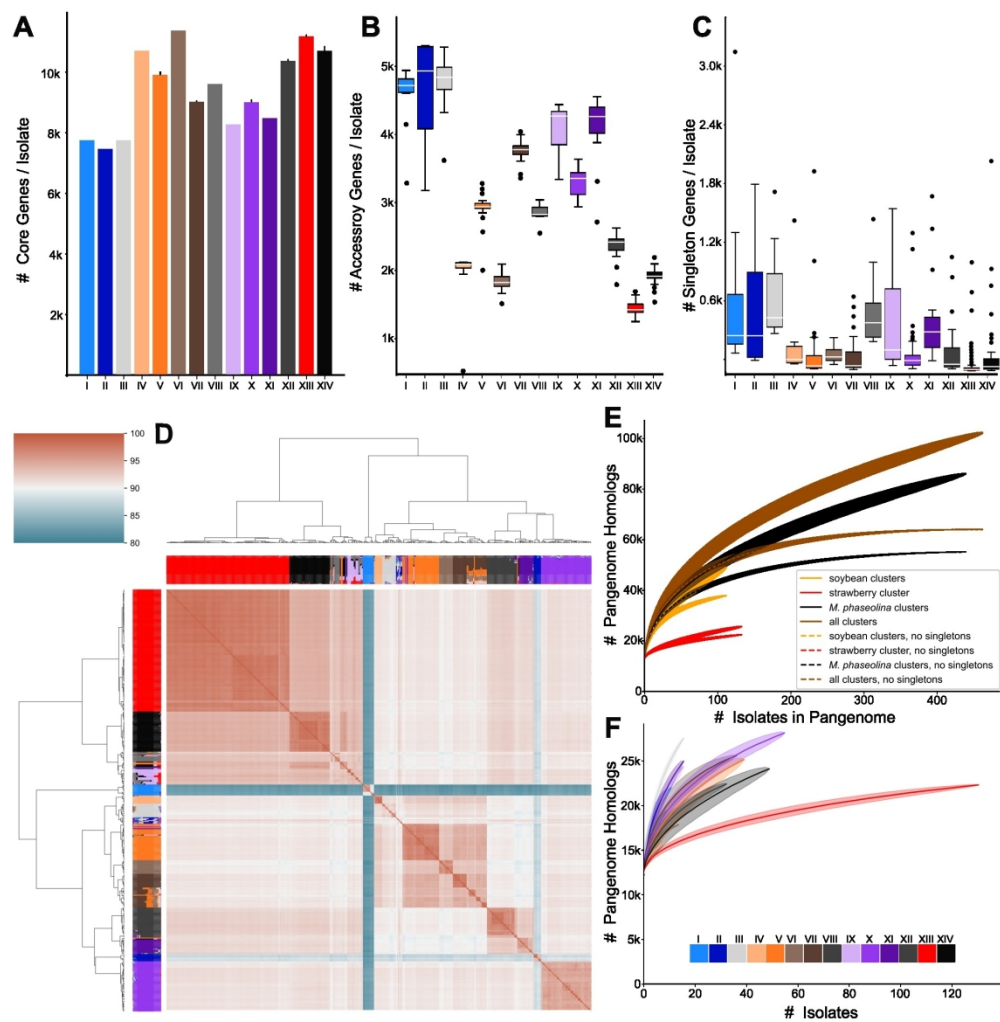
**Supplementary Information S2. Multiple sequence alignment of the concatenated ITS, *actin*, *beta-tubulin*, and *TEF1- $\alpha$*  sequences of *Macrophomina* spp. isolates and reference genomes.** Sequences were de-duplicated, and headers contain the IDs of represented isolate(s) and/or reference(s).

**Supplementary Information S3. Multiple sequence alignment of the *mspp* region of *Macrophomina* spp. isolates.** Sequences were de-duplicated, and headers contain the IDs of represented isolate(s).



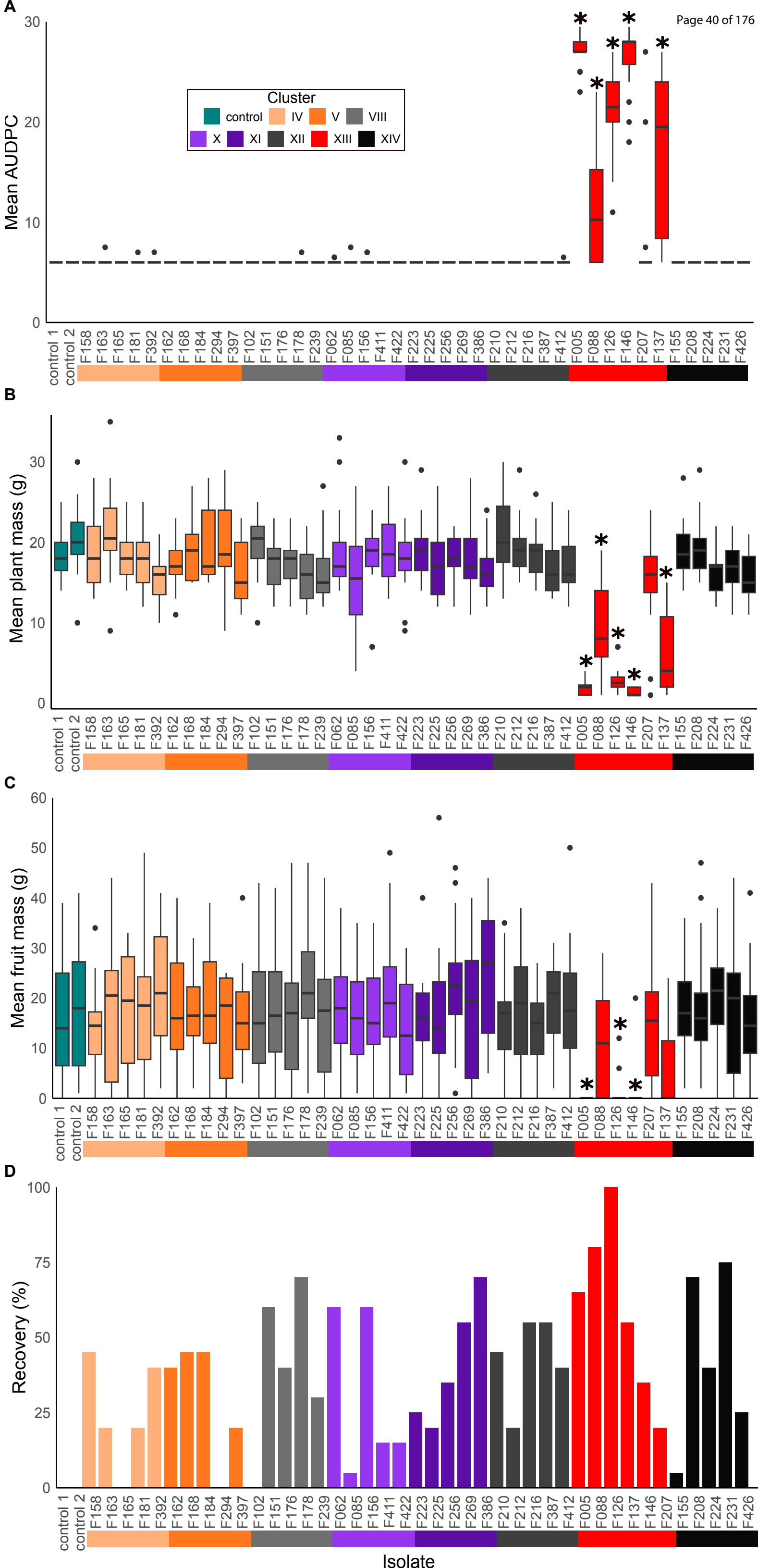
**Fig. 1.** Grouping of *Macrophomina* spp. isolates with Mp11-12-based SNPs. The admixture plot (A) was generated with sNMF at K=14. From left to right, the clusters are labeled with Roman numerals and color-coded. (B) Colored bars indicate isolation source for each isolate in the above admixture plot, where strawberry = red, soybean or soybean field soil = orange, and cluster borders = black. The reticulate phylogenetic tree (C) and PCA plot (D) use the same colors as the admixture plot to indicate the major cluster assignment. The Scree plot (E) shows the eigenvalues for the first 30 principal components.

226x151mm (300 x 300 DPI)

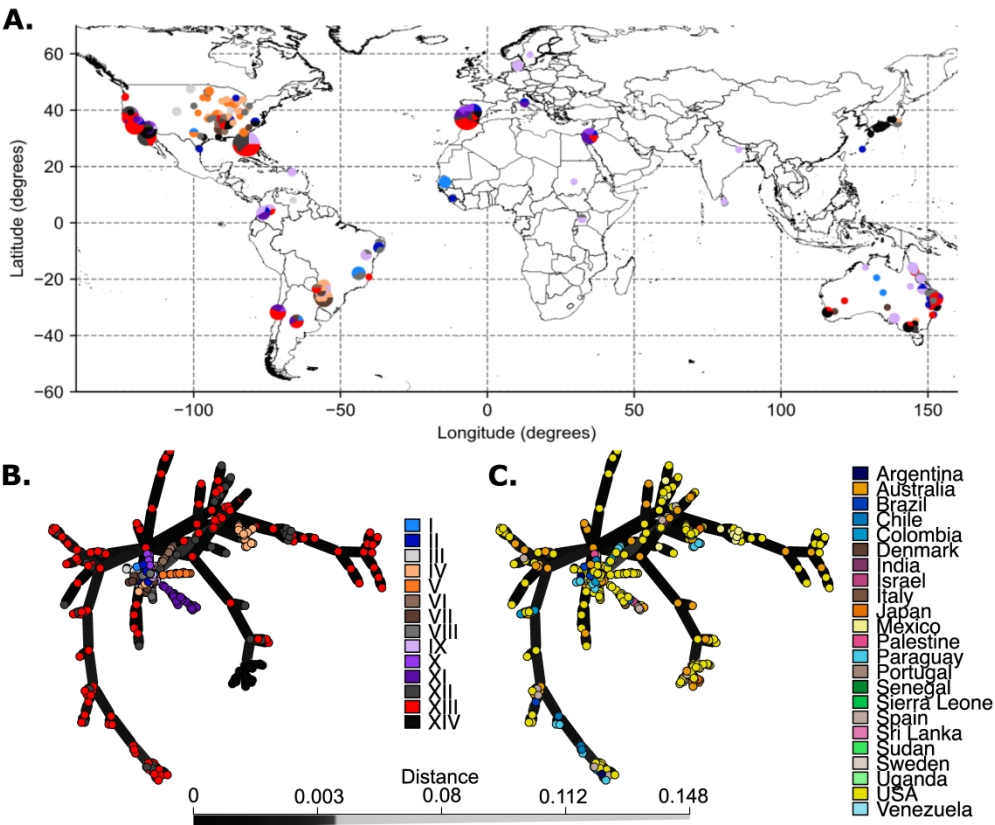


**Fig. 2.** Homolog presence/absence variation by population cluster. **A.** Barplot depicting the average number of core gene homologs (present in >95% of isolates) for each population cluster. Error bars represent one standard deviation above and below the mean. **B.** Boxplot depicting the number of accessory gene homologs (present in more than 2 isolates, less than 95%) for each population cluster. **C.** The number of singletons (genes present in a single isolate) for isolates in each population cluster. **D.** A hierarchically clustered heatmap of core and accessory homolog presence/absence variation. Color intensity within the heatmap indicates the percent of shared homologs for each pair of isolates. The corresponding Mp11-12 SNP-based admixture plot is displayed between the dendrogram and heatmap. **E.** Rarefaction plot of homologs among all isolates or in selected groups, with and without singletons. **F.** Rarefaction plot of gene homologs for each population cluster. In both E and F, shaded regions surrounding the means represent one standard deviation above and below the mean with 1,000 randomized replications.

176x180mm (300 x 300 DPI)

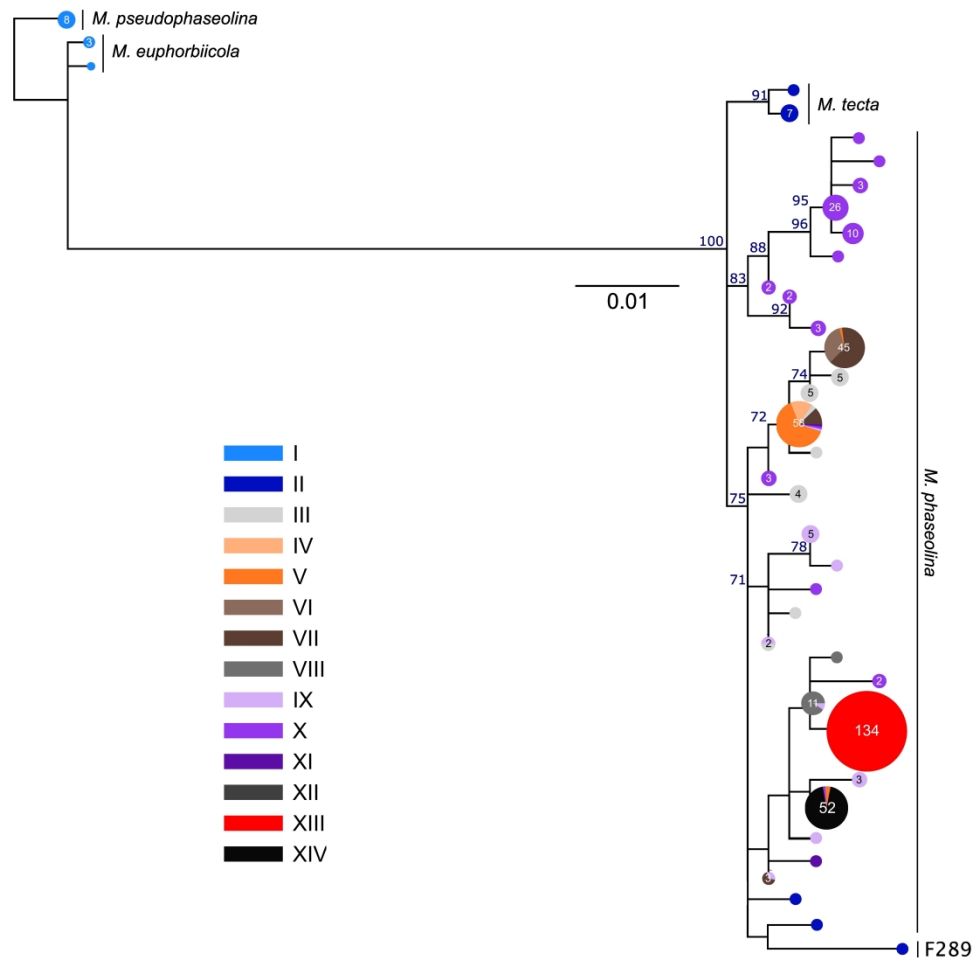






**Fig. 4.** Geographic distribution of *Macrophomina* spp. genotype clusters. **A.** Isolates without a known state, province, or locality were assigned longitudes and latitudes at the midpoints of their respective sub-country region (Supplementary Table 1). Pie charts scaled to the number of isolates recovered are shown for locations that had multiple isolates. The color-code for the pie charts is the same as provided for panel B of this figure. Minimum spanning trees for isolates in this study colored by genotype cluster (**B**) or the country of origin (**C**).

176x147mm (900 x 900 DPI)



**Fig. 5.** Phylogenetic tree built from sequences at the ~500 bp locus targeted by the "mspp" PCR primers. The phylogenetic tree was calculated by RAxML-NG version 1.2.1 with the GTR+G model of evolution and 1,000 bootstrap replicates. Only bootstrap values greater than 70 are shown. Species *M. euphorbiicola*, *M. pseudophaseolina*, *M. tecta*, and *M. phaseolina* are distinguishable and indicated by text annotations. All isolates from the strawberry-associated cluster XIII had identical sequences that were distinct from other *Macrophomina* spp. isolates. Used sequences are in Supplementary Information S3.

161x161mm (900 x 900 DPI)

>Bd02

GATCCTCCCACCCTTTG-TGTACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCC-  
 TCCGCGGCCGCCCCCCTCCCCGGG  
 GGGTGGCCAGCGCCCGCCAGAGGACCATCAAACCTCCAGTCAGTAAACGATGCAGTC  
 TGA AAA---ACATTTAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGAAGGGC  
 ATGCCTGTTCGAGCGTCATTACA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCTCAA  
 GCGTAGTAGAACATACATCTCGCTTCGGAGCGCAGGGCGTCGCCCCCGCGGACGAAC  
 CTTCTGAACTTTTCTCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F084

GATCCTCCCACCCTTTGTTATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTTCC  
 GCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTCGAGCGTCATTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F115,F114,F113,F448,F042,F331,F112,F449,F432,F264,Mpseudophaseolina,Meuphorbiicola

GATCCTCCCACCCTTTG-  
 TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTTCCGCGGCCGCCCCCGATTT  
 TGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTCGAGCGTCATTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F110

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTCCTGAACTTTTC  
 TCAAGGTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F057,F518,F224,F407,F259,F239,F273,F271,F178,F203,F347,F367,F353,F189,F311,F155,F2  
 10,F436,F013,F380,F426,F016,F021,F209,F060,F284,F342,F214,F056,F283,F230,F226,F280,F  
 049,F216,F361,F253,F379,F212,F368,F205,F022,F174,F176,F072,F296,F262,F026,F444,F297,  
 F025,F221,F236,F097,F337,F308,F437,F208,F102,F412,F151,F290,F291,F009,F024,F381,F286  
 ,F424,F030,F510,F295,F059,F293,F035,F172,F231,F077,F229,F019,F384,F511,F387,F020,F02  
 3,F017,F015,F206,F319,F292,F187,F413,F018,F063,F260,F440,F188,F079,F032,F307,F383,F3  
 66

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F080

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-GAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F270  
GATCCTCTCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGGAAACGTTGCAGTC  
TGAAAAAAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F158,F074,F175,F502,F485,F409,F501,F182,F494,F180,F152,F090,F492,F186,F477,F165,F3  
92,F509,F183,F163,F408,F104,F181,F329,F450,F423,F499,F170,F460,F496,F153  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGTTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F312  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAG--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F254

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA

AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCGGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAA--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F466

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA

AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGC-----

-----

-----

>F108

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAAAAAATATTAAATAAACTA

AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAA--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F495  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAACAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F465  
-----  
-----  
-----ATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F469  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGGC  
ATGCC-----  
-----  
-----

>F439,F306,F075,F048,F278,F433,F341,F039,F514,F255,F211,F040,F365,F472,F372,F193,F0  
85,F139,F125,F328,F103,F405,F410,F411,F107,F289,F374,F070,F483,F301,F513,F244,F470,F  
099,F390,F323,F421,F428,F314,F134,F351,F088,F195,F234,F332,F218,F378,F082,F130,F076,  
F220,F399,F246,F310,F242,F335,F185,F207,F454,F093,F268,F397,F011,F100,F285,F087,F334  
,F241,F484,F416,F119,F389,F325,F275,F354,F252,F474,F363,F375,F404,F489,F401,F395,F32  
0,F446,F364,F122,F417,F095,F003,F044,F267,F504,F370,F131,F144,F233,F190,F336,F194,F1  
68,F281,F137,F403,F141,F197,F058,F497,F343,F263,F517,F171,F133,F138,F247,F036,F121,F

393,F415,F468,F219,F167,F091,F414,F166,F303,F362,F391,F506,F330,F481,F148,F014,F394,  
 F498,F005,F038,F081,F352,F235,F340,F124,F348,F515,F443,F402,F279,F123,F062,F198,F476  
 ,F467,F452,F505,F147,F508,F046,F406,F149,F136,F313,F346,F227,F145,F500,F222,F355,F12  
 9,F516,F200,F482,F162,F047,F078,F192,F132,F120,F143,F001,F140,F419,F418,F146,F471,F1  
 59,F400,F377,F142,F382,F357,F105,F431,F173,F315,F245,F086,F486,F305,F265,F069,F257,F  
 360,F094,F232,F376,F316,F243,F451,F512,F503,F154,F266,F507,F422,F458,F199,F191,F493,  
 F396,F037,F157,F398,F008,F073,F128,F326,F322,F213,F309,F169,F318,F434,F333,F430,F106  
 ,F473,F344,F029,F045,F298,F338,F369,F164,F435,F373,F071,F135,F196,F294,F274,F126,F09  
 2,F089,F345,F054,F359,F349,F116,F117,F179,F300,F250,F282,F445,F287,F184,F202,F096,F1  
 27,F150,F033,F350,F427,F438,F459,F041,F055,F240,F201,F217,F324,F302,F083,F371,F429,F  
 161,F464,F118,F043,F111,F204,F238,F420,F156,F177,Mphaseolina,Mtecta,Mvacinnii  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT  
 >F249,F258,F386,F269,F277,F109,F248,F251,F225,F256,F237,F388,F223,F228  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGGAAACGTTGCAGTC  
 TGAAAAAAAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT  
 >Bd01  
 GATCCTCCCACCCTTTG-TGTACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCC-  
 TCCGCGGCCGCCCCCTCCCCGGG  
 GGGTGGCCAGCGCCCGCCAGAGGACCATCAAACCTCCAGTCAGTAAACGATGCAGTC  
 TGAAAA---ACATTAAATAAACTA



AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCTTTGGTATTCCGAAGGGC  
ATGCCTGTTTCGAGCGTCATTACA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAACATACATCTCGCTTCGGAGCGCAGGGCGTCGCCC GCCGGACGAAC  
CTTCTGAACTTTTCTCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT

>F265

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>Mvacinnii  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGTCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGC-----  
-----  
-----  
-----  
----CAGACGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAGCACACTTTTCGCACTCCTGCACATGAGGCCGTCGCTCCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F319

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCAATGCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
CCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F514

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F363  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC

GCATGTGGATTTT----TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F181,F182,F158,F183,F392  
GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGTTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTCACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG

CCGTATCTCATTTACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F108  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAAAAAAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG



GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTGCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGGCCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F285,F316,F335  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC

GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT----TTTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F518  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCGCAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA

CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F449  
GATCCTCCCACCCTTTG-  
TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTCCGCGGCGCCCCCGATT  
TGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCACCATGGGTATGTCC  
ACCTGTCCG  
CGTGCCCCGCATCGCACCGCCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGTC-----  
-----  
AGACGCGTCTTTGCCCCCGGCCGCAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCGAAACACGCCC  
GCC-CCCGCGCC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG

TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGCAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAGCATACTTTTCGTGCTCCTGCATATGAGGCCGCCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGAGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATATGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGCGCCCCCTCTACA--  
--TTTTTGGCCACTCATGTT  
C--CTTTGCAAGCATAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
>F254  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCGCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCGGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCCGAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC

GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT-TTTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA-  
---TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTACTATAGGAAGCCGCTGAGCTCGG  
>F466  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTGTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGC-----  
-----  
-----CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG

TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTCA-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGGCATT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
 GCATGTGGATTTTT---TTTGGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
 TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F448  
 GATCCTCCCACCTTTG-  
 TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTTCGCGGGCCGCCCCCGATT  
 TGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTGAGCGTCATTTC  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
 GATGCGCCCCGCGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
 CGCCCCGCCAGCGGCACCGCTG  
 CCCCAGATCCGC-  
 CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCACCATGGGTATGTCC  
 ACCTGTCCG  
 CGTGCCCCGCATCGCACCGCCC-----  
 -----  
 GCGTCTTTGCCCCCGGCCGCAACAACGGGGTCAACACCGCT  
 GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
 ATGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GTTTGTGCGCGAAACACGCCC

GCC-CCCGCGCC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCACTAATTGCAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAGCATACTTTTCGTGCTCCTGCATATGAGGCCGCCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGAGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCATATGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGCGCCCCCTCTCACA--  
--TTTTTGGCCACTCATGTT  
C--CTTTGCAAGCATAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
>F030  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCGCCCGATTGTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCAATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA

TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT----TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F331,F432,F115  
GATCCTCCCACCCTTTG-  
TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTCCGCGGGCCGCCCCCGATTT  
TGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTCGAGCCTCCATCGTCGGTCGTCTCGCCACCATGGGTATGTCC  
ACCTGTCCG  
CGTGCCCCGCATCGCACCGCCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
----TTGTACGTACCGGCC



TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCGCGCA-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCACTAATTGGAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAGGCACACTTTTCGCGCTTCTGTACATGAGGCCGTCGCTTCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT--  
TTGGTGGGGTTGATTCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCTTACGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCTCTCCCACA--  
--TTCTTGGCCACTCATGTT  
CTTCCCAGCGAGCACAAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
>F443  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCGGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG

CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
ACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTAT--CGTTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>Mtecta  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCCCCCAGCGGCACCGCTG

CCCCAAGATCCGCCCAGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGTGGA----  
-----TTGTAAGTACCGGCC  
TGCGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGCCAAAACACGCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCACTACACCGAGGGTGCGGAGAAGGTAAGCACACTTTTCGCACTCCTGCACATG  
AGGCCGTCGCTCCACCAGAATTTG  
CCGTATCGCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGGCCGATGCGCGAGCCCCCTCCACA--  
--TTTT-TGCCACTCATGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F434  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG

CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTAAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGC GTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTC AAAACACGCC  
GCC-CCCCGCCC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACA ACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACA ACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGGCCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F213  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAA ACTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG

TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGC GCCTGCTTGT TTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTACTATAGGAAGCCGCTGAGCTCGG  
>F114,F264  
GATCCTCCCACCCTTTG-  
TATACCTACCTCTGTTGCTTTGGCGGGGCCGCGGTCTTTCCGCGGGCGCCCCCGATTT  
TGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG  
 TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
 CGCCCGCCCAAGCGGCACCGCTG  
 CCCCAGATCCGC-  
 CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCACCATGGGTATGTCC  
 ACCTGTCCG  
 CGTGCCCCGCATCGCACCCGCCCCTGACGTACACCGCAGTATCATGATTGGT--  
 ATGGGCCAGAATTGTACGTACCGGCC  
 TCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACACCGCT  
 GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
 ATGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCCC  
 GCC-CCCGCGCA-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGGAAAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCTCGTTGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTC-----  
 GAGAAGGTAGGCACACTTTTCGCGCTTCTGTACATGAGGCCGTCGCTTCACCAGAAT  
 TTG  
 CCGTATCTCATTGAATGAGGGGCATTT--  
 TTGGTGGGGTTGATTCGCGCTAAGCCGCATTACGCCTTCGGCAAAATCTCC  
 GCTTACGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCTCTCCCACA--  
 --TTCTTGGCCACTCATGTT  
 CTTCCCAGCGAGCACAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
 >F016,F049,F236,F018,F407,F383,F015,F203,F283,F013,F072,F262,F097,F025,F384,F216,F0  
 59  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCTCAA

GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTCGCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGATGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT----TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>Meuphorbiicola  
GATCCTCCCACCCTTTG-  
TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTTCCGCGGGCCGCCCCCGATT  
TGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
 GATGCGCCCCGCGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
 CGCCCCGCCCAGCAGCACCGCTG  
 CCCCAGATCCGC-  
 CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCACCATGGGTATGTCC  
 ACCTGTCCG  
 CGTGCCCCGCATCGCACCGCCCCGCTGACGTACACCGCAGTATCATGATTGGT-----  
 -----  
 -----  
 CGCGTCTTTGCCCCCGGCCGCAACAACGGGGTCAACACCGCT  
 GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
 ATGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCCC  
 GCC-CCCGCGCC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGAGGAACATCAATGACTAAACAGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGCAAAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTTGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACTGGGCCAAG  
 GGTC-----  
 GAGAAGGTAAGCATACTTTTCGTGCTCCTGCATATGAGGCCGCCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGAATGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGAGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
 GCATATGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGCGCCCCCTCTCACA--  
 --TTTTGGCCACTCATGTT  
 C--CTTTGCAAGCATAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
 >F042  
 GATCCTCCCACCCTTTG-  
 TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTCCGCGGCCGCCCCCGATT  
 TGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA



AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCACCATGGGTATGTCC  
ACCTGTCCG  
CGTGCCCCGCATCGCACCGCCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGTC-----  
-----  
AGACGCGTCTTTGCCCCCGGCCGCAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCGCGCC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGCAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAGCATACTTTTCGTGCTCCTGCACATGAGGCCGCCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGATCCGAGCTAAGCCGCATTACGCCTTCGGCAAAATCTCC  
GCATATGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGCGCCCCCTCCCACA--  
--TTTTTGGCCACTCATGTT  
C--CTTTGCGAGCATAAGCTAACTGCGGCTGTTACAGGAAGCCGCTGAGCTCGG  
>Mphaseolina  
GATCTCCACCCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCGCCCGATTGTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTAAGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTACAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F333,F334

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
ACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F385

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAAAAAAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGTCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCCTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F178  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGG  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCTTGTTGCTGCTCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAATCTCC

GCATGTGGATTTTT----TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F436  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCAATGCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
CCCCCCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG

CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F188  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCTTGTTTCGCTGCCTTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-  
CCCCGCGCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG

GGTCA-----

GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG

CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC

GCATGTGGATTTTT---TTTTGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
TTTT-TGCCACTCGTGT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F100

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCGGATTTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA

AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA

GCGTAGTAGAA--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG

TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG

CCCCAAGATCCGCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG

CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC

TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT

GCAATGCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA

TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACAGCCC

GCC-CCCCGCGC-

CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG

TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG

CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA

CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGCACCATGGATGCCGTCC



GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCGATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT----TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F370,F416,F085,F093,F048  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA

CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTAT--CGTTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F107  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCGGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
ACAATGGCACCTCGGACCTCCAG

CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCAATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F106  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGG  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG

TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F110  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTCCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTACCCCCCGGCCG  
CAACAACGGGGTCAACGCCGT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC

GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTAGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCGCATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTACCTCATTGATTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTTTTTTTTTCGGCCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA-  
---TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F185  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCAGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT-GA-----  
-----TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCGATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA

TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGGCCAAAACACGCCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAGCACACTTTTCGCACTCCTGCACATGAGGCCGTCGCTCCACCAGAAT  
TTG  
CCGTATCGCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGGCCGATGCGCGAGCCCCCTCCCACA--  
--TTTT-TGCCACTCATGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F485,F494,F496  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGTTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT

GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGCGCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCGATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F003,F118,F156  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC

TGC GCCTGCTTGT TTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TCGGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTGTCGCAAAACACGCCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GTCA-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTAT--CGTTGACTGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
 --TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F307,F260,F424,F224,F214,F155,F353,F271,F021,F381,F205,F221,F297,F077,F209  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAATAAACTA  
 AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
 CGCCCGCCCCAGCGGCACCGCTG  
 CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG



CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT---TTTGTACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F410,F341,F320,F120,F318,F330,F062  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG

CCCCAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
 ---TTGTACGTACCGGCC  
 TGGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TCGGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTGTCGCAAAACACGCCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
 ACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTC-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
 -TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F361,F342,F379  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG

TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGC GCCTGCTTGT TTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCAATGCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
CCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACA ACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACA ACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F112  
GATCCTCCCACCCTTTG-  
TATACCTACCTCTGTTGCTTTGGCGGGGCCGCGGTCTTTCCGCGGGCGCCCCCGATTT  
TGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAA ACTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCACCATGGGTATGTCC  
ACCTGTCCG  
CGTGCCCCGCATCGCACCCGCCCCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCGCGCA-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCACTAATTGGAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCTCGTTGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAGGCACACTTTTCGCGCTTCTGTACATGAGGCCGTCGCTTCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT--  
TTGGTGGGGTTGATTCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCTTACGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCTCTCCACA--  
--TTCTTGGCCACTCATGTT  
CTTCCCAGCGAGCACAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
>F390,F081,F430,F389,F406,F429  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA

GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTTCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTTACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F386,F225,F256,F237  
GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGGAAACGTTGCAGTC  
TGAAAAAAAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
 CGCCCGCCCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
 ATGGGCCAGAATTGTACGTACCGGCC  
 TCGCCTGCTTGTTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TCGGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GTCA-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
 ATTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTTT---TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
 TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F472,F393,F394,F512,F468,F470,F473,F397,F471,F078,F159,F513,F484,F184,F150,F476,F5  
 05,F391,F493,F483,F489,F452,F451,F149  
 GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTTACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F201  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA

AACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
 CGCCCGCCCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
 ATGGGCCAGAAATTGTACGTACCGGCC  
 TGCGCCTGCTTGTTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TCGCTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTGTGTTGCCAAAACACGCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACTATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTC-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTTACTGAGGGGGCATT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
 -TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >Mpseudophaseolina  
 GATCCTCCCACCCTTTG-  
 TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTCCGCGGGCCGCCCCCGATT  
 TGGG



GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCACCATGGGTATGTCC  
ACCTGTCCG  
CGTGCCCCGCATCGCACCGCCCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTAAGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTCACCTCCAGACCGGCCA  
ATGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCGCGCA-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGGAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAGGCACACTTTTCGCGCTTCTGTACATGAGGCCGTCGCTTCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT--  
TTGGTGGGGTTGATTCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCTTACGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCTCTCCACA--  
--TTCTTGGCCACTCATGTT  
CTTCCCAGCGAGCACAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
>F105

GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----GAGAAGGTAAAC--  
ACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAATTTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F148

GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTCGCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
ACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAAAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F511,F151,F079

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCTTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAACCCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAATCTCC  
GCATGTGGATTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F293,F229,F308,F023,F022,F019,F172,F020,F290,F291,F189,F286,F187,F174,F026,F426,F0  
56,F208,F231,F366,F368,F063,F017,F057,F292,F035,F284,F296,F295,F230  
GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTGTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTACGCCTTCGGCAAAATCTCC

GCATGTGGATTTT----TTTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
 TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >Bd02  
 GATCCTCCCACCCTTTG-TGTACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCC-  
 TCCGCGGCCGCGCCCCCTCCCCGGG  
 GGGTGGCCAGCGCCCGCCAGAGGACCATCAAACCTCCAGTCAGTAAACGATGCAGTC  
 TGAAAA-A--CATTTAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCTTTGGTATTCCGAAGGGC  
 ATGCCTGTTGAGCGTCATTACA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAACATACATCTCGCTTCGGAGCGCAGGGCGTCGCCCCCGCGGACGAAC  
 CTTCTGAACTTTTCTCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATC-----  
 GCCGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGTGATAGCGCCTTTCGGC-  
 CGCGCCAGCTCGTTACTGCTGCACTGCCG  
 CAATGAGAATGGC-  
 TTGTTGACTTTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCACCATGGGTATGTCCA  
 CCTGTCCG  
 CGTGCCCCGCATCGCACCGCCCGCTGACGTACACCGCAGTATCATGATTGGT--AT-----  
 -----GGGCCA-----  
 -----GACGCGTC-  
 CTGCCCCCGCGCCGCAACAACAGCGCCAGCATCGCA  
 GCGCACCTTTGCTAACATCGCCTTTTCGTGTCCATAGGTTACCTCCAGACCGGCCA  
 ATGCGTAAGTCTCCTCCCGACCA  
 TCTTTTTCGCTGCATCGCGCTGACTTTGACTAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTGCAAAACACCC--  
 GCT-CCCGCGCC-  
 CCCGCTAACGCGAATCGACACCACAGGCAGACCATCTCCGGCGAGCACGGCCTGGA  
 CGGCTCTGGCG  
 TGTAAGTCTGCATCATTCTCAGCGTGGGAGAACATCAATGACTAAACTGTAGCAGCT  
 ACAATGGCACCTCGGACCTTCAG  
 CTCGAGCGCATGAACGTCTATTTCAACGAGGTACTCTCTCACTAATTAGACAAACAC  
 GTAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCGTCCAACAACAAGTACGTTCCCTCGTGCCGTCCTCGTCGACCTCGA  
 GCCCGGCACGATGGATGCCGTCC  
 GCGCCGGCCCCCTTCGGCCAGCTTTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTCA-----  
 GAGAAGGTAAGCACACATTTTCTGTGCCTGCACGTGTGCTGGGTTCTGCGCCGAAT  
 TTG

CCTTATCACTCTGGTGAGGGGCAATTT-  
CTTGGTGGGGCTGGCCCGCGCTAAGCCTCGTTTGGTCTTCGGCAAAATCTCC  
GCATCTGGATTTT-----TTGTGACCGGCGTGCGACCGACGCGAACACCCCTC-  
ACCAACGCTTCC-AGCCACTCACGTT  
CGTCTATGCGACCATATGCTAACCACCGCCACAACAGGAAGCCGCTGAGCTCGG  
>F163,F165,F329,F104,F180  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGGCCGCCCCCGATTTTGGG  
GGTTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG

GGTCA-----

GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTGCTGCACCAGAAT  
TTG

CCGTATCTCATTTACTGAGGGGCATTT-

TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC

GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---

-TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F336,F094,F343

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-

TCCGCGGGCCGCCCCCGATTTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT

GAAAA-AAATATTAAATAAACTA

AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC

GATAAGTAATGTGAATTGCAGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC

ATGCCTGTTGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC

TCGGCGGTGGCGTCTTGCTCAA

GCGTAGTAGAA--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGGACGAACCTTCTGAACTTTTC

TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----

CGCTGTCTTCCG

TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA

CGCCCGCCCCAGCGGCACCGCTG

CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCTGCCA

CCATGGGTATGTCCACCTGTCCG

CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--

ATGGGCCAGAATTGTACGTACCGGCC

TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG

CAACAACGGGGTCAACGCCGCT

GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA

TGCGTAAGTCTCGCCGCTGCCA

TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT

GGTTTGTTGCCAAAACACGCCC

GCC-CCCCGCGC-

CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA

CGGCTCCGGCG

TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC

TACAATGGCACCTCGGACCTCCAG

CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC

GCAAAGTATGGCAATCTTCTGAA

CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA

GCCCGGCACCATGGATGCCGTCC



GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F325,F096,F275,F268,F227  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA

CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTC-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCGCATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTACCTCATTGACTGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTTT---TTTTCGGCCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
 -TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F340,F123,F323,F418,F411,F130,F377,F428,F310,F324  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCGGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
 GATGCGCCCCGCGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
 CGCCCGCCCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
 ---TTGTACGTACCGGCC  
 TGGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTGTCGCAAAACACGCCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG

CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F414,F095  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGGATTTTGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGG  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG

TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTC-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAAAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGGCATT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
 GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
 --TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F171,F169,F166,F167,F315  
 GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
 GATGCGCCCCGCGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
 CGCCCCGCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
 ---TTGTACGTACCGGCC  
 TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
 CAACAACGGGGTCAACGCCGT  
 GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCC

GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCAATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F129,F346,F445  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTGTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA

TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
ACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCAATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F111  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT

GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTACCTCATTGACTGAGGGGCATTT--  
TTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGGCCGGCGTGCGACCGATCCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F220,F008,F071,F091,F177  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC

TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGC GTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTACGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTACTATAGGAAGCCGCTGAGCTCGG  
>F270  
GATCCTCTCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGGAAACGTTGCAGTC  
TGAAAAAAAATATTAATAAACTA  
AACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG



CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACACCGCT  
GCGATCCTTTGCTAACCTTGCCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F269,F258,F388,F251,F277,F223,F249,F109,F248,F228  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGGAAACGTTGCAGTC  
TGAAAAAAAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG

CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGC GTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTC AAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACA ACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACA ACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT----TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F332,F378  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAA ACTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG

TAAGTCTCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGC GCCTGCTTGT TTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACA ACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACA ACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCTGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F469  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAA ACTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCC-----  
-----  
-----  
-----GCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG

CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTCAAAACACGCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCGCAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTGTACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
TTTT-TGCCACTCGTGT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F113  
GATCCTCCCACCCTTTG-  
TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTCCGCGGCGCCCCCGATT  
TGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG

TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCACCATGGGTATGTCC  
ACCTGTCCG  
CGTGCCCCGCATCGCACCGCCCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
----TTGTACGTACCGGCC  
TGGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTTGCCAAAACACGCC  
GCC-CCCGCGCA-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGGAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCAGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAGGCACACTTTTCGCGCTTCTGTACATGAGGCCGTCGCTTCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT--  
TTGGTGGGGTTGATTCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCTTACGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCTCTCCCACA--  
--TTCTTGGCCACTCATGTT  
CTTCCCAGCGAGCACAAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
>F152,F450,F153,F074,F423,F477,F492,F175,F408  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGTTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACAGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTGCTGCTGCTCGTACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTTACTGAGGGGCAATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F157  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG  
TAAGTCTTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT-GA-----  
-----TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCAGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCGATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAGCACACTTTTCGCACTCCTGCATATGAGGCCGTCGCTCCACCAGAAT  
TTG  
CCGTATCGCATTGAATGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT----TTTTCGACCGGCGTGCGGCCGATGCGCGAGCCCCTCCCACA---  
-TTTT-TGCCACTCATGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F435  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGC GCCTGCTTGT TTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTGCTGCAACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>Bd01  
GATCCTCCCACCCTTTG-TGTACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCC-  
TCCGCGGCCGCCCCCTCCCCGGGG  
GGGTGGCCAGCGCCCGCCAGAGGACCATCAAACCTCCAGTCAGTAAACGATGCAGTC  
TGAAAA-A--CATTTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA



TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCTTTGGTATTCCGAAGGGC  
ATGCCTGTTCGAGCGTCATTACA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAACATACATCTCGCTTCGGAGCGCAGGGCGTCGCCCCGCCGGACGAAC  
CTTCTGAACTTTTCTCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCCGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGTGATAGCGCCTTTCGGC-  
CGCGCCAGCTCGTTACTGCTGCACTGCCG  
CAATGAGAATGGC-  
TTGTTGACTTTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCACCATGGGTATGTCCA  
CCTGTCCG  
CGTGCCCCGCATCGCACCCGCCGCTGACGTACACCGCAGTATCATGATTGGT--CG----  
-----CG-----  
-----TC-  
CTGCCCCCGGGCCGCAACAACAGCGCCAGCATCGCA  
GCGCACCTTTGCTAACATCGCCTTTTCGTGTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAGTCTCCTCCCGACCA  
TCTTTTTCGCTGCATCGCGCTGACTTTGACTAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACCC--  
GCT-CCCGCGCC-  
CCCGCTAACGCGAATCGACACCACAGGCAGACCATCTCCGGCGAGCACGGCCTGGA  
CGGCTCTGGCG  
TGTAAGTCTGCATCATTCTCAGCGTGGGAGAACATCAATGACTAAACTGTAGCAGCT  
ACAATGGCACCTCGGACCTTCAG  
CTCGAGCGCATGAACGTCTATTTCAACGAGGTACTCTCTCACTAATTAGACAAACAC  
GTAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCGTCCAACAACAAGTACGTTCTCGTGCCGTCCTCGTCGACCTCGA  
GCCCCGGCACGATGGATGCCGTCC  
GCGCCGGCCCCCTTCGGCCAGCTTTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAGCACACATTTTCTGTGCCTGCACGTGTGCTGGGTTCTGCGCCGAAT  
TTG  
CCTTATCACTCTGGTGAGGGGCAATTT-  
CTTGGTGGGGCTGGCCCGCGCTAAGCCTCGTTTGGTCTTCGGCAAAATCTCC  
GCATCTGGATTTT-----TTGTGACCGGCGTGCGACCGACGCGAACACCCCTC-  
ACCAACGCTTCC-AGCCACTCACGTT  
CGTCTATGCGACCATATGCTAACCACCGCCACAACAGGAAGCCGCTGAGCTCGG  
>F170,F090,F501,F186,F409,F460  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGTTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA

AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTTACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F415,F076  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTGTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTTGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCGCATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTACCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGGCCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F431,F083,F082,F398,F161,F396,F400,F073,F399

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTTCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTTACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F433

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGTCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTAAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTAAGTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT---TTTTCGGCCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F412,F280,F206,F273,F413,F210,F337,F226,F387,F024,F212,F259,F367,F009,F060,F253  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTGTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGG  
ATGCCTGTTTCGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCACGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTGTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGATGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC

GCATGTGGATTTT----TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F401  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGCGACTATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG

CCGTATCTCATTTACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F372,F374,F376,F373,F371,F362  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCAGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCGATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-  
CCCCGCGCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG



GGTCA-----  
GAGAAGGTAAGCACACTTTTCGCACTCCTGCACATGAGGCCGTCGCTCCACCAGAAT  
TTG  
CCGTATCGCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGGCCGATGCGCGAGCCCCTCCCACA--  
--TTTT-TGCCACTCATGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F515,F467,F516,F232  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCCTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC

GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTCA-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
 GCATGTGGATTTTT---TTTGGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
 TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F465

-----  
 -----  
 -----ATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGG  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG  
 TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
 CGCCCGCCCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
 ATGGGCCAGAATTGTACGTACCGGCC  
 TGGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TGGCTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG

GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTGCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCGATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F495  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAACAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCCGAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCCTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC

GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F439  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA

CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTAT--CGTTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT-TTTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA-  
---TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F239,F510,F102,F176  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCGGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCTTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-  
CCCCGCGCCCCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCA  
CGGCTTGGACGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG

CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTCA-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
 TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F040,F122,F252,F217,F207,F088,F005,F036,F240,F191,F303,F359,F245,F274,F137,F204,F5  
 06,F117,F243,F382,F365,F103,F142,F033,F222,F287,F116,F011,F211,F197,F302,F305,F046,F  
 241,F247,F351,F134,F144,F131,F140,F298,F369,F145,F313,F306,F127,F300,F282,F246,F121,  
 F202,F043,F143,F089,F219,F354,F375,F263,F497,F196,F193,F350,F503,F124,F119,F001,F421  
 ,F420,F322,F242,F070,F054,F438,F266,F069,F092,F218,F234,F309,F352,F138  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCGCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGGC  
 ATGCCTGTTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
 GATGCGCCTCGCGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
 CGCCCCGCCCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
 ---TTGTACGTACCGGCC  
 TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCAATGCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TCGGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTGTTGTTGCCAAAACACGCCC

GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGCCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCAATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTT---TTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F437,F440  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCAATGCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA

TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGGCCAAAACACGCCC  
CCCCCCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT----TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F311,F347  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT



GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT----TTTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F032  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTC  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCTCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC

TGCGCCTGCTTGT TTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCAATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTCACAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATTGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCGATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTT----TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F446  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTCGAGCCTCCATCGTCGGTCGTCTCGCCACCATGGGTATGTCC  
ACCTGTCCG

CGTGCCCCGCATCGCACCCGCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCGCGCA-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACAGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGGAAAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCCGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAGGCACACTTTTCGCGCTTCTGTACATGAGGCCGTCGCTTCACCAGAAT  
TTG  
CCGTATCTCATTGAATGAGGGGCATTT--  
TTGGTGGGGTTGATTCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCTTACGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCTCTCCACA--  
--TTCTTGGCCACTCATGTT  
CTTCCCAGCGAGCACAAAGCTAACTGCGGCTATTACAGGAAGCCGCTGAGCTCGG  
>F380,F444  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCCGCCAGCGGCACCGCTG

CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGTCGCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
ATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT----TTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F084  
GATCCTCCCACCCTTTGTTATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCTTTCC  
GCGGCCGCCCCCGATTGTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG

TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGC-  
CGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCACCATGGGTATGTCC  
ACCTGTCCG  
CGTGCCCCGCATCGCACCGCCC-----

-----  
GCGTCTTTGCCCCCGGCCGCAACAACGGGGTCAACACCGCT  
GCGATCATTTGCTAACCTCGCCTTTTCACCTCCATAGGTTACCTCCAGACCGGCCA  
ATGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTGGCCAAAACACGCCC  
GCC-CCCGCGCC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCCTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAGCATCAATGACTAAACAGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTACTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTTGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----

GAGAAGGTAAGCATACTTTTCGCGCTCCTGCACATGAGGCCGCCGCTGCACCAGAA  
TTTG  
CCGTATCTCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGATCCGCGCTAAGCCGCATTCTGCCTTCGGCAAAATCTCC  
GCATATGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGCGCCCCCTCCCACA--  
--TTTTTGGCCACTCATGTT  
C--CTTTGCGAGCATAAGCTAACTGCGGCTATTACAGGAAGCTGCTGAGCTCGG  
>F312

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTTAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAG--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG  
 TAAGTCTCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
 CGCCCGCCCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
 ATGGGCCAGAATTGTACGTACCGGCC  
 TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCC  
 GCC-  
 CCCC GCGCCCCCGCTAATGCGAATCAACACCGCAGGCAGACCATCTCTGGCGAGCA  
 CGGCTTGACGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GTCA-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGCATTTTTTTGGTGGGGTTGGTCCGCGCTAAGCCGC  
 ATTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTTT---TTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
 TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F326,F278,F360,F037,F349,F200,F058,F132,F338,F044,F041,F135,F128,F250,F199,F173,F1  
 92,F500,F190,F075,F194,F139,F039,F055,F133,F146,F038,F281,F195,F141,F029,F279,F086,F  
 235,F198,F126,F267,F233,F045,F136,F348,F147,F301,F344,F014,F238,F047,F364,F087,F244,  
 F357,F255,F355  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA  
 AAACCTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGC GCCTGCTTGT TTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCAATGCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GTGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGCCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTTGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F125  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
 CGCTGTCTTCCG  
 TAAGTCTCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
 CGCCCGCCCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCCGGCTCACTCTGTGCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
 ATGGGCCAGAATTGTACGTACCGGCC  
 TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
 ACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GTCA-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAATCTCC  
 GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
 --TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
 >F502,F509,F499  
 GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
 TCCGCGGCCGCCCCCGATTTTGGG  
 GGTTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-AAATATTAAATAAACTA



AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCCGCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTAAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F080  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
 GAAAA-GAATATTAATAAACTA  
 AAACCTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
 GATAAGTAATGTGAATTGCAGAA  
 TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
 ATGCCTGTTCGAGCGTCATTTCA  
 ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
 TCGGCGGTGGCGTCTTGCCTCAA  
 GCGTAGTAGAA--  
 TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
 TCAAGGTGA  
 CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
 GATGCGCCCCGCGCTGTCTTCCG  
 TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
 CGCCCCGCCAGCGGCACCGCTG  
 CCCCAGATCCGCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCGCCA  
 CCATGGGTATGTCCACCTGTCCG  
 CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
 ---TTGTACGTACCGGCC  
 TCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
 CAACAACGGGGTCAACGCCGCT  
 GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
 TGCGTAAGTCTCGCCGCTGCCA  
 TCGTGTTGCTGCCTCGCGCTGACTTTGCCCAGGGTAACCAAATCGGTGCTGCCTTCT  
 GGTTTGTTGCCAAAACACGCCC  
 GCC-CCCCGCGC-  
 CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
 CGGCTCCGGCG  
 TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
 TACAATGGCACCTCGGACCTCCAG  
 CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
 GCAAAGTATGGCAATCTTCTGAA  
 CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
 GCCCGGCACCATGGATGCCGTCC  
 GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACCTTCGTCTTCGGTCAGTCCG  
 GTGCCGGTAACAACCTGGGCCAAG  
 GGTC-----  
 GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
 TTG  
 CCGTATCTCATTGACTGAGGGGCATTT-  
 TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
 GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
 --TTTT-TGCCACTCGTGTT  
 CTTCTCTGCGAGCATAAGCTAACTGCGGCTACTATAGGAAGCCGCTGAGCTCGG  
 >F328,F345,F422,F314,F427

GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGGCATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F099,F482

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGG  
ATGCCTGTTTCGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCTGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTTCGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACAGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCCTGTGGAGGAACATCAATGACTAAACGGCAGCAGCT  
ACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTGAGCCTTCGGCAAATCTCC  
GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F404,F458,F507,F403,F179,F162,F405,F474,F395,F481,F294,F402,F154,F417,F459,F454,F5  
08,F168,F504,F498,F464,F486,F517,F164  
GATCCTCCCACCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTGTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG  
CCGTATCTCATTTACTGAGGGGCAATT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTACGCCTTCGGCAAAATCTCC

GCATGTGGATTTTT---TTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA---  
-TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F257  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGACGAACCTTCTGAACTTTTC  
TCAAGGTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----  
CGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
TGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCCGGCTCACTCTGTCTGCAGCCTCCATCGTCGGTCGTCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--  
ATGGGCCAGAATTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTACCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCCTTTTCGCCTCCATAGGTTCACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTATTGCTGCCTCGCGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTAGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACCTGGGCCAAG  
GGTCA-----  
GAGAAGGTAAACACACTTTTCGCACTCCTGCGCATGAGGCTGTCGCTGCACCAGAAT  
TTG

CCGTACCTCATTGACTGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT--TTTTCGGCCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--  
--TTTT-TGCCACTCGTGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG  
>F419  
GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-  
TCCGCGGCCGCCCCCGATTTTGGG  
GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT  
GAAAA-AAATATTAATAAACTA  
AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC  
GATAAGTAATGTGAATTGCAGAA  
TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCCTTGGTATTCCGGGGGGC  
ATGCCTGTTGAGCGTCATTTCA  
ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC  
TCGGCGGTGGCGTCTTGCTCAA  
GCGTAGTAGAA--  
TACACCTCGCTTCGGAGCGTAAGGCGTCGCCCCGCCGGACGAACCTTCTGAACTTTTC  
TCAAGGTTGA  
CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATATCCGGTTTCGCCGGTGAT  
GATGCGCCCCGCGCTGTCTTCCG  
TAAGTCTCCCCCGCCTCCTACTGAGCTGGCCGCGATGGCGCCTGTCTGCCACGACA  
CGCCCGCCCCAGCGGCACCGCTG  
CCCCAAGATCCGCCCGGCTCACTCTGTGCGAGCCTCCATCGTCGGTCGTCCTCGCCA  
CCATGGGTATGTCCACCTGTCCG  
CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--A-----  
---TTGTACGTACCGGCC  
TGCGCCTGCTTGTTTTTAACGCGTCCCCTGACTGCCCCGCGTCTTTGCCCCCGGCCG  
CAACAACGGGGTCAACGCCGCT  
GCGATCCTTTGCTAACCTTGCTTTTCGCCTCCATAGGTTACCTCCAGACCGGCCAA  
TGCGTAAGTCTCGCCGCTGCCA  
TCGTGTTGCTGCCTCACGCTGACTTTGCCAGGGTAACCAAATCGGTGCTGCCTTCT  
GGTTTGTTGCCAAAACACGCC  
GCC-CCCCGCGC-  
CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA  
CGGCTCCGGCG  
TGTAAGTCTGCGCCATTCCCACCGTGAGGAACATCAATGACTAAACGGCAGCAGC  
TACAAATGGCACCTCGGACCTCCAG  
CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC  
GCAAAGTATGGCAATCTTCTGAA  
CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA  
GCCCGGCACCATGGATGCCGTCC  
GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG  
GTGCCGGTAACAACTGGGCCAAG

GGTCA-----

GAGAAGGTAAACACACTTTTCGCACTCCTGCACATGAGGCTGTCGCTGCACCAGAAT  
TTG

CCGTATCTCATTGACTGAGGGGCATTT-

TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTTCAGCCTTCGGCAAAATCTCC

GCATGTGGATTTTT--TTTTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA--

--TTTT-TGCCACTCGTGTT

CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F289

GATCCTCCCACCCTTTG-TATACCTACCTCTGTTGCTTTGGCGGGCCGCGGTCT-

TCCGCGGGCCGCCCCCGATTTTGGG

GGGTGGCTAGTGCCCGCCAGAGGACTATCAAACCTCCAGTCAGTAAACGTTGCAGTCT

GAAAA-AAATATTAATAAACTA

AAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGC

GATAAGTAATGTGAATTGCAGAA

TTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCGGGGGGC

ATGCCTGTTTCGAGCGTCATTTCA

ACCCTCAAGCTCTGCTTGGTATTGGGCACCGTCCTTTGCGGGCGCGCCTCAAAGACC

TCGGCGGTGGCGTCTTGCCTCAA

GCGTAGTAGAA--

TACACCTCGCTTCGGAGCGTAAGGCGTCGCCC GCCGGACGAACCTTCTGAACTTTTC

TCAAGGTTGA

CCTCGGATCAGGTAGGGATACCCGCTGAACTTAAGCATAT-----

CGCTGTCTTCCG

TAAGTCTCCCCCGCCTCCTGCTGAGCTGGCCGCGATGGCGCCTGTCTGCCCACGACA

CGCCCGCCCCAGCGGCACCGCTG

CCCCAAGATCCGCCCGGCTCACTCTGTTCGCAGCCTCCATCGTCGGTCGTCCTCGCCA

CCATGGGTATGTCCACCTGTCCG

CGTGCCCCGCATCGCACCGTCCGCTGACGTACACCGCAGTATCATGATTGGT--

ATGGGCC-----

-----  
-----  
-----

-----AGAC-

CCCGCTAATGCGAATCAACACCACAGGCAGACCATCTCTGGCGAGCACGGCTTGGA

CGGCTCCGGCG

TGTAAGTCTGCGCCATTCCCACCGTGGAGGAACATCAATGACTAAACGGCAGCAGC

TACAAATGGCACCTCGGACCTCCAG

CTCGAGCGCATGAACGTCTACTTCAACGAGGTACTCTCTCGCTAATTGGACAAACCC

GCAAAGTATGGCAATCTTCTGAA

CGCGCAGCAGGCTTCCAACAACAAGTACGTTCCCTCGTGCTGTCCTCGTCGACCTCGA

GCCCGGCACCATGGATGCCGTCC

GCGCCGGCCCCCTTTGGCCAGCTCTTCCGCCCCGACAACTTCGTCTTCGGTCAGTCCG

GTGCCGGTAACAACTGGGCCAAG



GGTCA-----  
GAGAAGGTAAGCACACTTTTCGCACTCCTGCACATGAGGCCGTCGCTCCACCAGAAT  
TTG  
-CGTATCTCATTGAATGAGGGGCATTT-  
TTTGGTGGGGTTGGTCCGCGCTAAGCCGCATTCAGCCTTCGGCAAAATCTCC  
GCATGTGGATTTTT---TTTCGACCGGCGTGCGACCGATGCGCGAGCCCCTC-CACA----  
TTCT-TGCCACTCATGTT  
CTTCTCTGCGAGCATAAGCTAACTGCGGCTATTATAGGAAGCCGCTGAGCTCGG

>F153,F509,F454,F493,F494,F329,F450,F423,F474,F499,F398,F406,F179,F504,F152,F082,F501,F396,F512,F400,F081,F409,F090,F389,F473,F083,F502,F175,F073,F460,F496,F390,F408,F431,F430,F186,F074,F476,F492,F477,F170,F485,F161,F399,F429  
GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACATTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCATTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F148,F419,F445,F330,F343,F428,F118,F410,F332,F105,F310,F328,F099,F411,F340,F048,F334,F439,F370,F120,F107,F324,F333,F341,F443,F125  
GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCCTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTAGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTCGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CAGAGCTCTTGCTCAAGTTTCGCTACC  
>F336,F315,F130  
GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCCTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F433

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGG  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGGCAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F380,F285,F311,F347

GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F377,F093,F418,F085,F427,F129,F422,F416,F156,F062

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTAGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGAGAGATATCATTTAGCCCCGACACCATGGAAATCTCAATTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CAGAGCTCTTGCTCAAGTTTCGCTACC

>F318,F320,F346

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA

CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AGATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATA  
CCAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAGTTTGAGCACTAGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F312,F518,F469

GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F123,F003

GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCATCTA  
TCGAACATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCGTCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F106,F094

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA

GTCAAGAACCATAAAGACAATTTGAGCACTAGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F314,F323

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATA  
CCAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAGTTTGAGCACTAGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F084

GCCACATTGATGCCTTGAAGAGAGATTTCTGCATCGGGAACGGAAAACCTCCCATCTA  
TCGAATATATGGCTGACCAAGAG  
CCAACTGCAAATTACCCCGAGAATACCACTATCATCGAGACTGAATCCGACATCATG  
CTCCATTTACCCGGCTCCACGGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGTTCCATCATG  
GCGAGCTGCACTTTTTTCCTCTG  
AAACCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGCCAGCGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGCCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATATCAGTTGTCTTGACGCATCA  
TAAACAGCGGCGACGCAAGCCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F157

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACCGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCAT  
GCTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTAAGCACTAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TAAACAGCGGCGACGCAACCGG

CGGAGTTCTTGCTCAAGTTTCGCTACC

>F448,F449,F042

GCCACATTGATGCCTTGAAGAGAGATTTCTGCATCGGGAACGGAAAACCTCCCATCTA  
TCGAATATATGGCTGACCAAGAG  
CCAACCTGCAAATTACCCCGAGAATACCACTATCATCGAGACTGAATCCGACATCATG  
CTCCATTTACCCGGCTCCACGGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGTTCCATCATG  
GCGAGCTGCACTTTTTTCCTCTG  
AAACCATTCAAGAGGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGCCAGCGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGCCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATATCAGTTATCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F414

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTACTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F257,F363

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F345

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA

TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAACCCCTGGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTAGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTCGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CAGAGCTCTTGCTCAAGTTTCGCTACC  
>F071,F091,F008,F213,F220  
GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATTCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCATTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F275,F096,F268  
GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCATTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F362,F185,F372,F371,F373,F374,F376  
GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACCGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCAT  
GCTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG

AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
ATCAAGAACCATAAAGACAATTTAAGCACTAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F366,F467,F295,F424,F271,F020,F022,F335,F221,F290,F155,F172,F023,F232,F174,F466,F0  
57,F230,F187,F026,F260,F208,F056,F209,F214,F307,F035,F224,F368,F296,F297,F019,F286,F  
189,F426,F516,F292,F353,F021,F515,F205,F291,F229,F077,F063,F381,F284,F316,F017,F293,  
F231,F308  
GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATAACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCATCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F437,F440,F436  
GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGGTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATAACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCATCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F482  
GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATAACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG



AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTATAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTCGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CAGAGCTCTTGCTCAAGTTTCGCTACC

>F095

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTACCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTAGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTCGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F322,F306,F360,F147,F070,F041,F132,F354,F365,F197,F238,F086,F243,F136,F375,F014,F4  
97,F211,F207,F506,F139,F241,F128,F349,F087,F069,F038,F266,F145,F011,F351,F369,F359,F  
348,F143,F244,F240,F195,F298,F300,F263,F037,F192,F246,F357,F033,F278,F302,F055,F144,  
F141,F350,F252,F313,F364,F138,F255,F134,F040,F200,F088,F204,F235,F196,F326,F420,F274  
,F355,F103,F046,F117,F352,F126,F140,F092,F503,F219,F382,F287,F054,F303,F116,F043,F21  
8,F005,F190,F146,F194,F036,F199,F001,F191,F305,F281,F438,F121,F089,F075,F242,F193,F1  
27,F309,F133,F279,F202,F338,F344,F137,F500,F039,F135,F044,F421,F058,F301,F119,F245,F  
234,F267,F217,F142,F029,F045,F233,F250,F124,F122,F222,F247,F131,F047,F198,F282,F173  
GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTGC GAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F254,F177,F227,F080

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA

TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCAACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F434

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAAATTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAACACCACCATCATTGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F289

GCCACATTGATGCCTTGAAGAGAGATTTCTACGGCGGCGATGGAAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAACACCACCATATCGAGACTGAGTCCGACATCAT  
GCTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGGTCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCGTTATTC  
CAGTTGCTGGACAGTGCGTTCAA  
GTCAAGAACCATAGAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTCGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F110

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG

AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAT  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F325

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAAAGTGGTCATCTCAGAGCCATTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F435

GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATACCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F444

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACTCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC

TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F378,F111,F495

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACTGGAGTGGTCCTCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTCGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CAGAGCTCTTGCTCAAGTTTCGCTACC

>F458,F294,F391,F201,F472,F489,F165,F481,F154,F183,F163,F483,F395,F164,F505,F468,F1  
80,F076,F402,F459,F452,F171,F401,F498,F149,F415,F417,F513,F404,F167,F403,F159,F507,F  
078,F181,F182,F393,F394,F162,F392,F471,F405,F397,F470,F451,F517,F508,F150,F158,F168,  
F169,F464,F486,F166,F265,F184,F484,F104

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCATTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F385

GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCGTTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC

TGTGGAGATATCATTTAGCCCCGGCACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC

>F251,F277,F337,F256,F108,F413,F248,F203,F367,F258,F025,F018,F383,F216,F387,F280,F2  
69,F262,F223,F009,F465,F388,F249,F384,F206,F072,F514,F283,F253,F386,F049,F412,F024,F  
212,F236,F015,F259,F109,F225,F013,F407,F016,F228,F097,F273,F060,F237,F270,F210,F059,  
F226

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA

TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA

CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG

AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA

GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC

TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG

CGGAGCTCTTGCTCAAGTTTCGCTACC

>F446,F115,F113,F264,F331,F114,F112,F432

GCCACATTGATGCCTTGAAGAGAGATTTCTGCATCGGGAACGGAAAACCTCCCATCTA  
TCGAATATATGGCTGACCAAGAG

CCAACCGCAAATTACCCCGAGAATACCACTATCATCGAGACTGAATCCGACATCAT  
GCTCCATTTACCCGGCTCCACCGA

CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGCCATGTTCCATCATG  
GCGAGCTGCACTTTTTTCCTCTG

AAACCATTCAAGAGGGCAATCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGCCAGCGCGTTCAA

GTCAAGAACCATAAAGACAATTTGAGCACTATAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGCCCGAAAAAAGAC

TGTGGAGATATCATTTAGCCCCGACACCATGGAAATATCTGTTGTCTTGACGCATCA  
TCAAACAGCGGCGACGCAACCGG

CGGAGCTCTTGCTCAAGTTTCGCTACC

>F319,F379,F188,F342,F361

GCCACATTGATGCCTTGAAGAGAGATTTCTACGCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCAAAGAA

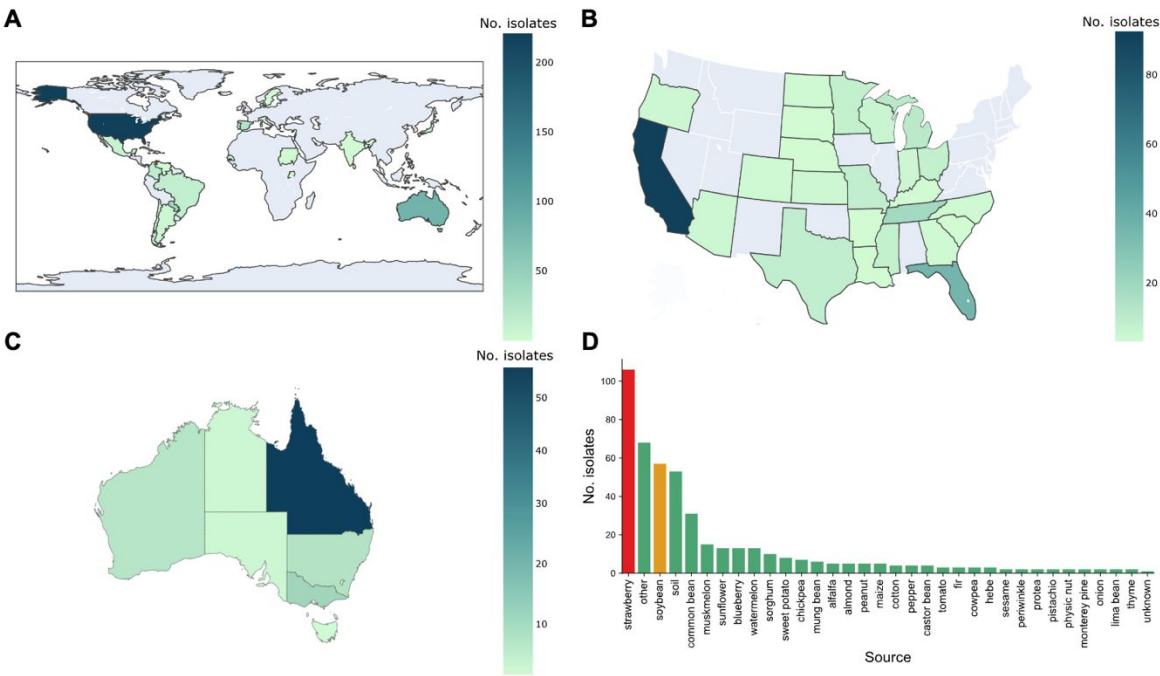
TCAACTGCAACCTACTCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA

CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCCATCGTG  
GCGAGCTGCACTTTTTTCCTCTG

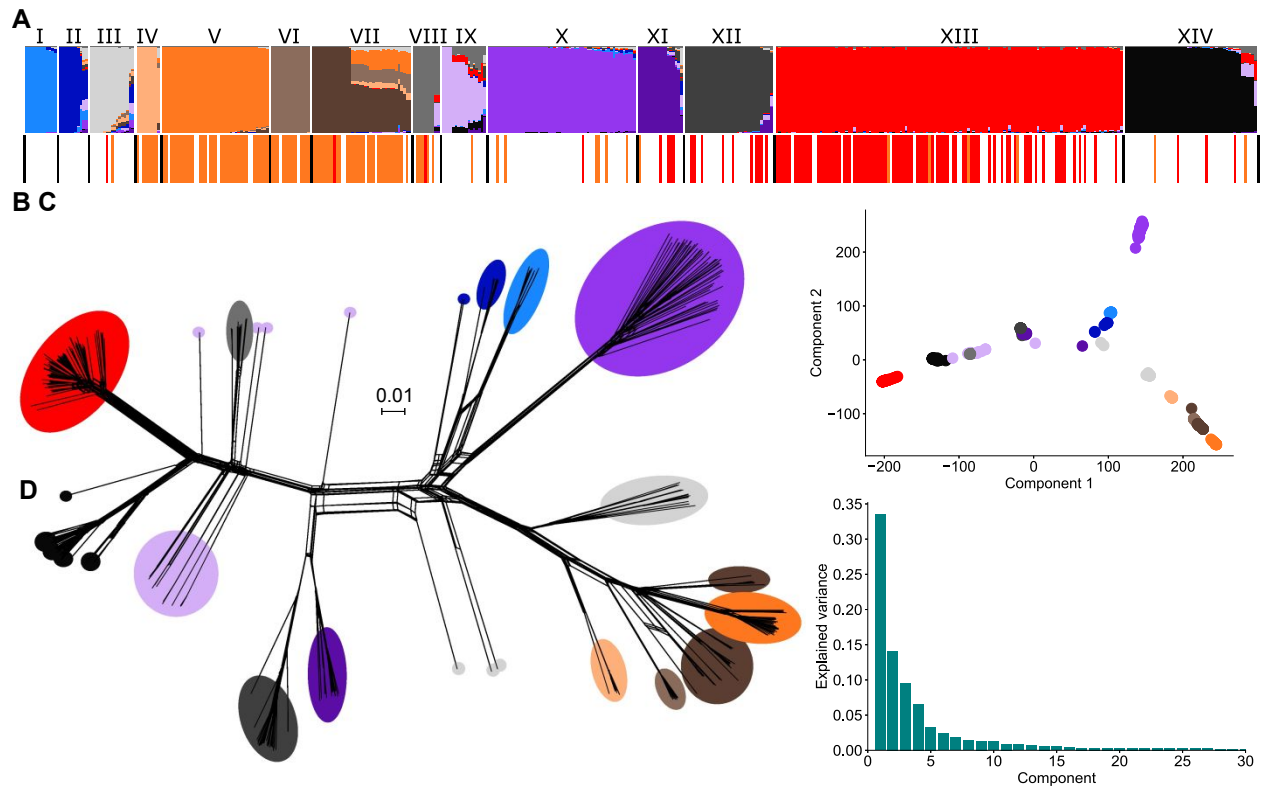
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGCGTTCAA

GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC

TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC  
>F030,F239,F510,F178,F100,F511,F032,F079,F102,F176,F151  
GCCACATTGATGCCTTGAAGAGAGATTTCTACCCCGGCGATGGAAATTTTCCGTCTA  
TCGAATATATGGCTGGCCAAGAA  
TCAACTGCAACCTACCCCGAGAATACCACCATCATCGAGACTGAGTCCGACATCATG  
CTCCATTTACCTGGTTCCACCGA  
CGATCTGCCTTCAGCTGAATCCGTCAGAAGGTTCTCGCTGGTCATGATCTATCGTG  
GCGAGCTGCACTTTTTTCCTCTG  
AAATCATTCAAGAGGGCAGTCGTTACCAGGAAAACCCCTTGAAACGCATCATTATTC  
CAGTTGCTGGACAGTGC GTTCAA  
GTCAAGAACCATAAAGACAATTTGAGCACCAGAGTGGTCATCTCAGAGCCACTGGG  
AATATCTCAAGTCCGAAAAAAGAC  
TGTGGAGATATCATTTAGCCCCGACACCATGGAAATCTCAGTTGTCTTGACGCATCA  
CCAAACAGCGGCGACGCAACCGG  
CGGAGCTCTTGCTCAAGTTTCGCTACC



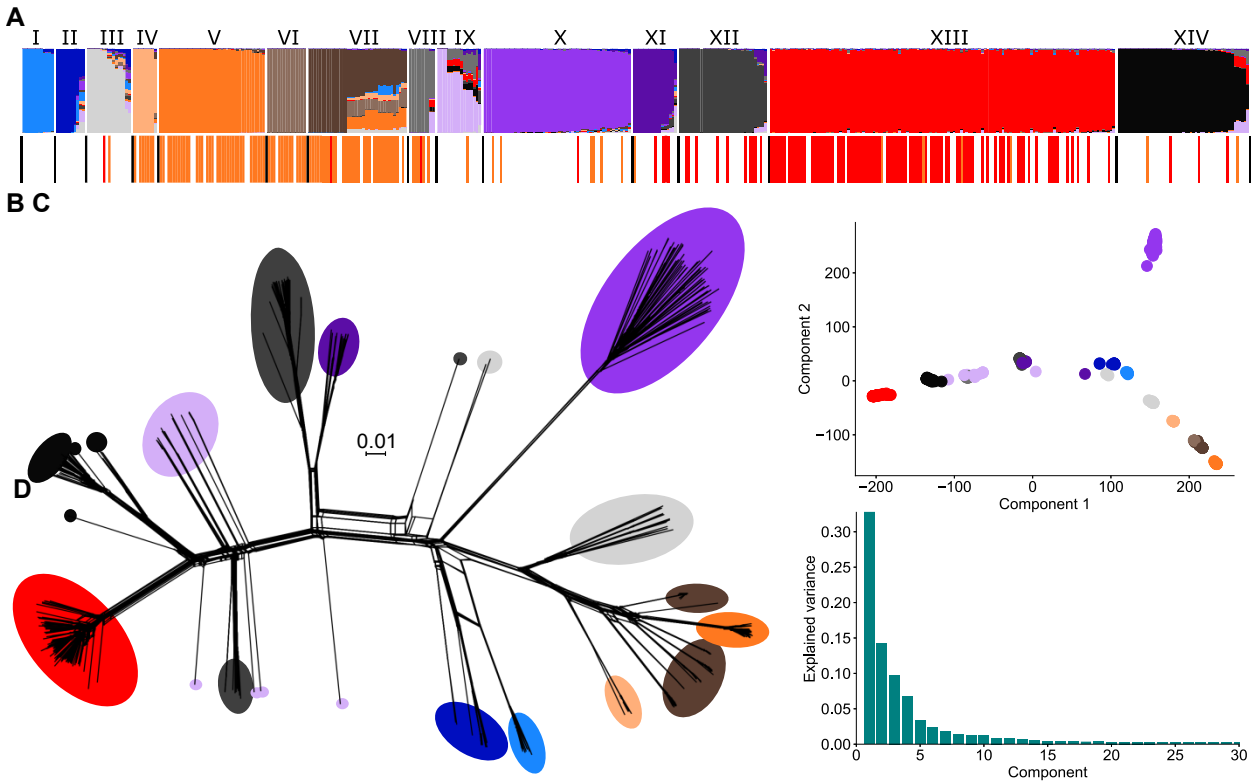
**Supplementary Fig. S1. Sources of *Macrophomina* spp. isolates.** **A, B, C** Geographic and **D** host origins of *Macrophomina* spp. isolates analyzed in the current study. Though **A** global collections, the plurality of isolates was collected in **B** the continental United States of America or **C** Australia, and **D** from strawberry or soybean plants. Choropleth colors correlate to the number of used isolates from a geographic region. Bar graph colors red and orange signify strawberry and soybean hosts.



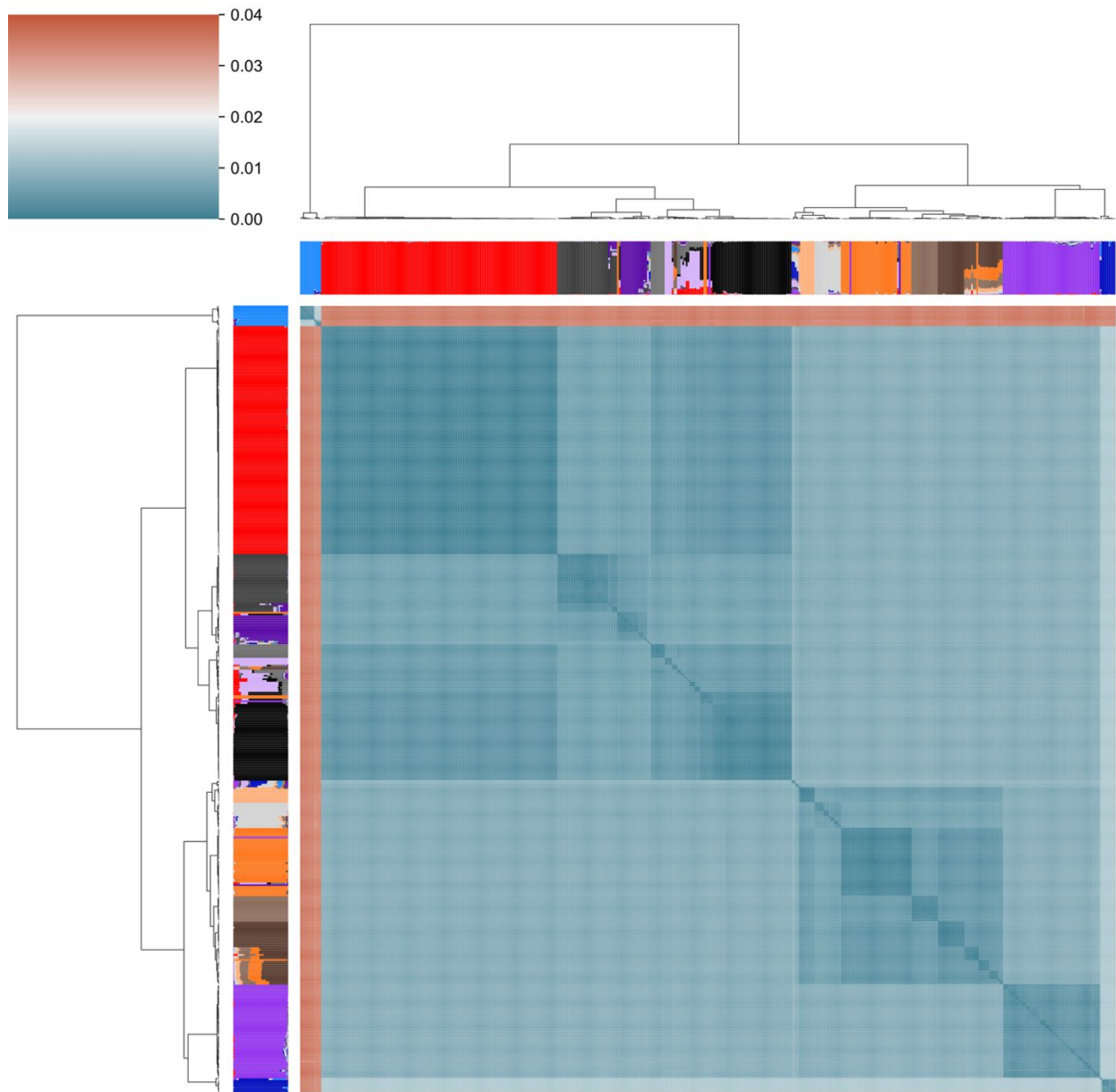
**Supplementary Fig. S2. Grouping of *Macrophomina* spp. isolates with AL-1-based SNPs.**

The **A** admixture plot was generated with sNMF at K=14. From left to right, the clusters are labeled with Roman numerals corresponding to those used in the main text, with Mp11-12. The **B** phylogenetic network and **C** PCA plots use the same colors as the admixture plot to indicate the major cluster assignment. The **D** Scree plot shows the eigenvalues for the first 30 principal components.

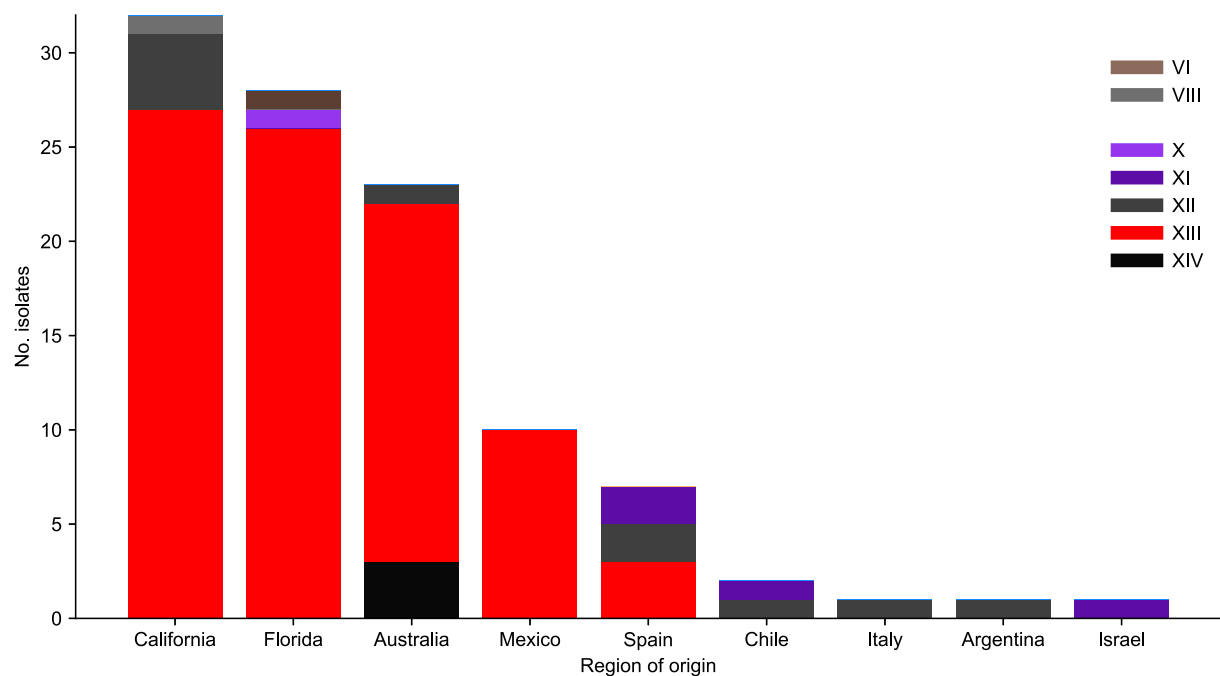




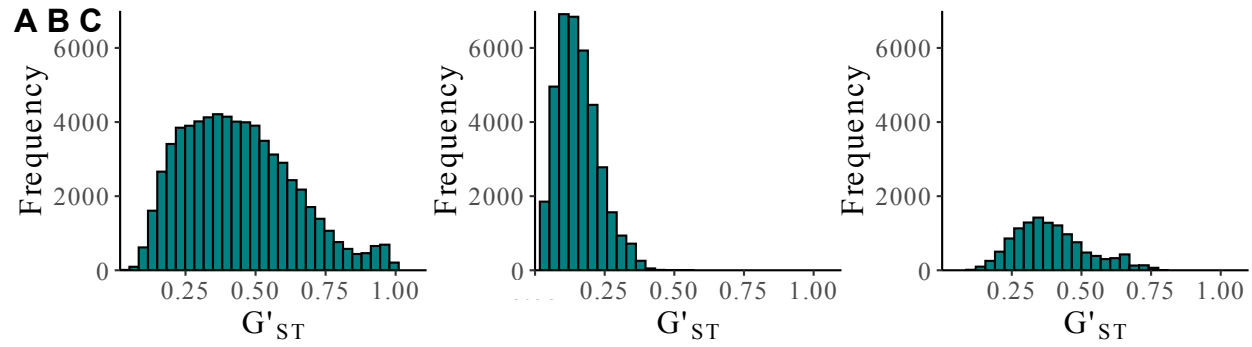
**Supplementary Fig. S3. Grouping of *Macrophomina* spp. isolates with mp117-based SNPs.** The **A** admixture plot was generated with sNMF at K=14. From left to right, the clusters are labeled with Roman numerals corresponding to those used in the main text, with Mp11-12. The **B** phylogenetic network and **C** PCA plots use the same colors as the admixture plot to indicate the major cluster assignment. The **D** Scree plot shows the eigenvalues for the first 30 principal components.



**Supplementary Fig. S4. *Macrophomina* spp. isolate relatedness according to pairwise anchor distances (ANDI).** Color value and intensity within the heatmap correspond to distance between a pair of isolates. Higher values indicate more dissimilarity. The corresponding Mp11-12 SNP-based admixture plot is overlaid on top utilizing the usual cluster-specific colors: I- light blue, II- dark blue, III- lightest grey, IV- light orange, V- orange, VI- light brown, VII- brown, VIII- grey, IX- lightest purple, X- purple, XI- darkest purple, XII- darkest grey, XIII- red, and XIV- black.

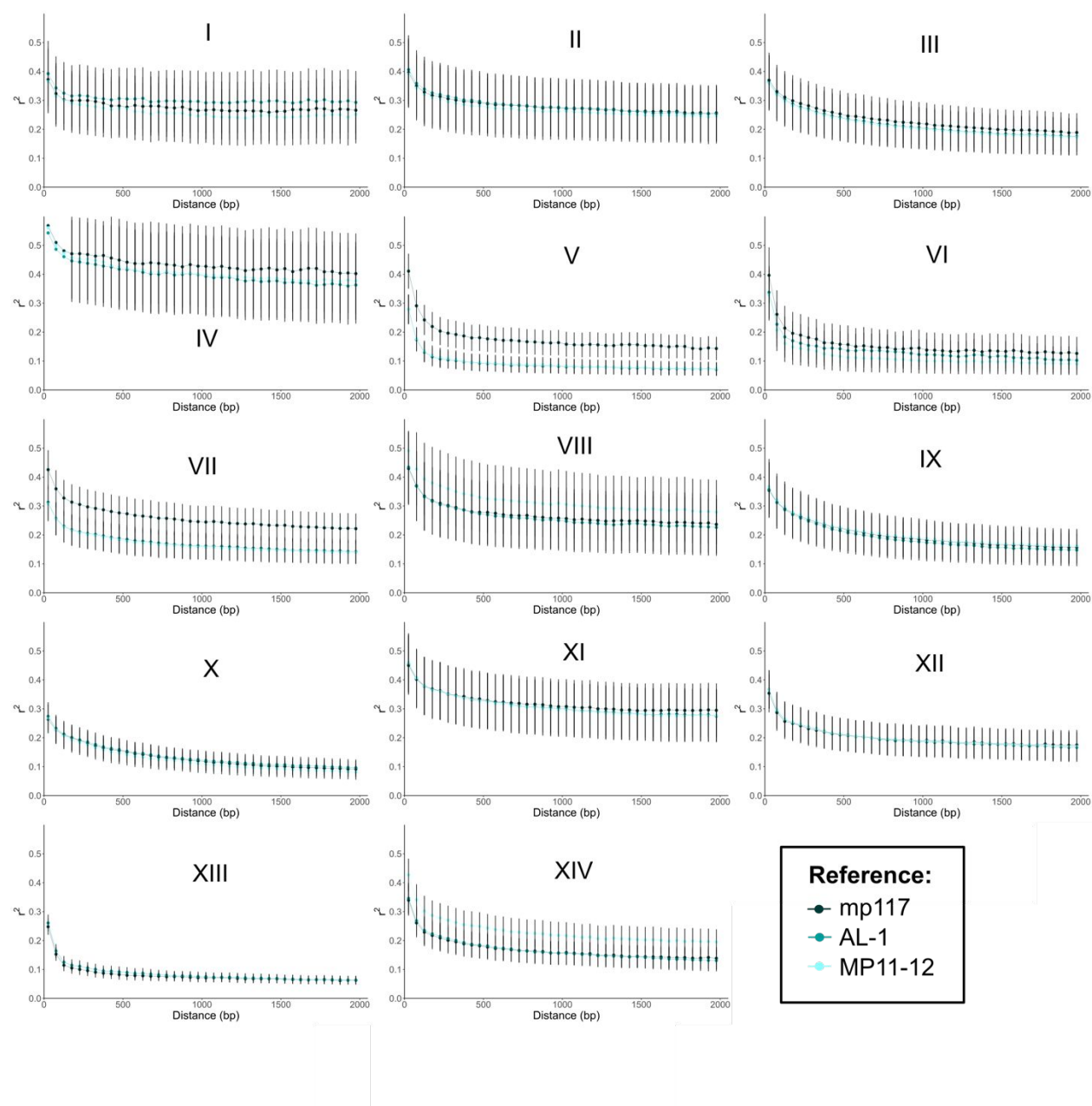


**Supplementary Fig. S5. Cluster assignments of isolates from strawberry by host country or USA state.** Most isolates from strawberry were assigned to cluster XIII.

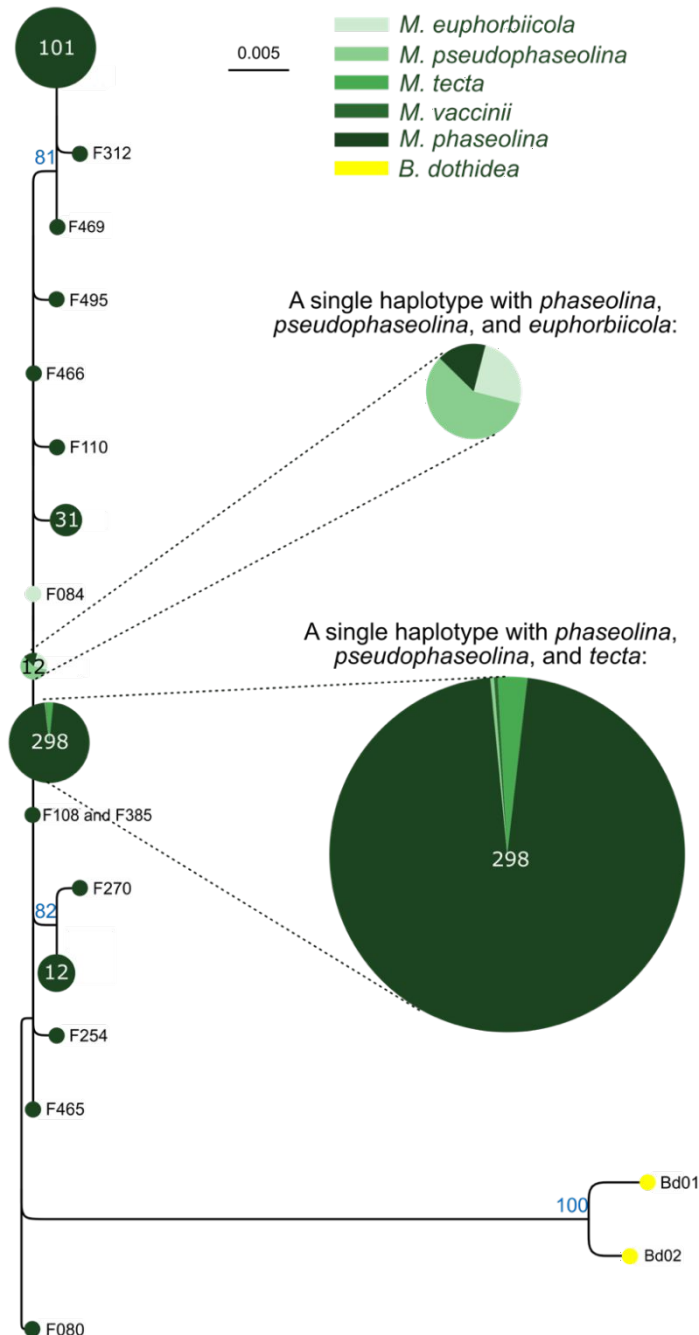


**Supplementary Fig. S6. *Macrophomina* spp. genetic variance in different subpopulations.**

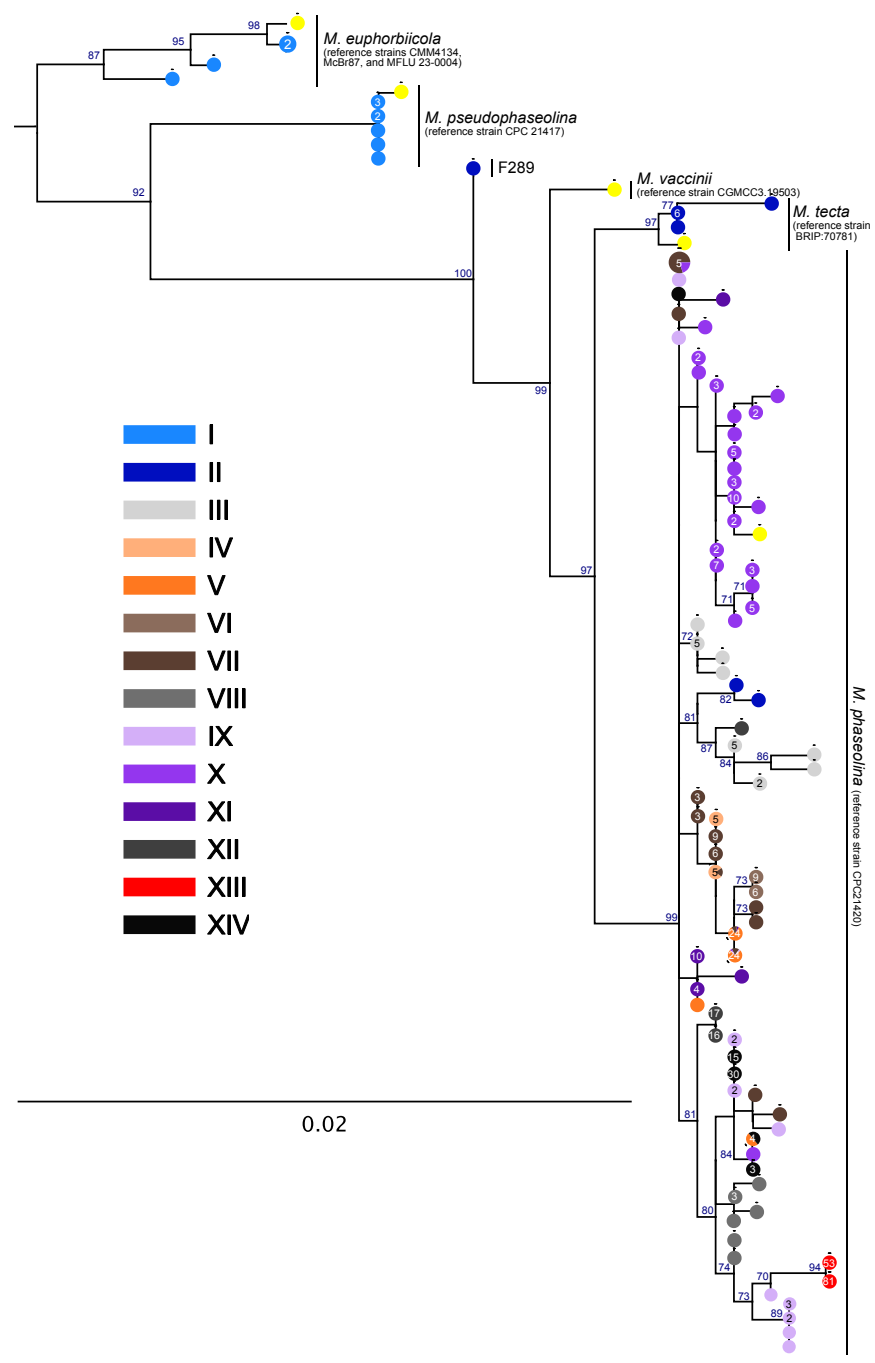
Fixation index distributions by **A** genotype cluster, **B** country of origin and **C** isolation source are shown. Histograms were generated with calculated  $G'_{ST}$  values from the R library vcfR using Mp11-12 SNPs.



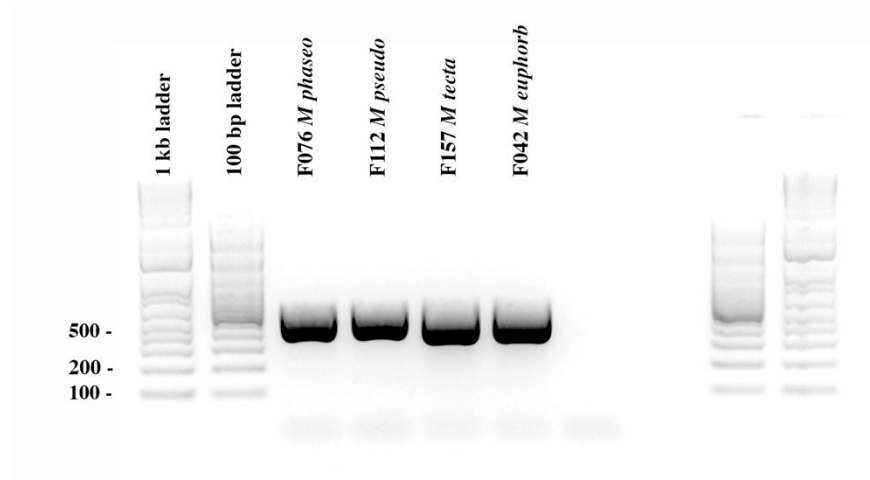
**Supplementary Fig. S7. Linkage associations of SNP loci.** Squared correlations calculated with SNPs based on *M. phaseolina* references Mp11-12, AL-1 and mp117 are shown. Error bars represent one standard error above and below the means. Linkage disequilibrium decay was analyzed separately for each cluster, and the cluster corresponding to each plot is indicated in roman numerals.



**Supplementary Fig. S8. Sequences similarities of the ITS regions of *Macrophomina* spp. isolates.** The phylogenetic tree was calculated by RAxML-NG version 1.2.1 with the GTR+G model of evolution and 1,000 bootstrap replicates. Only bootstrap values greater than 70 are shown. *Botryosphaeria dothidea* served as outgroups. Circles are approximately scaled to the number of isolates represented, which is also written in white text inside the circle. Circles representing isolates from more than one species are depicted as pie charts, with color codes shown in the upper right of the panel. Where a haplotype represents only one isolate, the isolate ID is shown to the right of the circle. The ITS region did not differentiate *Macrophomina* species, as most isolates had identical sequences. Used sequences are in Supplementary Information S1.



**Supplementary Fig. S9. Maximum likelihood phylogenetic tree constructed from concatenated ITS, *actin*, *beta-tubulin* and *TEF1- $\alpha$*  regions of *Macrophomina* spp. isolates.** The phylogenetic tree was calculated by RAxML-NG version 1.2.1 with the GTR+G model of evolution and 1,000 bootstrap replicates. Only bootstrap values greater than 70 are shown. Sequences from reference *Macrophomina* spp. genomes are denoted with yellow circles on branch tips. *Macrophomina* species could be differentiated by these combined sequences. Used sequences are in Supplementary Information S2.



**Supplementary Fig. S10. Gel showing 'mspp' target PCR amplicons from representatives four *Macrophomina* spp.**