STRUCTURAL GENOMICS OF *Fragaria*—WILD AND CULTIVATED STRAWBERRIES

By

DENISE CRISTINA MANFRIM TOMBOLATO

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2007

To:

- my father Vadir Tombolato, who has taught me the importance of moral integrity;
- my mother, Marlene Tombolato, who has, by example, taught me persistence;
- my professor, Kevin Folta, who permitted and encouraged me to exercise those virtues.

‘Yes,’ you will say, ‘but the plank is very long.’ That is true, and so if you do not have a sure foot and a steady eye, and are afraid of stumbling, do not venture down the path.

Jean de Léry, in "History of a Voyage to the Land of Brazil, Otherwise Called America", 1578

ACKNOWLEDGMENTS

I thank my parents Vadir and Marlene, and my brothers Eduardo and Ricardo, for their teachings, advice, support, and, above all, for their unconditional love. Though not content with my departure from Brazil, my family always supported my decisions. I appreciate their confidence in my choices and me, for it reaffirmed my personal mission in moments of doubt.

I am grateful to my professor Dr. Kevin M. Folta, who accepted me as his student in an altruistic gesture, and who has been a *lato sensu* adviser since. I thank the members of my committee for the enjoyable discussions about my project and about science in general: drs. A. Mark Settles, Natália A. R. Peres, and Craig K. Chandler. I also wish to thank my laboratory colleagues and friends drs. Philip J. Stewart and Amit Dhingra, Thelma F. Madzima, Stefanie A. Maruhnich, Jeremy Ramdial, Dawn Bies, and Maureen Clancy, as well as project collaborators drs. Thomas M. Davis and Daniel J. Sargent, for DNA sequences and plant material from the genetic linkage mapping population.

Many people made special the almost-9 years I spent in Gainesville, while I pursued part of my undergraduate training and two advanced degrees. I convey my gratitude to all those who facilitated not only my adaptation to a new country and language, but also the discovery of who I am and of matters I learned to be truly meaningful. I recognize Welch McNair Bostick III ("McNair"), whose short life was vastly fruitful. McNair caused positive impact into the lives of whomever surrounded him: his wife and my friend Carmen Valero, his neighbors (including myself), and his colleagues. I thank him for having shown to me the importance of treasuring the time shared with loved ones, expressing honest opinions and making a difference in society.

I express my appreciation for the time and assistance granted to me by professors and technicians with whom I worked since my arrival to the University of Florida: Richard D.

Berger, Terry A. Davoli, D. Pete Weingartner, Jeffrey A. Rollins, Ulla Benny, Valerie Jones, Jeffrey B. Jones, and Jerry Minsavage.

I thank these individuals for the attention they have dedicated to me: Balša Terzic', Sylvia Morais de Sousa, Gisele, Jens, and Gabriel Schöene, Mark D. Skowronski, Luciana C. B. Manfrim Bchir, Gustavo Ramirez, Juliana and Gustavo Astua, Aaron Hert, Botond Balogh, Abby Guerra, Ahu Demir, Petrônio Pinheiro, Ilka V. Araújo, Maggie Kellogg, Maria Beatriz Pádua, Melissa Webb, Bruno Maciel, Camila A. Brito C. Paula, Luiz Augusto de Castro e Paula, Hazar Dib, Marlise Klein, Marcus Martin, Michelle Bolton, Sonja I. Parisek, Penny E. Robinson, Anne Visscher, Ricardo da Costa Mattos, Claudia Riegel, Valerie Rodriguez-Garcia Schweigert, Lisa Olsen, Jared Greenberg, Wendy Gonzalez, and David Adato. Every one of them made my life in Gainesville a more enjoyable experience.

TABLE OF CONTENTS

page

APPENDIX

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

STRUCTURAL GENOMICS OF *Fragaria*—WILD AND CULTIVATED STRAWBERRIES

By

Denise Cristina Manfrim Tombolato

August 2007

Chair: Kevin M. Folta
Major: Horticultural Science

The extensive phenotypic variability and complex genetic makeup of the cultivated strawberry *Fragaria × ananassa* permits advances in plant improvement, a factor breeders have exploited to great benefit. However, the introgression of specific characters is complicated due to the cumbersome genetics and limited knowledge of genome structure and function of genes relevant to traits of interest. The present study represents the first genomics-level insight into strawberry genome structure and explores the hypothesis that a new type of molecular marker, the Gene-Pair Haplotype represents a transferable marker that may hasten linkage mapping in the diploid and octoploid strawberry.

My research presents the findings of four related research activities. First, an efficient and unified method for genomic DNA isolation was derived from over 100 experimental tests and conditions. Next, 1% of the *Fragaria* genome was sequenced and functionally annotated, using a bioinformatics approach and computational tools. Over 120 kb of intergenic regions were sequenced using the Gene-Pair-Haplotype approach, allowing for some initial relationships to be formulated concerning the diploid subgenome contribution to octoploid strawberry. Finally, Gene-Pair Haplotypes were used to add a suite of alleles to the growing *Fragaria* linkage map. These findings provide a starting point for further analyses of the strawberry genome.

CHAPTER 1
STRAWBERRY AND THE GENOMICS ERA

Introduction

The cultivated strawberry, *Fragaria* × *ananassa* Duch*,* belongs to the family Rosaceae as do the also economically important crops rose, apple, pear, peach, cherry, plum, raspberry, and almond. Linnaeus named the genus *Fragaria* due to its fragrant properties, whereas the odor, taste and berry shape was thought to be similar to pineapple, or “ananas”, in Latin (Darrow, 1966). In 1765, the *F.* × *ananassa* parentage was proposed by Antoine Nicolas Duchesne, whose father worked at the Court of Louis XV (Darrow, 1966). *F.* × *ananassa* was first observed in several countries in Europe since the 1750’s and it originated from a spontaneous hybridization between *F. virginiana* and *F. chiloensis,* both from the American continent. *F. virginiana* is thought to have been imported to Europe by two routes (Wilhelm and Sagen, 1974): to France by the explorer Jacques Cartier during his first expedition to the Quebec Canadian Province in 1534; and to England, by Thomas Hariot, who visited the “New Found Land of Virginia” in 1588. Later, in 1714, *F. chiloensis* was taken to France by the engineer Amédée François Frézier. During his mission to study the defense fortifications of Chile and Peru, Frézier noticed the large-fruited berries at Concepción, Chile, and collected several plants to take back to his country (Darrow, 1966). The result of the accidental cross between the two *Fragaria* species was the basis for the creation of the fruit cultivated and appreciated throughout the world today.

Profitable strawberry production is challenged by several factors: diseases, pests, market competition, and, arguably most importantly, by the phase-out of methyl bromide. This fumigant is considered essential for the production of many crops, including strawberry (Rosskopf et al., 2005), but because methyl bromide has great stratospheric ozone depletion ability, the Montreal Protocol mandates that its use be reduced (Anonymous, 1998). Although traditional plant breeding has been used to remedy several of the above-mentioned challenges, the knowledge of the *Fragaria* genome structure may streamline the variety improvement process, potentially permit discovery of gene function, and ultimately lead to more diverse and hypothesis-based solutions to traditional and contemporary problems not only for the strawberry but also for other Rosaceous crops.

Molecular Markers for Strawberry

The cultivated strawberry has a complex (2n=8x=56) (Ichijima, 1926), (Fedorova, 1946) and poorly understood genome. Despite strawberry's commercial value of 1.4 billion dollars as a fruit crop (Folta et al., 2005), substantial knowledge of *Fragaria* structural genomics before this project was virtually nonexistent. Sequence information facilitates the development of molecular markers that can be used for marker-assisted selection (Haymes et al., 1997), (Van de Weg, 1997), (Albani et al., 2004), (Sugimoto et al., 2005), (Haymes et al., 2000), (Lerceteau-Köhler et al., 2002), clone characterization in germplasm banks (Harrison et al., 1997), (James et al., 2003), identification of cultivar proprietary (Arulsekar et al., 1981), (Bringhurst et al., 1981), (Gidoni et al., 1994), (Nehra et al., 1991), (Bell and Simpson, 1994), (Hancock et al., 1994), (Levi et al., 1994), (Parent and Page, 1995), (Landry et al., 1997), (Degani et al., 1998), population genetics studies (Degani et al., 2001), (Harrison et al., 1997), (Graham et al., 1996), (Arnau et al., 2003), (Hadonou et al., 2004), and construction of genetic linkage maps (Williamson et al., 1995), (Yu and Davis, 1995), (Davis and Yu, 1997), (Deng and Davis, 2001), (Lerceteau-Köhler et al., 2003), (Sargent et al., 2003), (Sargent et al., 2004).

Pioneer molecular markers were based on polymorphisms observed on punctual loci or the whole genome: isozymes and intron length polymorphism; Randomly Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP), and Amplified Fragment Length Polymorphism (AFLP). More recently, Simple Sequence Repeats (SSRs) have been employed to address the challenge of marker transferability (Monfort et al., 2005), (Nourse et al., 2002), (Ashley et al., 2003). The present work discusses the creation of a novel marker type that, in addition to responding to the transferability necessity of modern markers, also attaches functional information to markers generated.

The Genomics Era

“Genomics” has been defined as “the study of all nucleotide sequences, including structural genes, regulatory sequences, and noncoding DNA segments, in the chromosomes of an organism.” (The American Heritage®, 2006)

The complexity of plant genomes began to be investigated in the mid- to late-1970’s using quantitative DNA reassociation kinetics (i.e. Cot curves) (Goldberg, 2001). It was determined that plant genomes had families of repetitive sequences and that these repeats varied in copy number and arrangement in the genome (Flavell et al., 1974), (Goldberg, 1978).

By the end of the 1970’s, with the ability to construct cDNA clones, there was the surprising finding that the coding regions of eukaryotic genes were interrupted by introns (Gilbert, 1978), what led to investigation of posttranscriptional splicing mechanisms (Jeffreys and Flavell, 1977).

The first plant gene was cloned in 1979 (Bedbrook et al., 1980), demonstrating that plant DNA was not different from the DNA of other organisms and therefore could be manipulated using the same enzymes, cells, and vector systems. The result was the construction of both plant genomic and cDNA libraries of many plants and organs (Goldberg, 2001).

The demonstration that *Agrobacterium tumefaciens* tumor DNA (T-DNA) integrates into the chromosomes of plant cells (Chilton et al., 1977) created the opportunity to generate transgenic plants, the first one being sunflower cells expressing bean phaseolin seed storage protein gene (Murai and Sutton DW, 1983). In addition to being a vector to foreign genes, T-

DNA began to be used to generate transformed *Arabidopsis* lines with mutant phenotypes to identify and clone important plant genes, such as genes involved in the control of meristem identity and hormone perception (Feldmann, 1991), (Feldmann and Marks, 1987). A second method to clone plant genes was devised upon the isolation of the Ac and Ds transposable elements (Fedoroff et al., 1983).

The beginning of the sequencing era can be attributed to the determination of a bacteriophage RNA gene sequence in 1972 (Min Jou et al., 1972). The first whole-genome sequencing was also from a virus, *Haemophilus influenza*, completed in 1995 (Fleischmann et al., 1995), whereas a draft of the Human Genome was released in 2001 (Venter JC, 2001). The first plant genome sequenced was *Arabidopsis thaliana*, completed in 2000 (The Arabidopsis Genome Initiative, 2000). In 2007, approximately 2300 sequencing projects are being carried out or completed, of which about 130 are plant genomes, according to the Genomes Online database (Liolios et al., 2006). Our collaborators in this *Fragaria* genomics project have successfully completed 1% of the *F. vesca* genome. Recently, *Malus* (apple) was selected for full sequencing by an Italian sequencing effort. Peach also will be sequenced through a US Department of Energy initiative. Although these genomes are much larger than the strawberry genome, their completion will have important ramifications to *Fragaria*, as annotation will provide a list of components that are similar to those in strawberry. The work presented here is a complementary effort to those in other rosaceous crops, providing an initial glimpse into the genome of one of the world's most prized horticultural crops.

CHAPTER 2
DNA EXTRACTION FROM RECALCITRANT SPECIES

Introduction

Strawberry (*Fragaria* × *ananassa*) is an important crop worldwide, and it supports many regional economies in the United States. However, relatively little is known about the genes that govern agriculturally important traits or their expression. Contemporary genomics tools have the potential to accelerate study of strawberry and bring additional resolution to strawberry gene form and function. Strawberry belongs to the genus *Fragaria*, a genus that includes a number of species of varying ploidy with a small haploid genome size. These facets make strawberry an excellent candidate for genomic studies representing the Rosaceae family. Because it is easily transformable, it is particularly well suited for translational-genomics studies.

Any genomics effort, whether translational, structural or functional, is generally dependent on a reproducible and effective means to isolate quality genetic material. Although protocols have been streamlined over the last several decades, it is challenging to isolate large amounts of quality DNA from strawberry (Manning, 1991; Porebski et al., 1997). A similar problem has been encountered in other species. Plants like cotton (Katterman and Shattuck, 1983; Dabo et al., 1993; Chaudhry et al., 1999; Li et al., 2001), sugarcane (Aljanabi et al., 1999), conifers (Crowley et al., 2003), tomato (Peterson et al., 1997), grape (Collins and Symons, 1992; Lodhi et al., 1994), and the rosaceous chestnut rose (Xu et al., 2004) have been reported to be recalcitrant to DNA extraction. The high content of polysaccharides and polyphenols either limit DNA isolation or inhibit downstream enzymatic reactions.

The DNA Extraction Procedure

A typical DNA extraction is accomplished by three basic steps: lysis of the cell, removal of proteins, and separation of nucleic acids from other cellular compounds. Cell lysis is easily achieved by removal of membrane lipids with detergents such as sodium dodecyl sulfate (SDS), triisopropylnaphthalenesulfonic acid (TIPS) (Bies and Folta, 2004), and N-laurylsarcosine (sarkosyl) when extracting DNA from bacterial or animal cells; however, because plants have a solid cell wall in addition to the cellular membrane, solvents alone are not enough to expose organelles, and mechanical force must be applied. Samples can be sonicated but generally are either treated with ethyl ether (Watson and Thompson, 1986; Peterson et al., 1997; Folta and Kaufman, 2000; Peterson et al., 2000), lyophilized or frozen in liquid nitrogen to make the material more friable prior to manual grinding. Additional homogenization is performed with a Polytron or comparable tissue disruptor.

Cell lysis is carried out either as a single step, breaking open all cellular compartments simultaneously, or in a stepwise fashion, first rupturing outer membranes to expose the nucleus, then solubilizing the nuclear envelope to free nucleic acids. The first membrane lysis is induced by osmotic pressure generated by 0.35M sorbitol (Fulton et al., 1995; Hanania et al., 2004) , 0.35M glucose (Chaudhry et al., 1999) or Triton X-100 (which lyses chloroplasts and mitochondria, but does not solubilize nuclear DNA) (Watson and Thompson, 1986; Peterson et al., 1997), while the second lysis is performed by detergents and ethylenediaminetetraacetate (EDTA). During this perturbation of the cell, DNA-degrading enzymes must be inhibited, which is accomplished by manipulating pH and removing divalent cations. Since DNAses act at pH 7.0, Tris is added to raise the pH to between 7.5 and 8.0. The chelation of divalent cations (Ca^{2+}, Mg^{2+}) by EDTA prevents the activity of metal-dependent enzymes.

Cellular and histone proteins can be dissociated by SDS (Kay and Dounce, 1953), proteases, chaotropic agents, chloroform (Sevag et al., 1938), and phenol. Because phenol solubilizes proteins (Cohn and Conant, 1926), it has been used to deproteinize preparations of carbohydrates (Westphal et al., 1952; Westphal and Jann, 1965) and nucleic acids (Kirby, 1956). Chaotropic agents denature proteins by increasing the solubility of nonpolar substances in water (Voet et al., 1998). Hofmeister (Hofmeister, 1888) defined the series of anions and cations with increasing protein destabilizing properties when he measured the concentration of various salts needed to precipitate proteins from whole egg white (translated by (Kunz et al., 2004)). According to the Hofmeister series, urea, guanidinium, thiocyanate (Sawyer and Puckridge, 1973) and perchlorate (Wilcockson, 1973) are extremely chaotropic agents. Thus, high concentrations of urea (Settles et al., 2004), guanidine hydrochloride (Logemann et al., 1987), and guanidine thiocyanate have been used in isolation of RNA (Cox, 1968; Chomczynski and Sacchi, 1987) and DNA (Chomczynski et al., 1997).

Chemical or physical means such as precipitation by isopropanol, ethanol, butoxyethanol (Manning, 1991), acetone (Vogelstein and Gillespie, 1979), adsorption to silica (Vogelstein and Gillespie, 1979), paramagnetic particles (Anonymous, 1980, 2001; Koller and al., 2001), and ion exchange resin (QIAGEN Anion-Exchange Resin manual) can be utilized to retrieve DNA from solution. The resin is coated with diethylethanolamine (DEAE), and DNA recovery is due to interaction between negatively charged phosphates of the DNA backbone and positively charged DEAE groups. In the case of silica columns, DNA is recovered from solutions because it tends to adsorb to silica in the presence of chaotropic salts, such as sodium iodide (NaI) (Vogelstein and Gillespie, 1979), guanidine thiocyanate, and guanidine hydrochloride. The binding capacity depends on the solution’s ionic strength and pH, being higher in concentrated solutions and at pH<7.5 (GeneClean Manual). Silica columns have been used to eliminate polysaccharide contaminants, and the ratio A260/230 increases as polysaccharides are removed (Abdulova et al., 2002).

DNA Extraction from Plants

Pioneer methods to isolate genetic material of plants used DNA-rich matter such as germ tissue (Lipshitz and Chargaff, 1956; Shapiro and Chargaff, 1960). Early attempts to extract DNA from leaves resulted in degraded product due to the extreme pHs used by the procedure for removal of RNA (Thomas and Sherratt, 1956). The currently most used protocol for plant DNA isolation, developed by Murray and Thompson (Murray and Thompson, 1980), takes advantage of the selective precipitation of DNA by cetyltrimethylammonium bromide (CTAB), a phenomenon observed by Jones during DNA isolation from bacteria (Jones, 1953). CTAB is a cationic detergent that, in high ionic strength solutions (e.g. >0.7M NaCl), complexes with proteins and non-acidic polysaccharides, whereas at low ionic strength it precipitates nucleic acids and acidic polysaccharides, leaving proteins and neutral sugars in solution (Sambrook and Russell, 2001). Multiple variations of Murray and Thompson's protocol have been used by researchers to adapt the original process to different plant species. A protocol designed by Doyle and Doyle (Doyle and Doyle, 1987) is also frequently used for plant DNA extraction and is ultimately a variation of the Murray and Thompson procedure. Doyle and Doyle's protocol uses fresh tissue in place of lyophilized material and a higher concentration of CTAB and salt to compensate for the greater water content of fresh tissue.

Although CTAB is the reagent of choice to purify DNA from organisms that produce many polysaccharides (Sambrook and Russell, 2001), even high quantities of the cationic detergent seem insufficient to free DNA preparations from sugar contamination. In attempt to circumvent this problem, boric acid is added to the extraction buffer. Boric acid forms complexes with polyphenols at pH 7.5 (King, 1971) and with carbohydrates (Gauch and Dugger Jr., 1953), making these complexes more soluble. An additional approach to avoid co-purification of polysaccharides during DNA isolation is to differentially precipitate the sugars by manipulating the 2-butoxyethanol concentration (Manning, 1991).

Cytoplasmic compounds come into contact with nuclei contents when cells are disrupted and the oxidized polyphenols covalently link to DNA (Loomis, 1974), restraining subsequent DNA manipulation (Katterman and Shattuck, 1983). Reducing agents like ß-mercaptoethanol, dithiothreitol, ascorbic acid, sodium bisulfite, and diethylcarbamic acid can be added to the extraction buffer to inhibit the oxidation process and protect DNA from quinones, disulfutes, peroxidases, and polyphenoloxydases. Polyvinylpyrrolidone (PVP) and its insoluble, cross-linked form, PVPP (Gegenheimer, 1990), also protect DNA from phenolics and alkaloids by sequestering them. Additional approaches to avoid problems caused by phenolics like freezing tissue prior to homogenization (Katterman and Shattuck, 1983; Leutwiler et al., 1984), purification by cesium chloride gradient (Travaglini and Meloni, 1962; Williamson, 1969; Murray and Thompson, 1980), and extraction of DNA from isolated nuclei (Hamilton et al., 1972; Katterman and Shattuck, 1983; Watson and Thompson, 1986; Peterson et al., 1997) have been used.

As genomics tools become more common in strawberry research, it is imperative to devise a standard protocol that is effective across cultivars and species of different ploidy levels. Examination of the literature on strawberry (*Fragaria* spp.) indicates that the many published DNA isolation methods are not universally transferable between cultivars or species. An optimal protocol should use readily available plant material (such as mature leaves), be inexpensive, rapid, reproducible, and have high yields of high molecular weight DNA, amenable to downstream manipulation. Of all these traits, quality is most important, yield second in importance, followed by cost and ease of protocol.

Material and Methods

Thirty-three DNA extraction protocols, totaling 103 treatments, were tested using either lyophilized or liquid nitrogen-frozen leaf tissues. A broad range of genotypes were tested, including tissue from *F. nubicola*, *F. vesca* cultivars Yellow Wonder, Alexandria, and Hawaii-4, *F. chiloensis* CA 1367*, F. virginiana* NC 96-35-2*, F.* × *ananassa* cultivars Sweet Charlie, Tristar, Camarosa, Quinault, Diamante, Strawberry Festival, and the laboratory transformation genotype LF9 (Folta et al., 2006). The detailed protocols can be found in Appendix A, whereas further below is a summary of the approaches adopted. When at least 15µg of DNA were obtained, digestion of 5µg of DNA with at least 2 separate restriction enzymes were carried out. The uncut and enzyme treated samples were loaded on 1% agarose gel for assessment of DNA quality (integrity and amenability to use of restriction enzymes), and correlation to spectrophotometric readings. Phenols are known to absorb at 260nm as does DNA, and high readings may be attributed to the presence of phenols, particularly when the DNA pellet has brown coloration, caused by oxidation of phenolic compounds (phenylpropanoid and flavonoids) to quinones (Loomis, 1974). To further test the quality of the DNA preparations, PCR was carried out using primers for *F.* × *ananassa* 18S ribosomal DNA. The primers (forward: 5’ TAT GGG TGG TGG TGC ATG GC 3’; reverse: 5’ TTG TTA CGA CTT CTC CTT CC 3’) were designed utilizing as sequence source the accession gi|18448|emb|X15590.1|FA18S. The fragment to be amplified by this primer pair is not large (510bp from cDNA, ~1kb from genomic) and should be easily amplified, since many copies of ribosomal DNA are present in the genome. If a product was observed, a second set of primers (forward: 5’ CAC TGC CAA GGA GCG TGG TG 3’; reverse: 5’ TCA GTA GGG CAG CTG ATG 3’) targeting a single-copy region, the *Leafy* gene, was used to provide a more challenging test. This second primer pair was designed utilizing *F. vesca* ‘Pawtuckaway’ sequence provided by our collaborator, Dr. Thomas

M. Davis, and encompassed a 770-nucleotide region. Both PCR reactions were carried out for 35 cycles, with 55ºC as annealing temperature, and 1min as extension at 72ºC.

The original CTAB protocol designed by Murray and Thompson (Murray and Thompson, 1980) is extremely laborious, requiring a long centrifugation period in a cesium chloride (CsCl) gradient. Since the aim of this project was to develop a rapid, practical method to extract DNA, the CsCl step was omitted from all DNA extraction attempts. Further modifications of the protocol were tested systematically to pyramid the beneficial aspects of each preparation into a unified and effective means to generate high-quality DNA for downstream analysis as described bellow:

- CTAB was tested at 1, 2, 6, and 20%
- Inclusion of one or combinations of the following reagents to prevent DNA oxidation: 0.01% -1% sodium (bi)sulfite, 5mM ascorbic acid, 1-4% PVP
- EDTA concentration from 10mM (as proposed by Murray and Thompson) to 200mM
- Tris concentration ranged from 50mM (as in original protocol) to 200mM. The pH was adjusted to 8.0 by addition of HCl. In cases where boric acid was used to adjust the pH, the Tris-borate solution was brought to pH 7.6 because at that pH, boric acid forms complexes with polyphenols
- The original protocol removes proteins by treating the solution with 24:1 chloroform:octanol. Alternative deproteination methods tested were: 25:24:1 phenol:chloroform:isoamyl alcohol, 1M sodium perchlorate, and 150µg/ml proteinase K
- DNA was recovered by either adsorption to silica, or precipitation by ethanol, isopropanol, 2-butoxyethanol, or 5M potassium acetate. In Murray and Thompson's original protocol, DNA is precipitated by decreasing salt concentration
- Attempts to remove water-soluble contaminants by adsorption to silica column (QIAGEN DNeasy kit) and by dialyses of DNA solution into TE pH 7.0 at 4°C
- Instead of adding buffer subsequent to grinding the plant tissue, an additional tissue/buffer homogenization step was performed. An aliquot of the final buffer was used to either produce a tissue/buffer paste in the mortar and pestle or Polytron homogenizer
- In place of the standard incubation in buffer at 50-60ºC for 20-30 minutes, incubation was carried out at 4, 20, 42, and 65ºC for 0, 5, 30, and 60 minutes. In order to eliminate

variability that may be induced because of the leaves of various ages, leaves were cut with a hole puncher, mixed, and split into 4 portions, one for each temperature treatment. Enough plant tissue was ground per temperature treatment so that 2 experimental replicates for each time treatment were derived from a single test tube (see figure 2-1).

In addition to variations of the CTAB protocol, other approaches adopted included use of the chaotropes 8M urea, 4M guanidine thiocyanate (alone or in combination with 2% CTAB, simultaneously or sequentially); DNA isolation using kits: QIAGEN DNeasy Plant Mini Kit (charged resin-based), Molecular Research Center DNAzol® Extra Strength (guanidine thiocyanate-based), Epicentre MasterPure^{TM} Plant Leaf DNA Purification, MoBio PowerPlant^{TM} DNA Isolation Kit; 0.5% SDS, Tris-borate extraction buffer; and crude and fine isolations of nuclei prior to DNA extraction. Five DNA extraction procedures, QIAGEN DNeasy kit, 2% CTAB, 2% SDS, 4M guanidine thiocyanate/1% sarkosyl, and 5% SDS/1%TIPS, were tested on Percoll gradient-isolated nuclei. Refer to table 2-1 for all the treatments.

The amount of tissue necessary to obtain the highest DNA extraction efficienty was determined by keeping the volume of buffer constant at 5ml and varying the tissue weights at 50, 200, 500, and 1,000mg. Once the best tissue-to-buffer ratio was determined, an attempt to extract DNA from 10 species within the genus *Fragaria* was made to test the universality of the method. Each treatment had 2 replicates for both experiments. Expanded leaf tissue was ground in liquid nitrogen, added to the buffer, and the mixture was incubated at 4ºC for 5 minutes. An equal volume (5ml) of 24:1 chloroform:octanol were added to the tubes after incubation, agitated, and centrifuged at 4,000rpm for 5 minutes. The aqueous phase was transferred to a new tube, and nucleic acids precipitated by 1/10 volume of 5M NaCl and 7/10 volume of isopropanol. After a second centrifugation, the supernatant was decanted, the pellet air-dried, and resuspended in 500µl water. RNAse was added to final concentration of 50µg/ml. The solution was transferred to 1.5-ml tubes and DNA was precipitated as described above. The dry DNA pellet was resuspended in 200µl water and DNA quantities were estimated by a NanoDrop ND-1000 spectrophotometer.

Nucleic acids were extracted from 96 individuals that belong to a diploid *Fragaria* mapping population. Minimal quantities of lyophilized tissue were processed, ranging from 3 to 14mg (average = 6.44mg, standard deviation =1.98). Because the buffer volume was kept constant, there was an opportunity to further study tissue-to-buffer ratios, under different conditions from those tested above. This time, tissue was macerated in buffer after having been ground in liquid nitrogen and incubated at 65ºC for 1hour. The absorbance values at 230, 260, and 280nm were determined by a NanoDrop to make inferences about nucleic acid purity. Absorbance ratios A260/A230 and A260/A280 are measures of contamination by polyphenols or carbohydrates (Craigie and McLachlan, 1964; Logemann et al., 1987), cited by (Manning, 1991) and protein, respectively. The ultimate usefulness of each sample was determined by PCR with two primer pairs in separate reactions—*leafy* primers amplify a short fragment of 770 nucleotides; 72E18 challenged amplification, for it is a relatively long fragment of 2622 nucleotides. Like primers for *leafy*, primers 72E18 (Fb: GCT AGG GAA AAC AGC TCG TG; Rb: TGG GTT TGG TTT TGG GAT AA) were designed for *F. vesca* cv. 'Pawtuckaway' and are transferable to *F. nubicola*.

Results

The majority of the protocols tested either failed to render appreciable amounts of DNA from mature plant leaf tissue, or yielded plenty of material that was not amenable to further manipulations, such as restriction digestion or PCR (data not shown). However, a variable previously considered minor had an unexpectedly great impact in the retrieval of nucleic acids: further maceration of tissue in extraction buffer. Most of these preparations do not separate DNA from RNA, so quantification is generally a combination of nucleic acids. This is important for two reasons. First, the RNA isolation protocols for strawberry are principally revisions of DNA extraction methods. Those that yield high amounts of RNA also contain proportionate amounts of DNA, and RNA is removed with selective LiCl precipitation. In these preparations RNA and DNA recovery is generally parallel and so quantification of both as "nucleic acids" provides a general measure of DNA recovery. Also, in an attempt to identify an efficacious method, the step of removing RNA, and verifying its removal would limit the number of protocols and experimental conditions that could be tested.

Table 2-1 lists yields from the different DNA isolation protocols described in the Appendix A. Different numbers of treatment replications and amounts of plant tissue were used in the DNA extraction attempts. Therefore, to allow comparison between treatments, values for yield shown in the table are averages of replications, standardized using 1 g of plant tissue as the denominator. Table 2-2 ranks the four methods that had highest nucleic acid returns per g of tissue. Control samples were excluded from the calculation of averages. For example, T85 was a control in protocol 30—tissue was not macerated in buffer. Because the factor in question was the formation of slurry due to maceration, T85 was excluded from the calculation of the average for "slurry" protocols.

Although the "strawberry protocol" permitted extraction of nucleic acids 10 times greater than CTAB-based methods, DNA obtained through the former protocol cannot be digested by restriction enzymes or PCR-amplified by primers for the 18S ribosomal DNA. The DNA remains intractable even after treatment with proteinase K and subsequent dialysis. Similar situations occurred with DNA extracted by CTAB/Tris-borate or guanidine thiocyanate. Only after purifying the guanidine thiocyanate prep utilizing the DNeasy Plant Mini kit, did the DNA become PCR-amplifiable. It is interesting to note that the difference in spectrophotometer readings before and after the purification was minor (treatments 8 *versus* 10), suggesting that the kit may be a viable alternative to oher methods used to purify PCR-recalcitrant DNA.

The 4th highest ranked protocol type in table 2-2 is in fact the only one of the four listed that resulted in tractable DNA. Of the many CTAB protocols that were investigated, the ones that required maceration of plant tissue <u>in buffer</u> cluster together at the top in terms of µg of nucleic acid obtained per gram of tissue (presented later in Figure 2-6).

Components of the "Strawberry Protocol"

Because the strawberry protocol had such high yield relative to the other methods tested, attempts to determine the reason for its superiority were made. The objective was to discover the variable responsible and incorporate it into a protocol that would yield DNA amenable to enzymatic reactions. The factors tested were: i, nucleic acid precipitation by 2-butoxyethanol; ii, boric acid (rather than HCl) used to adjust the pH of Tris for the extraction buffer; iii, second round of extraction from plant tissue after chloroform treatment; iv, dilution of upper phase with Na^+ solution before DNA precipitation.

Treatments T30-T37 (comparing precipitations by isopropanol against 2-butoxyethanol) verified that the latter has a detrimental effect on DNA precipitation. Considering all 4 experimental variables, 65 to 200% more nucleic acids were recovered by isopropanol rather than by 2-butoxyethanol precipitation.

The absolute importance of boric acid to nucleic acid isolation has not been tested, though borate appears to contribute to higher in yields when in combination with other factors. In the extractions using guanidine thiocyanate, borate-containing buffer (T36) had on average 10x higher yield than HCl-containing buffer (T8, T9). However, this increase may be attributed to the different tissue-to-buffer ratios among treatments. A second comparison, this time between

CTAB buffers, strengthens the argument for the contribution of borate: T30 (Tris-borate) *versus* T82 (Tris-HCl), where T30 had a tissue-to-buffer ratio = 16mg/ml and T82 had the ratio that was determined to be optimum for DNA extraction (illustrated in figure 2-2). Perhaps borate was at least partially responsible for T30's 25x greater yield than with T82. When used in substitution to Tris, though, boric acid alone was not able to increase the retrieval of nucleic acids. T38 (1M boric acid, no Tris) was a similar treatment to T30, but the yield was 60x lower.

A second round of extraction from plant tissue increased approximately 50% the DNA recovery relative to a single incubation in extraction buffer. T34 and 35 yielded 60 and 45% of single-extraction treatments T32 and T33, respectively. Although this may be a considerable increase, it is not the sole factor responsible for the dramatic advantage of the strawberry protocol (3 times higher yield than the 2nd highest ranked protocol).

The dilution of the aqueous phase also plays an important role in the recovery of nucleic acids. Observing the results for treatments T24, T27: no dilution; T25, T28: dilution by 2.5 volumes of Na^+ solution (detailed in Appendix A); T26, T29: dilution by 4 volumes, it became apparent that the 2.5 volumes were superior to the other two, in a ratio of 50:125:1 (no dilution : 2.5vol : 4vol).

Optimization of the CTAB Protocol

Protocols containing CTAB in the extraction buffer produced the highest yield of tractable DNA. Therefore, an optimum protocol was devised to further investigate the following factors: leaf tissue state, incubation temperature and duration, tissue-to-buffer ratio, leaf tissue maceration.

Leaf tissue state

DNA was extracted from the same mass of fresh and lyophilized tissues. As expected, yield per gram of sample was generally higher from lyophilized samples. However, this likely is due to the higher number cells that contained in freeze-dried samples in comparison to the same weight of fresh tissue. While yield from T58 was not different from that of T59, increases of 73 and 50% were observed in T13-T16.

There was concern that the lyophilization process might compromise DNA quality. This was addressed by running uncut genomic DNA on agarose gel, and the integrity of all lyophilized samples (T13, T14, T23, and T57) appeared preserved. Therefore, lyophilization may be a good solution for storing material that does not require immediate DNA extraction, but it is not indispensable.

Incubation temperature and duration

Utilizing fresh 'Strawberry Festival' leaf tissue, the effects of temperature and duration of incubation of tissue in extraction buffer were investigated. The treatment that relinquished the most DNA was incubation at 65ºC for 1 hour (figure 2-2), which is the treatment specified in most plant DNA extraction protocols. However, the resultant preparation at this temperature is atypically viscous, complicating mechanical and enzymatic downstream manipulations.

Tissue-to-buffer ratio

Tissue-to-buffer ratios were tested for four protocols (2, 5, 14, 23; ratios and yields shown in table 2-1), and yielded inconsistent results. For protocols 2 and 14, the lower the ratio, the higher the yield, whereas for protocols 5 and 23, the opposite was true. Since all of the ratios (10-200 mg/ml) tested did not use the same protocol, a last DNA extraction experiment was conducted using leaf tissue of 'Strawberry Festival'. Volumes of extraction buffer were kept constant at 5 ml, whereas the treatments were 50, 200, 500, or 1000 mg of fresh tissue. Each treatment included two replicates, and incubation was carried out at 4ºC for 5 min. Samples were treated with RNAse A, DNA was precipitated by isopropanol and resuspended in deionized water. Figure 2-3 illustrates the result of the optimization of the tissue-to-buffer ratio, where the optimum observed was at 40 mg of fresh tissue per milliliter of buffer.

Using the optimum tissue-to-buffer ratio determined in the experiment above (40 mg/ml), the procedure of extracting DNA with incubation at 4ºC for 5 min was tested on ten strawberry cultivars, 2 replicates each. 'Strawberry Festival' was included as a control. DNA recovery was dependent of plant species and cultivars (table 2-3). Plants with rigid leaves, such as *F. chiloensis* and the more *F. chiloensis*-like *F. × ananassa* 'Diamante' had negligible yields. Perhaps solely grinding leaves in liquid nitrogen is not sufficient to break down the cells and expose contents to the extraction buffer solvents.

An attempt to determine the optimum tissue-to-buffer ratio for lyophilized tissue was made utilizing material from a *Fragaria* diploid mapping population. Tissue weights varied from 3 to 14 mg, with average of 6.8 mg and standard deviation of 2 mg. Tissue was macerated in liquid nitrogen and, subsequently, in extraction buffer for approximately 30 s. Grinding in buffer was conducted until the material was the consistency of paste. Incubation was performed at 65ºC for 1 hour. No correlation between amount of tissue processed and DNA recovered was apparent (figure 2-4).

PCR was performed using 1µl of the extracted DNA at variable nucleic acid concentrations (40ng/µl to 4.5µg/µl) and the primer pairs designed for the *Leafy* gene: FvLFYintron2F (5' CAC TGC CAA GGA GCG TGG TG 3') and FvLeafy3' (5' TCA GTA GGG CAG CTG ATG 3'). Due to inability to PCR-amplify 50% of the diploid mapping population samples, an effort was made to monitor for correlations between PCR outcomes and i, nucleic acid concentration in the sample (figure 2-4); ii, tissue-to-buffer ratio during DNA extraction (figure 2-4); and iii, A260/A230 ratios (figure 2-5) that could be indicative of carbohydrate contamination. The absorbance ratios at 260nm and 230nm wavelengths were grouped into seven categories, and the number of samples in each category is indicated in figure 2-5.

No conclusive correlation between success of amplification reaction and any of the three variables cited above could be determined. Although not statistically analyzed, subjective evaluation indicated no need to apply statistical techniques. Surprisingly, there was no pattern suggesting a relationship between template concentration and PCR amplification. This outcome indicates that other factors are contributing to inhibition of the process. In an attempt to dilute a possible polymerase inhibitor, lower tissue-to-buffer ratios were tested. However, no correlation between ratios and PCR outcome was apparent, since all permutations were detected: amplification was observed for both low and high tissue-to-buffer ratios; lack of amplification was also observed for both low and high ratios. Regarding the A260/230, according to Manning (Manning, 1991), the ratio 1.8 indicates the purest nucleic acid sample. From the samples that were classified in this category (47 samples out of 91), 2/3 of them were amenable to amplification. Amplification was also observed for both extreme A260/230 ratios: 0.6 and 6.2. Therefore, the ratio either is a poor estimator of polysaccharide inhibition, or the polymerase inhibition was caused by polyphenols or other indeterminate factors. These trials indicate that there is no simple measure that serves as an indicator of a sample's potential to be used successfully in downstream applications.

Tissue maceration method

The processes of breaking leaf tissue down solely in liquid nitrogen *versus* preliminary pulverization in liquid nitrogen with subsequent grinding in buffer were compared. Formation of slurry by maceration of tissue in buffer not only increased the yield by many fold (table 2-4 A), but also permitted the extraction of allegedly purer DNA, indicated by the lower absorbance at 230nm (figure 2-6). The most prominent absorbance peak at 260nm was observed for samples that were processed at 60ºC and ground in buffer (figure 2-6). Samples macerated this manner and incubated at 4ºC appear to contain many polysaccharide contaminants, as a peak is seen at 230nm. The desired A260:A230 and A260:A280 ratios are equal to 1.80. Samples that were ground in liquid nitrogen only and incubated at 4ºC absorbed more at 230nm than 260nm (ratio = 0.61, table 2-4), indicating that they probably had low content of nucleic acids.

Due to the extraordinary increase in DNA content by the maceration procedure, several treatments combining speed (1/2, full) and duration (5, 15, 30, 60, 120 seconds) of homogenization with a Polytron were investigated. Incubations post-homogenization were carried out at 65ºC for 1hour. The more aggressive the treatment, the higher the amount of DNA obtained (figure 2-7). None of the samples appeared degraded on 1% agarose gel, DNA was digestible by restriction enzymes and amenable to PCR amplification with *Leafy* primers.

Discussion

The profound effect on nucleic acid yield by the aggressive maceration method suggests that the cell wall plays a major role in preventing DNA isolation. This hypothesis is further substantiated by the lower DNA yields observed for genotypes that contain harder leaves with a glossy, conspicuous cuticle, such as *F. chiloensis* and 'Diamante' (table 2-3). However, when the cell wall was removed prior to DNA extraction, DNA extraction from isolated nuclei did not present appreciable yields. It is possible that the isolated nuclei were not pure and therefore the number of nuclei used for DNA extraction was overestimated, explaining the low yield observed.

Guanidine thiocyanate has been used in nucleic acid isolation for a variety of plants. The compound is known to act as protein denaurant by breaking intramolecular hydrogen bonds (Kauzmann, 1954) and, therefore, it causes inhibition of enzyme activity. We hypothesized that the lack of amplification by PCR and digestion by restriction enzymes occurred due to the presence of this chaotropic salt in the DNA preparation. To test this hypothesis, two approaches were adopted to purify the DNA from the guanidine thiocyanate: DNA adsorption to a silica column and dialysis of the DNA preparation. DNA purified by the first method rendered tractable DNA, whereas DNA remained unsuited for enzymatic reaction after dialysis. When isolated by the “strawberry protocol” proposed by Manning, DNA was also intractable even after treatment with proteinase K and dialysis. Therefore, it is possible that the co-purified guanidine thiocyanate or other inhibitors are retained in the dialysis tube. A modification of DNA during the extraction procedure was considered as a possible explanation to enzyme activity inhibition, but the fact that previously intractable DNA purified by a silica column permits amplification by PCR refutes this idea.

The disappearance of an absorbance peak at 230nm when incubation was carried out at higher temperatures (figure 2-6) may be explained by the solubilization of sugars. At lower temperatures, the sugars are present and are not solubilized by the extraction buffer, therefore are carried throughout the remaining steps of the DNA extraction protocol. Their solubilization in the early phase favors production of a purer product.

When considered together it is clear that many variables have no effect on yield. Whereas many protocols alter CTAB concentration, Na concentration, method of precipitation, additional organic extraction and use of affinity matrices, it is clear that concurrent physical and chemical disruption of cells is the most critical parameter in the generation of pure genomic DNA suitable for downstream manipulations.

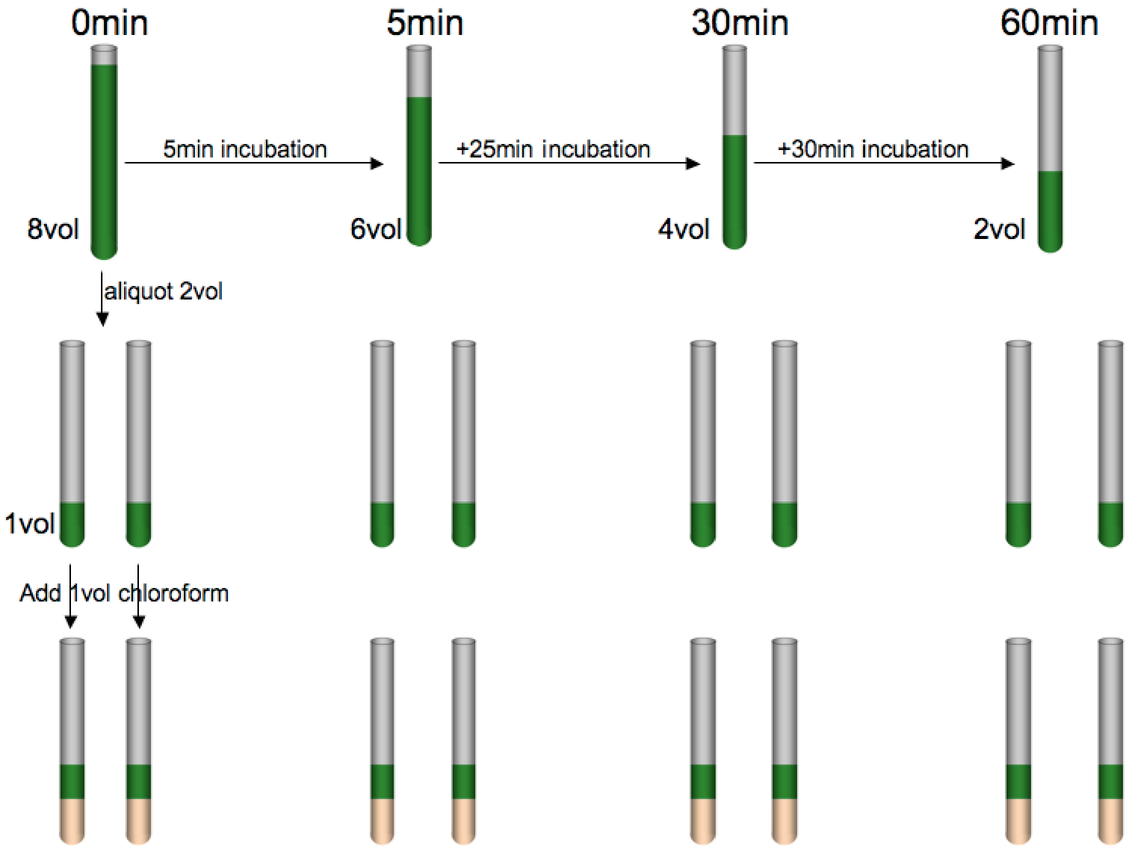

Figure 2-1. Design of incubation temperatures and durations experiment. The scheme illustrated above was followed for each of the incubation temperatures of 4, 20, 42, and 65ºC. Samples for a specific temperature were ground and homogenized together to decrease random variation between time points.

Table 2-1. Nucleic acid yields from isolation protocols. P: Protocol number as listed in Appendix A; T: Treatment number; Status: condition of leaves prior to DNA isolation. F: fresh, L: lyophilized; T/B: tissue-to-buffer ratio (mg of tissue per ml of buffer). n/a: not aplicable; Yield: µg of nucleic acids obtained if 1 g of tissue had been used for DNA isolation

P	T	Status	T/B mg_{tissue} $/ml_{buffer}$	Yield $\mu g_{nucl\ ac}$ $/g_{tissue}$	Brief Description
1	1	F	100	0	Nuclei crude isolation
	2	F	200	0	Nuclei crude isolation
	3	F	400	0	Nuclei crude isolation
2	4	F	100	3	PEG
	5	F	100	1	PEG
	6	F	10	235	PEG
	7	F	10	232	PEG
3	8	F	200	112	Guanidine thiocyanate, newly expanded leaf
	9	F	200	774	Guanidine thiocyanate, unexpanded leaf
	10	F	200	96	T8 cleaned by QIAGEN kit
	11	F	200	11	T8 cleaned by dialysis
4	12	F	1000	35	Guanidine thiocyanate, CTAB consecutively
5	13	L	20	450	Guanidine thiocyanate, CTAB simultaneously
	14	L	200	750	Guanidine thiocyanate, CTAB simultaneously
	15	F	20	260	Guanidine thiocyanate, CTAB simultaneously
	16	F	200	500	Guanidine thiocyanate, CTAB simultaneously
6	17	L	66	0	DNAzol® kit by Molecular Research Center, Inc
	18	L	333	0	DNAzol® kit by Molecular Research Center, Inc
	19	F	66	0	DNAzol® kit by Molecular Research Center, Inc
	20	F	333	0	DNAzol® kit by Molecular Research Center, Inc
7	21	F	70	15	Pine tree *minus* lithium chloride
8	22	F	400	30	Urea
	23	L	50	580	Urea + antioxidants
9	24	F	15	5000	No dilution
	25	F	15	15000	2.5vol dilution
	26	F	15	120	4vol dilution
	27	F	30	8200	No dilution
	28	F	30	18800	2.5vol dilution (not amenable to restriction digestion, even after treatment with proteinase K and dialysis)
	29	F	30	150	4vol dilution
10	30	F	16	5700	Tris-borate, isopropanol
	31	F	16	3130	Tris-borate, 2-butoxyethanol
11	32	F	25	3515	1st extraction, isopropanol
	33	F	25	1190	1st extraction, 2-butoxyethanol
	34	F	25	2135	2nd extraction, isopropanol
	35	F	25	545	2nd extraction, 2-butoxyethanol
12	36	F	16	4300	Guanidine thiocyanate/Tris-borate, isopropanol
	37	F	16	2600	Guanidine thiocyanate/Tris-borate, 2-butoxyethanol
13	38	F	20	90	1M Boric acid, no Tris
14	39	F	33	50	Epicentre kit

Table 2-1. continued

P	T	Status	T/B mg_{tissue} $/ml_{buffer}$	Yield $\mu g_{nucl\ ac}$ $/g_{tissue}$	Brief Description
	40	F	100	15	Epicentre kit
	41	F	333	5	Epicentre kit
15	42	F	635	0	Mo Bio kit
16	43	F	125	8.5	Qiagen DNeasy kit
17	44	F	2.5	150	Silica
	45	F	25	40	Silica
18	46	F	n/a	18	Isolated nuclei, Qiagen DNeasy kit
	47	F	n/a	5	Isolated nuclei, Qiagen DNeasy kit
19	48	F	n/a	12	Isolated nuclei, CTAB
	49	F	n/a	14	Isolated nuclei, CTAB
20	50	F	n/a	3	Isolated nuclei, SDS
	51	F	n/a	1	Isolated nuclei, SDS
21	52	F	n/a	0	Isolated nuclei, guanidine thiocyanate
22	53	F	n/a	0	Isolated nuclei, SDS, TIPS
23	54	F	14	0	Murray and Thompson + solid CTAB, ppt by low ionic strength
	55	F	70	25	Murray and Thompson + solid CTAB, ppt by low ionic strength
	56	L	14	0	Murray and Thompson + solid CTAB, ppt by low ionic strength
	57	L	70	1300	Murray and Thompson + solid CTAB, ppt by low ionic strength
24	58	F	66	45	Murray and Thompson, 6% CTAB, ppt by low ionic strength
	59	L	66	50	Murray and Thompson, 6% CTAB, ppt by low ionic strength
25	60	L	1.6	1250	Murray and Thompson, precipitation by ethanol
	61	L	8	60	Murray and Thompson, precipitation by ethanol
	62	L	16	100	Murray and Thompson, precipitation by ethanol
26	63	F	250	0	Murray and Thompson, 5% CTAB, ppt by isopropanol
	64	F	250	1	Murray and Thompson, 5% CTAB, ppt by isopropanol
27	65	F	100	0	CTAB + SDS
	66	F	100	0	CTAB + SDS
28	67	F	40	16	4ºC, 0min
	68	F	40	80	4ºC, 5min
	69	F	40	98	4ºC, 30min
	70	F	40	69	4ºC, 60min
	71	F	40	34	20ºC, 0min
	72	F	40	28	20ºC, 5min
	73	F	40	95	20ºC, 30min
	74	F	40	142	20ºC, 60min
	75	F	40	28	42ºC, 0min
	76	F	40	41	42ºC, 5min
	77	F	40	34	42ºC, 30min
	78	F	40	57	42ºC, 60min
	79	F	40	41	65ºC, 0min
	80	F	40	25	65ºC, 5min
	81	F	40	76	65ºC, 30min
	82	F	40	211	65ºC, 60min
29	83	F	75	387	Unexpanded leaf
	84	F	75	28	Expanded leaf

Table 2-1. continued

P	T	Status	T/B mg_{tissue} $/ml_{buffer}$	Yield $\mu g_{nucl\ ac}$ $/g_{tissue}$	Brief Description
30	85	F	100	92	Powder
	86	F	100	400	Slurry
31	87	F	50	2476	Slurry, 4ºC
	88	F	50	300	Powder, 4ºC
	89	F	50	3048	Slurry, 60ºC
	90	F	50	700	Powder, 60ºC
32	91	F	40	1400	2% CTAB
	92	F	40	1450	6% CTAB
	93	F	40	665	20% CTAB
33	94	F	40	660	No polytron
	95	F	40	1000	½ speed, 5sec
	96	F	40	940	½ speed, 15sec
	97	F	40	1155	½ speed, 30sec
	98	F	40	1605	½ speed, 60sec
	99	F	40	955	Full speed, 5sec
	100	F	40	975	Full speed, 15sec
	101	F	40	1335	Full speed, 30sec
	102	F	40	1455	Full speed, 60sec
	103	F	40	2245	Full speed, 120sec

Table 2-2. Ranking of 4 best nucleic acid extraction protocols

Average μg nucleic acid/g tissue	Treatments included in average calculation	Protocol #	Protocol type
11,750	T24, T25, T27, T28	9	Strawberry
4,415	T30, T31	10	CTAB with tris/borate
3,450	T36, T37	12	Guanidine thiocyanate
1,232	T83, T84, T86, T87, T89, T91-T103	29-33	CTAB with slurry

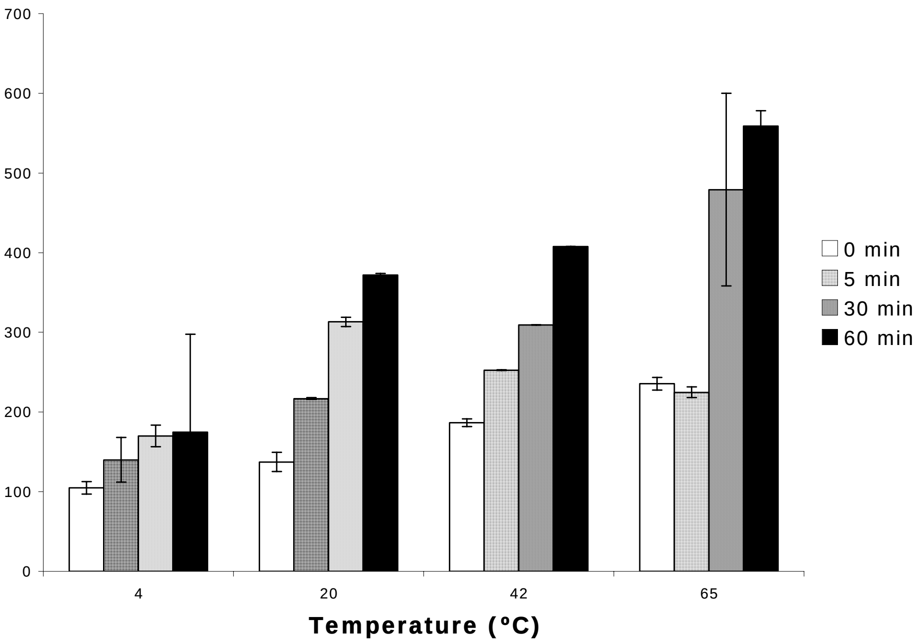

Figure 2-2. Effect of incubation temperature and time on DNA yields. The standard plant DNA extraction procedure of carrying out incubation at 65ºC for 1 hour displayed, as expected, superior yields to other incubation time lengths and temperatures.

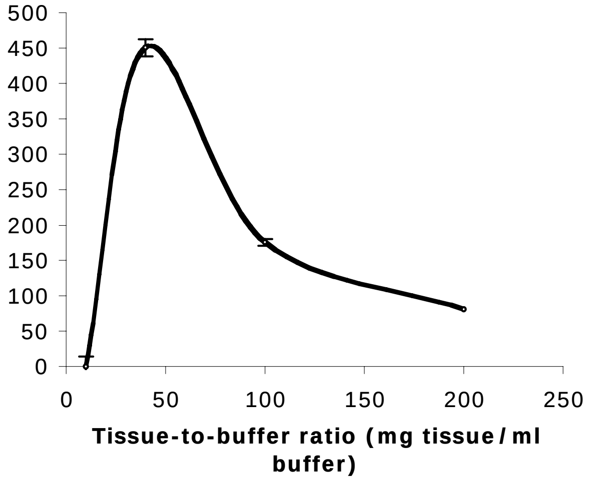

Figure 2-3. Effect of tissue-to-buffer ratios on DNA yields. The optimum ratio for DNA isolation was 40 mg of leaf tissue per ml of extraction buffer. The yield declined rapidly as more tissue was processed by the same volume of buffer.

Table 2-3. DNA yields (µg DNA) from ten strawberry genotypes. Plant tissue incubation with the extraction buffer was carried out at 4ºC for 5min. Averages of 2 replicates, 200mg tissue each, extracted by 5ml buffer.

Genotype	µg DNA/200mg tissue
F. vesca cv Yellow Wonder	127
F. vesca cv Alexandria	59
F. virginiana	54
F. chiloensis	0.85
F. × *ananassa* cv Diamante	0.65
F. × *ananassa* cv Strawberry Festival	50
F. × *ananassa* Laboratory Festival #9	52
F. × *ananassa* cv Camarosa	100
F. × *ananassa* cv Sweet Charlie	64
F. × *ananassa* cv Quinault	55

Table 2-4. Impact of interactions between maceration methods and incubation temperatures on DNA yield and purity. The ratio between absorbance at 260nm and 230nm ($A_{260/230}$) estimate contamination by polysaccharides, whereas the ratio $A_{260/280}$ estimate contamination by proteins. Pure samples have both ratios equal to 1.80.

	Yield µg DNA/50mg tissue		$A_{260/230}$		$A_{260/280}$	
	4ºC	60ºC	4ºC	60ºC	4ºC	60ºC
slurry	31	38	1.02	1.78	1.71	1.91
no slurry	3.8	8.8	0.61	1.46	1.67	1.95

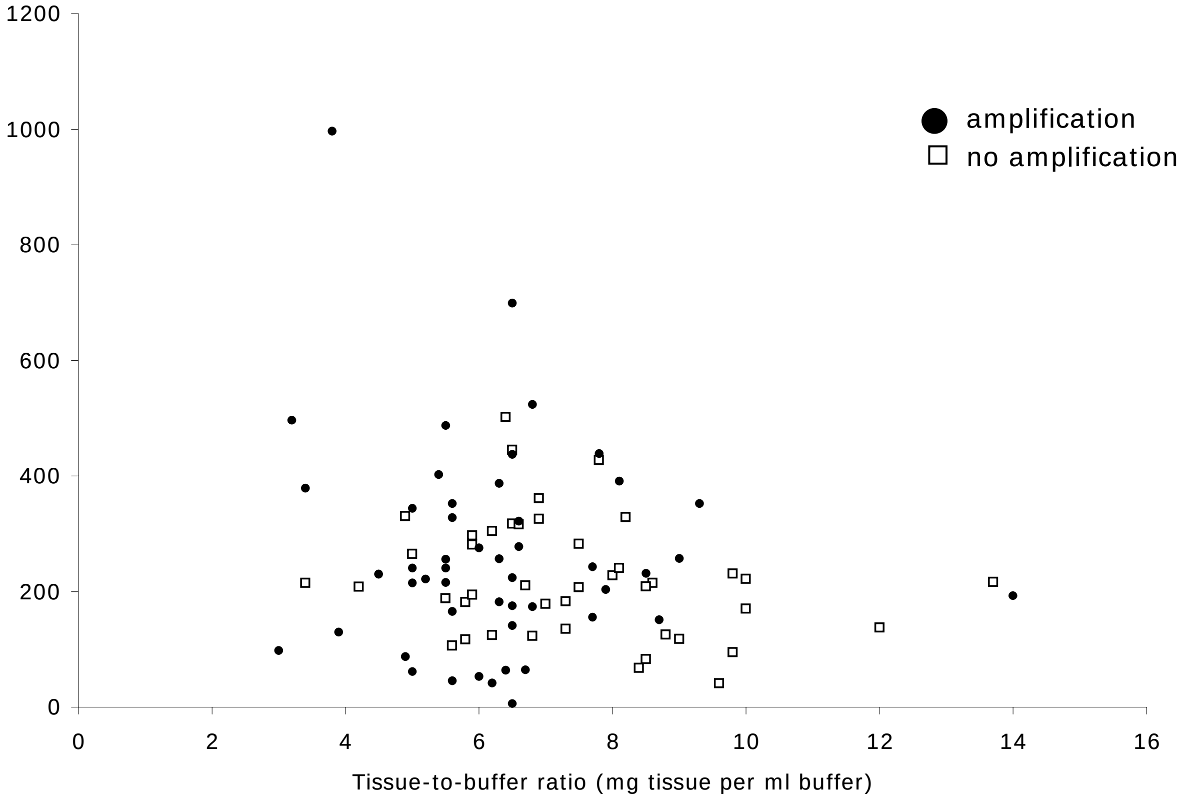

Figure 2-4. Relationships between DNA yield, tissue-to-buffer ratios, and sample amenability to amplification by PCR. DNA was extracted utilizing lyophilized tissue from 94 F2 individuals from a *Fragaria* diploid mapping population. The range of tissue weights was 3-14mg, with average of 6.7mg and standard deviation of 2mg. Because the volume of extraction buffer was kept constant at 1ml, the tissue-to-buffer ratios also represent the amount of tissue (in mg) processed per sample. Correlations between amount of tissue processed, tissue-to-buffer ratio, DNA yield, and PCR outcomes were not apparent.

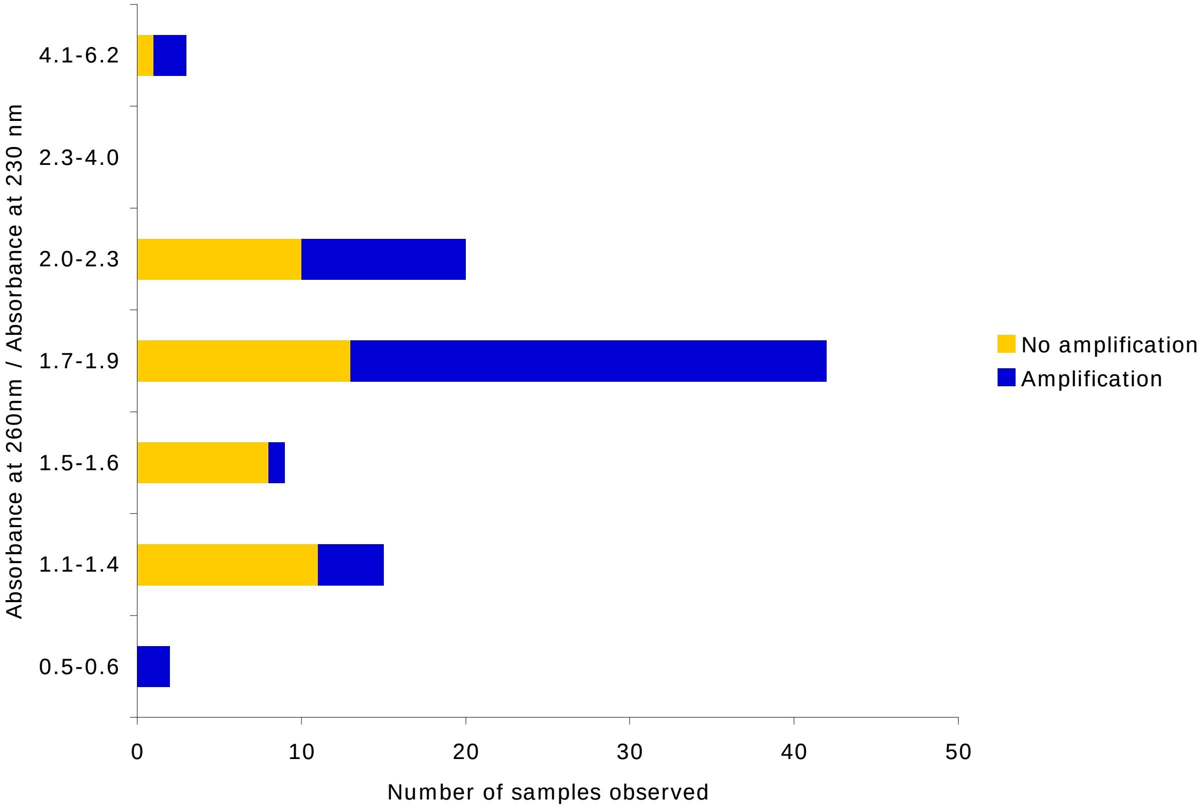

Figure 2-5. DNA contamination by carbohydrate (estimated by the ratio between absorbance at 260nm and 230nm) and its influence on PCR outcome. Absorbance at 230nm and 260nm wavelengths were observed for 94 samples from a genetic linkage mapping population. The A260/230 ratio was calculated for each sample and the ratio data were grouped into 7 categories, varying from 0.5 to 6.2. Most samples presented ratio in the 1.7-1.9 range (1.8 is the optimum for DNA purity from carbohydrates). However, even within the purest DNA category, amplification by PCR was not observed for 1/3 of the samples. Therefore, contamination by carbohydrates may not be considered the sole responsible for the polymerase inhibition.

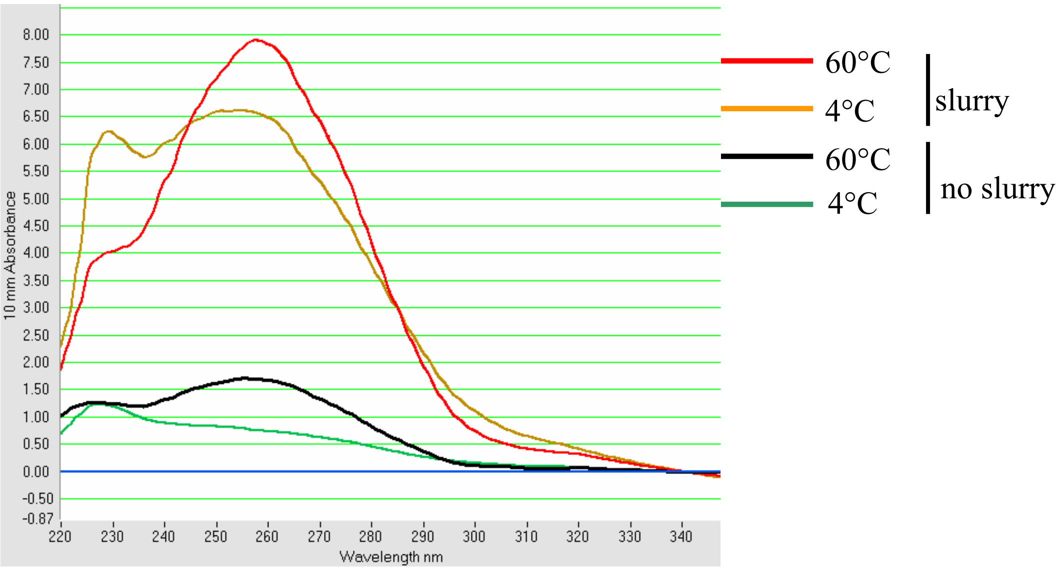

Figure 2-6. Effect of interactions between maceration method and incubation temperature in the absorbance at 220-340nm. The most desirable product from a DNA isolation procedure has a peak at 260nm. A peak at 230nm indicates contamination by polysaccharides. The more aggressive maceration method, combined with higher temperatures, appears to be the best combination of treatments.

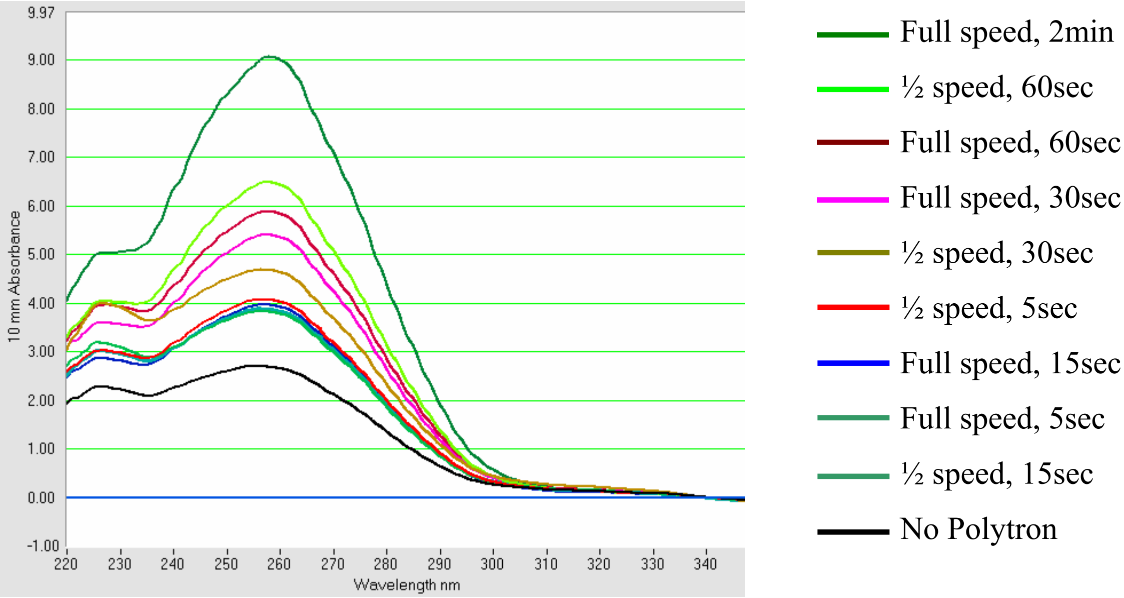

Figure 2-7. The effect of Polytron homogenization on nucleic acid recovery. Leaf tissue was ground in liquid nitrogen and further blended with buffer by utilization of a Polytron. The full uniformization promoted by higher speeds and prolonged durations yielded the best results on DNA isolation.

CHAPTER 3
PRIMARY ANALYSES OF *Fragaria* GENE DISTRIBUTION

Introduction

Although the cultivated strawberry genome is complex and polyploid, its monoploid genome is particularly small and tractable (approximately 200 Mb (Folta and Davis, 2006)). When compared to other rosaceous species, the strawberry genome is exceptionally well suited for rapid elucidation of its sequence, leading to meaningful descriptions of gene distribution and content. Here, small portions of the genome may be sampled and annotated to describe the basis of the *Fragaria* genome. These studies may then be extended to other rosaceous species or utilized in comparative genomics efforts. The goal of the research described in this chapter is to provide a basic description of the first expanses of the *Fragaria* genome. The sequences obtained originate from a fosmid library constructed by Dr. Thomas M. Davis. Individual fosmids were selected by hybridization to genes of interest, and some were randomly selected. These studies provide a primary characterization of the *Fragaria* genome, revealing an understanding of gene content and placement as well as other features of the genome of strawberry.

Genome annotation has been defined as “the process of taking the raw DNA sequence produced by the genome-sequencing projects and adding the layers of analysis and interpretation necessary to extract its biological significance and place it into the context of our understanding of biological processes.” (Stein, 2001) The first challenge to annotate any genomic sequence information is to discriminate between two types of sequences: coding (DNA sequences encoding a protein) and non-coding (DNA is not transcribed into RNA or it is transcribed but not translated into a protein). Regulatory sequences such as promoters and enhancers are examples of non-coding DNA sequences. Other non-coding DNAs are transfer RNA, ribosomal RNA, small RNAs (snoRNAs, microRNAs, siRNAs, piRNAs), and long RNAs (Xist, Evf, Air, CTN, PINK).

The second challenge for annotation is to ascertain or predict gene function, how gene products might interact, and how they are regulated (Salamov and Solovyev, 2000). Gene finding can be accomplished by similarity-based or *ab initio* gene prediction software. Similarity is defined by the NCBI glossary as “the extent to which nucleotide or protein sequences are related.”

Similarity-based algorithms provide information on alternative transcription (Li et al., 2006), translation start sites, and slicing and are more specific than *ab initio.* However, the latter is more sensitive than the former because it does not bias findings based on prior descriptions (Birney et al., 2004).

Similarity-based algorithms like GeneWise (Birney et al., 2004) predict genes by testing putative translation products for similarity to known proteins. A nucleotide comparison against cDNA, to an expressed sequence tag (EST), or a protein database using the Basic Local Alignment Search Tool (BLAST) are also similarity-based gene predictions Non-coding rRNAs are also identified using this approach (Stein, 2001). In contrast, the *ab initio* approach attempts to predict genes from sequence data without prior information on gene characterization. Most gene predictors attempt to define a gene using neural networks (modeled according to the learning process in cognitive systems), rule-based systems (algorithms that use an explicit set of rules to make decisions), or hidden Markov models (HMMs). HMMs are statistical algorithms typically utilized in natural language processing. In gene prediction, they are trained with known gene structures (Stein, 2001; Yandell and Majoros, 2002). A Markov model is a statistical model in which the system being modeled is assumed to be a Markov process, i.e., a stochastic (random) process in which the conditional probability distribution of future states of the process depends on previous states. While in the Markov model one or more states can be directly observed, in the *hidden* Markov model, they cannot. HMMs are popular because they are relatively simple, and efficient methods that exist for training and testing HMMs, these being the Baum-Welch and the Viterbi algorithms, respectively (Mark D. Skowronski, *personal communication*). For a review on HMMs, refer to Rabiner, 1989 (Rabiner, 1989). Examples of *ab initio* HMM gene prediction software are GenScan (Burge and Karlin, 1997), GeneMark (Besemer and Borodovsky, 1999), and FGENESH (Salamov and Solovyev, 2000). When used to annotate the rice genome, FGENESH was more sensitive and more specific than GeneMark and GenScan (Yu et al., 2002).

Plant genomic annotation mechanisms gained favor in the year 2000, shortly after the completion of sequencing of *Arabidopsis thaliana*, a widely used genetic, developmental and physiological model for plants (The Arabidopsis Genome Initiative, 2000), followed by rice in 2002 (Yu et al., 2002). The initial annotation of the *Arabidopsis* genome was submitted by numerous centers, each of them utilizing their own annotation method and terminology. The genome has been re-annotated and classified using Gene Ontology terms as a solution to the cumbersome handling of the information that had resulted from non-centralized annotation (Haas et al., 2005).

Since the completion of the first draft of the rice genome, sequencing of many plants has progressed: high-quality finishing of rice and deep draft coverage of maize, alfalfa (*Medicago truncatula*, the model legume), tomato (*Lycopersicon esculentum*) (National Plant Genomics Initiative, 2002), and black cottonwood (*Populus trichocarpa*) (Tuskan and Difazio S, 2006). Despite the high commercial value of strawberries, there is extensive more nucleotide sequence information for the above-mentioned species than for *Fragaria*. The availability of strawberry nucleotide sequences was so scarce in 2004 that, if one searched for "*Fragaria*" in public databases, only 58 gene sequences were retrieved (Folta and Davis, 2006). In 2007, this number jumped to over 20,000 sequences, of which 50% are Expressed Sequence Tag (EST) sequences. Collaborative work between the laboratories of Drs. Thomas M. Davis (University of New Hampshire), Kevin M. Folta (University of Florida), Jeffrey L. Bennetzen (University of Georgia), and Phillip SanMiguel (Purdue University) have added an additional 50 genomic DNA sequences, constituting slightly less than 2 megabases of genomic information. The sequences are derived from a *Fragaria vesca* 'Pawtuckaway' genomic library and represent 1% of *F. vesca*'s 200Mbp haploid genome (Folta and Davis, 2006). Due to its minute genome size and to the facts that *F. vesca* is the most geographically predominant diploid *Fragaria* species (Folta and Davis, 2006) and it is a plausible ancestor of the cultivated, octoploid strawberry (Ichijima, 1926; Davis and DiMeglio, 2004), this diploid serves as a valuable model for development of molecular markers and comparisons amongst several *Fragaria* species, as well as other genera of the Rosaceae family.

This study aimed to annotate the newly sequenced parcels of the *F. vesca* genome. This represents the first opportunity to explore the gene distribution and the composition of the *Fragaria* genome, which, at 200 Mbp, is comparable to the 157 Mbp (Bennett et al., 2003) genome size of the model plant *A. thaliana*.

Materials and Methods

Dr. Thomas M. Davis at the University of New Hampshire used fosmids (CopyControl^{TM} pCC1FOS^{TM} from Epicentre) as vectors to produce a *F. vesca* genomic library with 8x coverage. The theory is that if the genome was digested into 35kb fragments, approximately 45,000 colonies would be necessary to represent the 200Mbp haploid genome 8 times. Fosmid vectors were developed by Kim *et al.* (Kim et al., 1992) to address undesirable recombination during cloning in multicopy cosmid vectors. Due to the single-copy F-factor replicon, DNA inserted into fosmid vectors underwent a lower rate of rearrangements and deletions than did fragments inserted into cosmids.

In order to annotate the newly available *F. vesca* sequence, a complement of *ab initio* and similarity-based approaches was utilized. Preliminary screening for putative genes was executed by using the gene prediction software FGENESH (accessible at http://www.softberry.com) for each of 26 fosmid insert sequences, using *Medicago* as the gene model. Subsequently, a series of different types of sequence similarity searches were performed using BLAST algorithms (http://www.ncbi.nlm.nih.gov/BLAST/), as illustrated in figure 3-1.

The amino acid sequences from each gene predicted by FGENESH were used as query sequences against the non-redundant protein sequences database for “all organisms” using the BLASTP algorithm. Significant similarities between a query sequence and a sequence in the database, termed “hits”, were indicated by an expectation value (E value) lower than 10^{-15}. (The lower the E value, the more significant is the score because the E value ultimately represents how likely two sequences are of being similar by chance alone.) The threshold of 10^{-15} was defined based on thresholds used in the *Arabidopsis* genome annotation (The Arabidopsis Genome Initiative, 2000), where BLASTP E values < 10^{-20} and 10^{-10} were adopted to identify protein families and functional roles between different organisms, respectively.

The BLASTP results that produced significant hits were used to guide the subsequent BLAST interrogations because they determined which nucleotide fragments should be further analyzed. Though the entire 30-45kb sequence could conceivably be analyzed at once, it is more convenient to do the analysis in sequence parcels. The response to a BLAST submission of sequences larger than 12kb may require protracted time frames and the process may get aborted before the result is retrieved (T. M. Davis, *personal communication*). A second reason to perform searches in parcels is that if two genes are contained in the large query nucleotide sequence and one of them has very high similarity to more than 100 hits, this condition may mask the similarity results to the second gene, appearing as if the second gene was non-coding sequence.

Similarity searches with BLASTX were performed using sequence segments for which BLASTP detected amino acid matches. The translated nucleotide query was delimited to sequence fragments of 8kb whereas the non-redundant (nr) amino acid database against which the *F. vesca* sequences were compared was confined to green plants (green algae and embroyphytes) "Viridiplantae". BLASTX was carried out to determine coding sequence orientation, to assign tentative gene identification and function to the query sequence, and to note the accession and locus tag numbers for the best *Arabidopsis thaliana* orthologs. The *Arabidopsis* loci are sequentially tagged according to their physical position in the genome. Therefore, the tag numbers could be used to assess microcolinearity between *Arabidopsis* and *F. vesca.*

The BLASTN algorithm was utilized in separate searches against the EST and the "nr" nucleotide collection databases. The query sequences originated from fragments for which a gene had been predicted by FGENESH. EST databases searched were delimited to the Rosaceae, in an attempt to detect homologs (sequences that display similarity due to their shared ancestry) and the best *Fragaria, Malus, Prunus, Rubus,* and *Rosa* hits were noted. When no identities were detected within this botanical family, the search was expanded to the Viridiplantae database to detect ESTs that would facilitate detection of genes in the genomic sequence. BLASTN against the "nr" database was executed to detect repetitive elements and non-translated sequences features such as rRNA, tRNA, and was also useful to detect duplications within the query sequence. If two different regions from a single query were similar to single subject from the database, that indicated a duplication in the fosmid insert sequence under investigation.

To address the sensitivity aspect of the FGENESH gene predictor software, a second search utilizing the BLASTN algorithm was carried out. This time, the query sequences were 8-12kb fragments of genomic sequence (regardless of whether or not genes were predicted in that segment), compared against Rosaceae ESTs. The objective was to determine if *Medicago* suffices as a gene model for gene prediction in *Fragaria*.

A survey of the simple sequence repeats (SSRs) present in the newly accessible *F. vesca* sequences was carried out and their location, composition and predominance were noted. The tool used, SSRIT (Temnykh et al., 2001), is available online at the Gramene website: http:// www.gramene.org/ db/searches/ssrtool.

Results

The average fosmid insert fragment size was 35kb and FGENESH predicted 235 genes from the 26 fosmid insert sequences. Of the total number of nucleotides, 42% were predicted to belong to genic sequences. A list of the numbered predicted genes and their corresponding BLASTX results is available in Appendix B. The software specificity was 55%, since out of the 235 genes predicted, 129 had hits in the amino acid database having as threshold $E < 10^{-15}$.

Enzymes related to mobile elements like transposase, integrase, polyprotein, retrotransposon polyprotein, transcriptase, and reverse transcriptase were putatively present ubiquitously: 14 out of 26 fosmids contained at least one of those types of enzymes. In some cases, several of these enzymes were present in tandem, as depicted for fosmid 18A19 in figure

3-2. The second fosmid diagrammed in figure 3-2, fosmid 34D20, contained putative protein-encoding sequences, including inverted repeats of a gene next to a transposase.

Expressed Sequence Tags (ESTs)

ESTs facilitate genome annotation (The Arabidopsis Genome Initiative, 2000) because they are strong evidence that a sequence is transcribed. In the case of *Fragaria,* only a small percentage of protein hits was supported by EST hits (32 of 129), exacerbating the need for more rosaceous ESTs. Three classes of ESTs were identified (figure 3-3): i, those that displayed identity to predicted, putative protein-encoding genomic sequence; ii, those that were FGENESH-predicted genes, but for which there was no protein hit; and, more interestingly, iii, those that were identified spanning DNA sequences for which no ORF was predicted.

Simple Sequence Repeats (SSRs)

Due to their widespread presence, SSRs have been used to construct a linkage map in diploid strawberry (Sargent et al., 2004). Here, SSRs were identified in all fosmid insert sequences, except three: 11D02, 15B13, and 32L07. It is interesting to note that these fosmids putatively contain plastid and RNA genes and belong to the 50% class that did not contain any putative transposon-related enzymes.

A total of 195 SSRs containing at least 5 motif repetitions were identified. Of the nearly 4,000 nucleotides contained in the SSRs, 71% occurred in regions that were predicted to be intergenic. The great majority (92%) of the repeats were dinucleotides (table 3-1). The numbers of times a specific motif was observed are listed in table 3-2.

Discussion

Amplification of repetitive elements, together with polyploidy, are the mechanisms responsible for genome expansion (Bennetzen and Kellogg, 1997). Evolutionary mechanisms for genome downsizing also exist, though they are less well characterized. Bennetzen *et al.*

(Bennetzen et al., 2005) proposed that retrotransposon removal as well as small deletions caused by unequal homologous recombination and illegitimate recombination, lead to genome shrinkage. Grasses like rice, maize, sorghum (Bennetzen et al., 1998), and wheat (Li et al., 2004) are known to have large gene-empty regions and abundant transposons in the intergenic sequences of gene clusters (Barakat et al., 1998). Fosmid insert 38H05 appeared to be one such gene-empty space, since the only similarity detected between its 32kb sequence and the protein database was to polyprotein, which comprised only a small portion of the fosmid sequence.

The pattern of gene distribution was more similar to *Arabidopsis* than to grasses. *Arabidopsis* has been determined to have 15 to 32 Open Reading Frames (ORFs) per 100 kb (Barakat et al., 1998), or 1 gene per 3-6.6kb, whereas rice has one gene per 6.46 kb (Yu et al., 2002) and barley has one gene per 15-20 kb (Keller and Feuillet, 2000). The *Fragaria* average gene distribution was calculated as 1 gene/4kb or 1 gene/9kb, depending on the prediction method used: *ab initio* gene prediction software FGENESH or BLASTX similarity-based approach at $E<10^{-15}$, respectively. In either case, strawberry ranks among the more gene-dense species. Since a portion of the fosmid sequences analyzed arose from non-random, "gene of interest" selections, it is possible that the sample was biased toward genic regions, and that the number of kb containing one gene will increase as more random expanses of the genome are sequenced.

The number of putative genes per fosmid ranged from 6 to 15 (identified by *ab initio*) or from 1 to 11 (according to homology to protein database). The discrepancy between the numbers from the two methods may be attributed to the following possibilities: i, the gene structure used for prediction was from *Medicago*, not *Fragaria.* There is a possibility that the gene structures between these two organisms are distinct enough that a sequence encoding a protein in *Medicago* is not coding in *Fragaria*; ii, the gene prediction is correct, but the putative amino acid sequence is not represented in the protein database because the transcript is not translated (RNA genes in fosmid 15B13, for example), or because the protein has not yet been described; iii, the amino acid sequence is indeed represented in the database, but it is not conserved with *Fragaria*, so the E value threshold chosen as a threshold is too stringent. If a less stringent threshold is used (E value 10^{-10}, rather than 10^{-15}), the number of BLASTX hits increases from 129 to 166 and, therefore, software specificity rises from 55 to 70%.

Half of the ESTs that were identified in genomic regions for which no gene was predicted (figure 3-3) were detected in fosmids that either contained sequence similar to chloroplast DNA (11D02 and 32L07) or to ribosomal RNA (26S in fosmid 15B13). One of the ESTs displayed identity starting in the first nucleotide of the fosmid insert. Perhaps the gene predictor failed to perceive this ORF because the query sequence did not contain transcription initiation signals. The other half of the ESTs that were identified but not predicted was similar to genomic sequences from other species, and the reason why the gene prediction software failed to predict them is not clear. This may suggest some facet of *Fragaria* gene structure that is not recognized by other conditioned algorithms. The detection of putative genes through homology-based similarity search reveals the need to utilize various homology search methods in combination to *ab initio* gene prediction for the optimum genome annotation. This finding is exceedingly important as the genomes of peach and apple will soon be sequenced. Accurate genome annotation will depend on the capacity to adapt current gene prediction methods to these genomes.

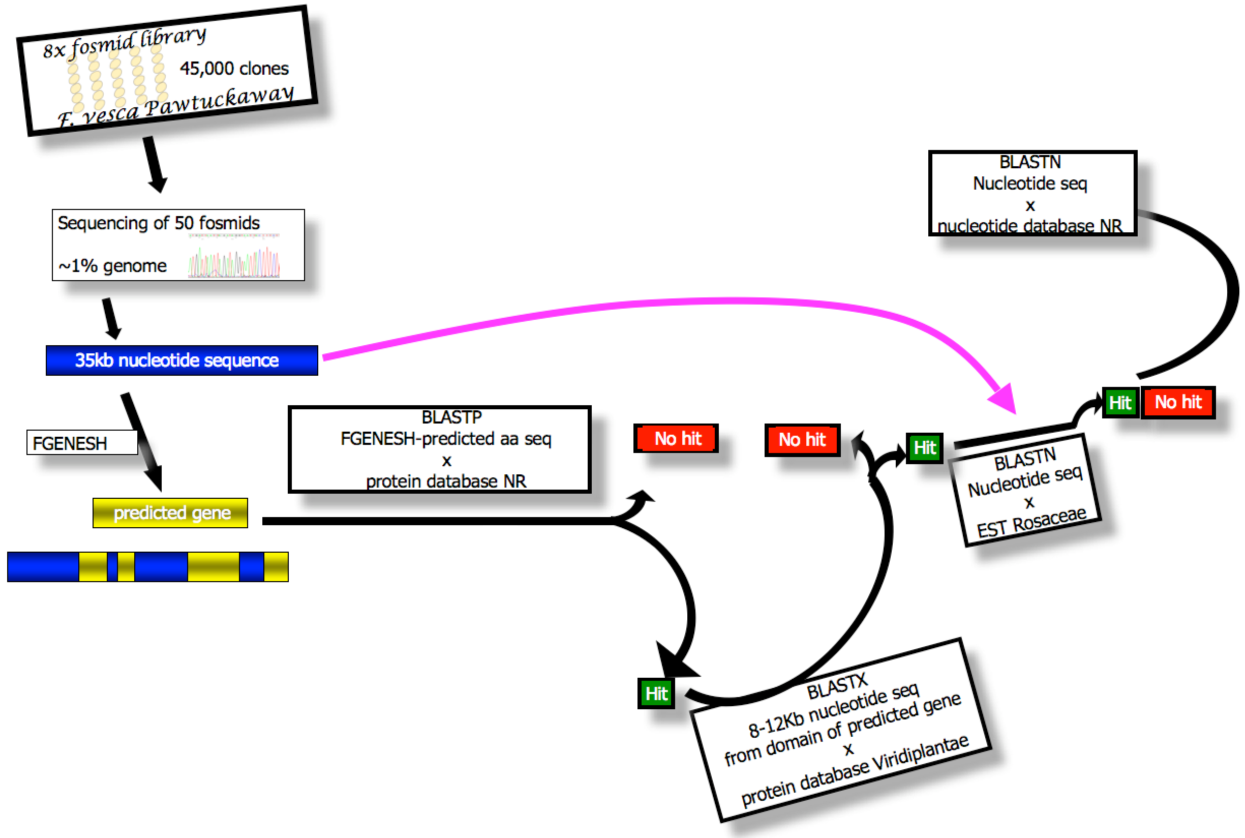

Figure 3-1. Flowchart of genomic DNA sequence annotation scheme. The software FGENESH was used with *Medicago* as the gene model to predict possible gene positions in the genomic sequence. BLASTP algorithm was utilized as preliminary validation FGENESH prediction, whereas BLASTX was used to determine coding sequence orientation and assign tentative gene function. Putative homologs within Rosaceae, conservation amongst various taxonomical families, as well as sequence repeats and duplications were detected by different homology searches utilizing BLASTN. Finally, putative genes that had not been predicted by FGENESH were identified by searching similarities between large fragments of genomic sequence (containing or not FGENESH-predicted genes) and Rosaceae EST.

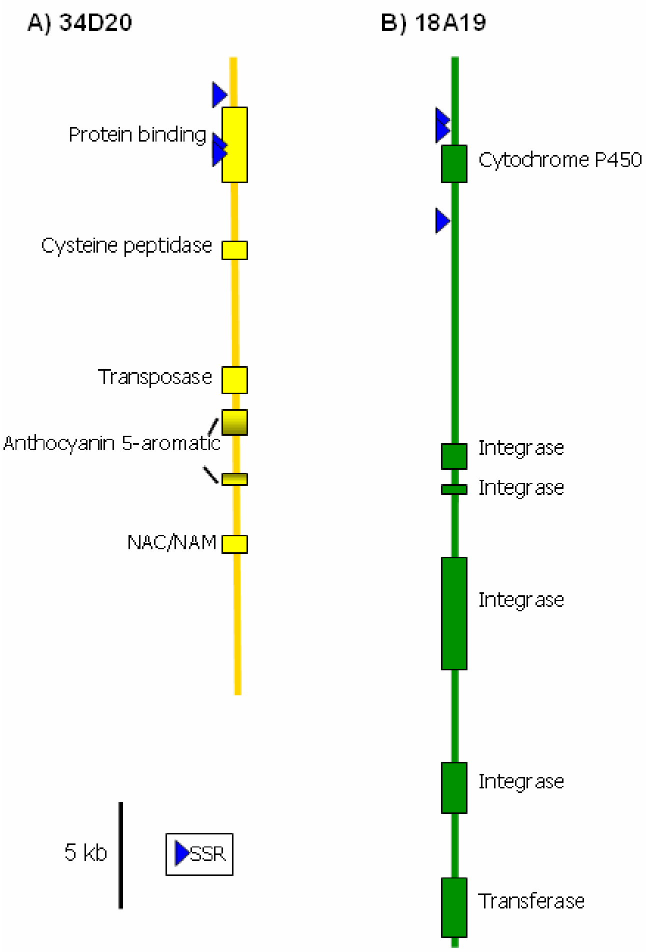

Figure 3-2. Diagram of two fosmid inserts of variable length, with their putative proteins and Simple Sequence Repeats (SSRs). Fosmid 34D20 contained an inverted repeat of an anthocyanin gene next to a transposase, in addition to other protein-encoding genes. Fosmid 18A19 contained mostly transposon-related enzymes, integrase and transferase. SSRs were identified both in genic and inergenic spaces.

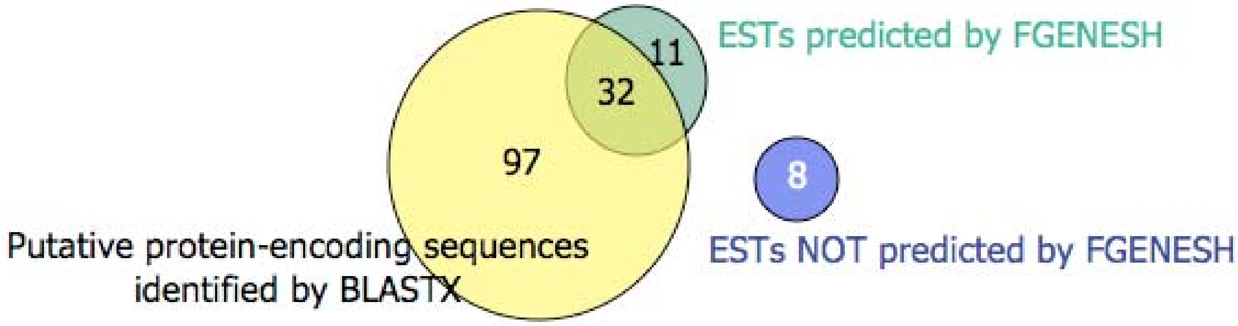

Figure 3-3. EST classes identified by homology searches between large genomic *F. vesca* sequence and Rosaceae ESTs. BLASTX similarity searches were carried out between genomic sequence and the Viridiplantae protein database. A fraction (25%) of the protein matches was validated by BLASTN-detected similarities to the Rosaceae EST database. Other 19 ESTs present no functional information, since no similar amino acid was identified. Of these, a set of 8 ESTs belong to genomic sequence for which there were no genes predicted by FGENESH utilizing *Medicago* as the gene model.

Table 3-1. Number of simple sequence repeats (with a minimum of 5 repeats) observed in *Fragaria vesca* genomic sequence

Motif length	Number of Repeats	Frequency
2 bp	1864	92.1%
3 bp	149	7.4%
4 bp	10	0.5%

Table 3-2. Different types of dinucleotide and trinucleotide repeats observed in *Fragaria vesca* genomic sequence

Motif	Number of Repeats	Frequency
AG/GA/CT/TC	1105	54.9
AT/TA	658	32.7
AC/CA/TG/GT	91	4.5
AGA/GAA/CTT/TCT	79	3.9
CAC/GGT/GTG/TGG	21	1.0
AAC/ACA/GTT/TGT	19	0.9
AAT/TTA	15	0.7
GC/CG	10	0.5
ATG/CAT/TGA//TCA	10	0.5
AGG/CCT	5	0.2

CHAPTER 4

GENE-PAIR HAPLOTYPES: NOVEL MOLECULAR MARKERS FOR INVESTIGATION OF THE *Fragaria* × *ananassa* OCTOPLOID GENOME

Introduction

The cultivated strawberry, *Fragaria* × *ananassa*, contains 8 copies of a set of 7 chromosomes (2n=8x=56). The amount of DNA contained in a single complete strawberry chromosome set is approximately 200 million bases (Nehra et al., 1991; Akiyama et al., 2001; Folta and Davis, 2006), a very small genome size relatively to other angiosperms. There is some controversy as to which angiosperm contains the lowest C-value (or "C*x*-value", terminology proposed by Greilhuber (Greilhuber et al., 2005) to specify the monoploid genome of polyploids), due to different size standards used among various flow cytometry studies. However, likely candidates to the smallest genome are *Arabidopsis thaliana* with 157Mb (Bennett et al., 2003), and perhaps the green strawberry *Fragaria viridis* with 0.108pg (Antonius and Ahokas, 1996). If the formula proposed by Doležel (Doležel et al., 2003) (where 1pg = 978 Mb) is applied, the estimate for *F. viridis* genome size is 105Mb. However, according to the correction proposed by Bennett (Bennett et al., 2003), *F. viridis* current C-value estimate is 206 Mb (Folta and Davis, 2006). Considering that angiosperm C-value varies approximately 1000-fold between species (Bennett and Leitch, 2005), the difference in monoploid genome sizes between *F.* × *ananassa* and *A. thaliana* is negligible.

Although strawberry's small basic genome size makes *Fragaria* species attractive organisms for genomic studies, the process of sorting out segregation in an octoploid background is an extremely complex task, posing a formidable barrier to development of molecular markers and genetic linkage mapping. In a polyploid where reassortment amongst all homologous chromosomes occurs, the number of possible genotypes for a single locus would be $1+\frac{1}{2}\left(C_{\frac{a}{2}}^{a}\right)^{2}$, where a = number of distinct alleles. For an octoploid containing 8 different alleles for a single locus, the number of different combinations would be 2,451. However, this estimate is artificial, since most polyploid plants are considered to be alloploids, and therefore display a degree of fixed, non-segregating heterozygosity (Soltis and Soltis, 2000).

The *F.* × *ananassa* genome structure is not well understood. The first proposed genome structures were derived from cytological analyses of meiotic pairing chromosomes. The genome composition was first described as AABBBBCC (Fedorova, 1946), whilst the model AAA'A'BBB'B'(Bringhurst, 1990) is currently the accepted one. More evidence gathered through the use of molecular markers (Arulsekar et al., 1981; Haymes et al., 1997; Viruel et al., 2002; Ashley et al., 2003) supports the fully diploidized model. In a single study using molecular markers (Lerceteau-Köhler et al., 2003), the authors have observed some polysomic inheritance in a F1 octoploid population. However, the deviations from disomic ratios observed may not be due to polysomic inheritance, as segregation distortions have also been observed in diploid segregant populations (Davis and Yu, 1997; Sargent et al., 2004; Sargent et al., 2006).

The identification of genome-specific polymorphisms may permit the monitoring of segregation of each genome in the complex polyploid background. The "Gene-Pair Haplotype" (GPH) is a tool developed to fingerprint the alleles present in the contributing genomes in the octoploid strawberry. It is defined as a suite of intergenic polymorphisms—Simple Sequence Repeats (SSRs), Single Nucleotide Polymorphisms (SNPs), insertion or deletions (InDels), and changes in restriction sites (RFLP) that present a complex genetic marker for a given locus within the diploid genomes. The types of polymorphisms likely to be detected in a GPH locus and their respective expected location in the genome (within versus between genomes) are summarized in figure 4-1.

GPH markers were also used to investigate polymorphisms in diploid *Fragaria* species, in an attempt to identify genome contributors and trace the diploid ancestors. The genus *Fragaria* contains 23 species of different ploidy numbers. Most of the *Fragaria* species are represented in figure 4-2, with locations based on maps and descriptions published elsewhere (Hancock et al., 2004) (Darrow, 1966) (Staudt, 1973) (Hummer et al., 2005) (Staudt, 2003) (Staudt, 2005). *F. × ananassa* is not included in the figure, as cultivated strawberry is ubiquitous.

According to T. M. Davis, "Lake Baikal marks a major geographical boundary for strawberry distribution. *F. vesca* and *F. viridis* are not found in India, Tibet, China, Japan, or southeast Asia. Likewise, no Asian species grow to the west of Lake Baikal." (http://www.strawberrygenes.com/map.html)

Fragaria species have been cultivated for a long time. The French started transplanting fraise des bois, or the wood strawberry *F. vesca* ("vesca" means "little", in Latin (Fay, 1903)) from the wilderness to gardens in the 1300's, whereas the hexaploid *F. moschata,* the "musky" strawberry, was common in gardens in the 1700's (Darrow, 1966). The modern cultivated strawberry has a very well documented history. It was first cited by Philip Miller in the 1759 edition of the *Gardener's Dictionary* (Darrow, 1966) and it received the name of *F. × ananassa* due to its resemblance to pineapple in odor, taste, and berry shape. In 1765, Duchesne correctly proposed that the new species' parents were *F. virginiana* and *F. chiloensis.* Although both parents are native to America, the spontaneous hybridization occurred in Europe. *F. virginiana* with its rather small fruits was transported overseas in the 1600's. Because of its relatively large fruits, *F. chiloensis* was collected by the Frenchman Amédée François Frézier, during a reconnaissance mission to the Spanish West Indies ordered in 1714 by the king Louis XIV. Disappointingly, no fruits were observed during the first years, probably because, in an attempt to collect only the largest-fruited plants, Fréezier imported only female plants. About 50 years later, the product of the pollination of *F. chiloensis* by *F. virginiana* was observed in Germany, Switzerland, Holland, and the Trianon gardens in France (Darrow, 1966).

F. × *ananassa*'s nuclear genomic content can be traced to fifty-three founding clones (Sjulin and Dale, 1987), whereas as few as seventeen cytoplasm donors are represented in the cultivated strawberry (Dale and Sjulin, 1990). Wild accessions from the octoploid parents have been used relatively recently in strawberry breeding programs for introgression of various characteristics (Hancock, 1999), including day neutrality into California cultivars (Ahmadi et al., 1990).

Athough the identities of the direct ancestor of *F.* × *ananassa* are known, their genome constitutions and evolution are not. The present research investigated polymorphisms in the intergenic regions of diploid species, as well as the cultivated octoploid to attempt to trace ancestry and make inferences about the octoploid genome mode of inheritance.

Materials and Methods

Before the commencement of this study in the year 2004, virtually no *Fragaria* genomic sequence was available. Therefore, it was necessary to develop a means to capture useful sequences for analysis. Two different approaches were adopted: i, inference of gene adjacency by putative micro-colinearity between *F.* × *ananassa* and *Arabidopsis thaliana*; ii, construction and annotation of a *F. vesca* genomic library (discussed in Chapter 3).

Potential micro-colinearity was detected using the approach described in figure 4-3. This approach was possible because the genome of *Arabidopsis* has been completely sequenced and the genes were numbered in such fashion that their locus tags indicate their position on the chromosomes. The hypothesis was that if two genes were adjacent in *Arabidopsis*, they would also be adjacent in *Fragaria*. Similarity between *F.* × *ananassa* ESTs and *A. thaliana* transcripts was tested using the FASTA software available at The *Arabidopsis* Information Resource (TAIR)'s website (http://arabidopsis.org/cgi-bin/fasta/nph-TAIRfasta.pl) and the best match for *Arabidopsis* was recorded. The sequences of each of the *Arabidopsis* genes adjacent to the *Arabidopsis* matches were retrieved from the Salk Institute Genomic Analysis Laboratory (SIGnaAL) "T-DNA Express" *Arabidopsis* Gene Mapping Tool website (http://signal.salk.edu/cgi-bin/tdnaexpress). The next step was to detect *Fragaria* sequences that were similar to each of the *Arabidopsis* gene sequences retrieved. The Basic Local Alignment Search Tool (BLAST) was used to search the *Fragaria* translated nucleotide database using the *Arabidopsis* translated nucleotide query, since TBLASTX is the most sensitive algorithm to detect sequence similarities. Results with an E-value $< 10^{-4}$ were considered positive hits and primers were designed for the putative gene pair to amplify the presumably intergenic sequence flanked by the conserved *Fragaria* and *Arabidopsis* primers. In addition, forward and reverse primers (table 4-1) were designed to amplify at least 100 bp of the EST. This allowed validation that the amplification sequenced was specific to the target regions.

A second approach was adopted to increase the micro-colinearity detection level. Two-hundred and fifty *F.* × *ananassa* EST sequences were randomly selected for similarity searches against *Arabidopsis* utilizing FASTA and all (rather than only the best match) of the *Arabidopsis* sequences that had a similarity E-value $< 10^{-4}$ were considered for further analysis. The loci tags were recorded on two separate tables, one table keeping the correspondence between *F.* × *ananassa* and *Arabidopsis* similar sequences, and the other table had the *Arabidopsis* loci tag sorted in crescent order. When the difference between two consecutive *Arabidopsis* loci tag numbers was equal to or lower than 10, a putative gene pair was detected and the *F.* × *ananassa* EST sequences were retrieved from the non-sorted table.

In addition to the amplification of unknown regions, sequences were gathered by sample sequencing genomic DNA. Both random and targeted sequences were studied in a *F. vesca* fosmid library. The annotation scheme is described in Chapter 3 of this dissertation. Forty combinations of PCR primer pairs were tested to amplify 18 loci, since different primer combinations were required to amplify some of the loci. The primer pairs generated for the putative intergenic regions are listed in table 5-1 of Chapter 5.

Following determination of location and design of PCR primer pairs, PCR was carried out to amplify 28 loci, of which 10 gene pairs (listed in table 4-1) were inferred by the *F.* × *ananassa*/*Arabidopsis* micro-colinearity approach and 18 gene pairs (listed in table 5-1) were inferred from gene prediction from *F. vesca* 'Pawtuckaway' genomic sequence. The optimizations of PCR conditions were carried out utilizing as template DNA from the species for which primers had been designed: *F.* × *ananassa* and *F. vesca* 'Pawtuckaway' for micro-colinearity- and genomic-DNA-based approaches, respectively. Once optimum conditions were determined, the reaction was carried out for seven *Fragaria* species, which included the respective control species: *F.* × *ananassa* 'Strawberry Festival', *F. vesca* 'Pawtuckaway', FRA341 *F. viridis*, FRA377 *F. iinumae,* FRA520 *F. nubicola*, FRA1318 *F. nilgerrensis*, and FME *F. mandshurica*.

The PCR products were cloned using the plasmid cloning vectors pJET1 (GeneJet™ PCR cloning kit by Fermentas Life Sciences) or pCR®2.1-TOPO® (Invitrogen Life Technologies™). The ligation reaction was carried out according to manufacturer's directions and 1µl of the ligation reaction was used to transform 50µl of competent cells. The chemically competent *Escherichia coli* bacterial cells (Invitrogen One Shot® TOP10) were purchased with the TOPO cloning kit whereas XL1-Blue competent cells (Bullock et al., 1987) were prepared in the laboratory using and the rubidium chloride method (Hanahan, 1985). The putative recombinant plasmids and competent cells were gently mixed, iced for 30min, heat-shocked at 42ºC for 2min, and immediately iced again for at least 5min. XL1-Blue cells with no vector were included in each transformation round as a negative control. When large fragments were cloned, a separate treatment with a smaller fragment for which transformation had been successful before was included as a positive control. Two-hundred µl of Luria-Bertani (LB) broth (10g tryptone, 5g yeast extract, 10g NaCl, per litter of deionized water) were added to the transformed cell and were incubated in a shaker for 1 hour at 37ºC, with agitation of 220rpm, after which 100µl of cells were spread onto LB-agar plates containing 50µg ampicillin/ml medium. The TOPO vector has the ß-galactosidase reporter gene. Therefore, when this vector was used, an overlay of 50µl of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (X-gal) at 20mg/ml dissolved in N-N'-dimethyl-formamide and 10µl of the filter-sterilized inducer isopropylthiogalactoside (IPTG) at 1M were added to the LB-agar plates before the transformed cells were plated. IPTG and X-Gal were not added to the LB plates when the pJET1 vector was used. This vector contains a gene for a restriction endonuclease in the cloning site. If disrupted by an insert, the lethal endonuclease is not expressed and the transformants are able to propagate. After the cells were plated, they were incubated at 37ºC overnight and single colonies were selected for screening for transformants—white colonies for TOPO and, supposedly, any colony for pJET1. The screening procedure was carried out by setting up individual PCR reactions for each colony using annealing temperature of 55ºC and primers specific for the vector (pJET1F: 5' GCC TGA ACA CCA TAT CCA TCC 3', pJET1R: 5' GCA GCT GAG AAT ATT GTA GGA GAT C 3'; TOPO, M13F: 5' GTA AAA CGA CGG CCA GTG AAT TGT A 3'; M13R: 5' CAG GAA ACA GCT ATG ACC ATG ATT AC 3'). Approximately 10 colonies were initially selected from each plate with transformants containing PCR products from diploids. Because several different alleles were sought for the octoploid, 30 colonies were selected from the plates that had transformants with inserts amplified from 'Strawberry Festival'. The tested colonies were streaked on a separate LB/ampicillin plate during the set up of the colony PCR reactions. The PCR products were resolved in 0.8% agarose gel with 1x TAE buffer. PCR-confirmed transformants were grown in 3ml LB broth containing 50µg amplicillin/ml for approximately 4 hours at 37ºC, with agitation at 220 rpm. Plasmids were extracted from 1.5ml culture by the alkaline lysis method, followed by 24:1 chloroform extraction. Isolated plasmids were resuspended in 50µl of deionized water and 5 µl were digested with 1 unit of restriction enzymes: EcoRI or XbaI/XhoI for amplicons ligated to TOPO or pJET1, respectively. Transformants carrying distinct alleles were detected by digestion with EcoRI. The digested bands were resolved in 2% Metaphor/1xTBE or 2% agarose/1xTAE. Clones with similar restriction patterns were grouped into different classes and a representative clone of each class was sent to DNA sequencing facilities. A list of primers generated for sequencing reactions can be found in Appendix C, whereas the sequences generated are included in Appendix D.

Sequences obtained were analyzed for conservation between diploid and octoploid alleles. Alignments were performed using the global alignment tool ClustalW available at the European Bioinformatics Institute's website at http://www.ebi.ac.uk/clustalw/. Except for the penalty for gap extension, which was set at 0.05 instead of the default 6.66, all other penalty settings were the default ones: gap open: 15; end gap: -1; gap distance: 4.

Results

Considering the number of *Fragaria* ESTs available at the time this study was initiated (approximately 1,500 ESTs), and the estimated 26,000 genes in the *Arabidopsis* genome (Sterck et al., 2007), if micro-colinearity indeed existed, the chance that two adjacent genes would be detected in the pool of 1,500 was approximately 10% as calculated by $1{,}500\times\left(\frac{1}{25{,}999}+\frac{1}{25{,}998}\right)$. Therefore, the amplification of 2 loci out of the 10 investigated (table 4-2) may be regarded as fairly successful. A third set of primers (GPH4) permitted amplification, though after fragment cloning and sequencing, amplification was determined to be unspecific. There was no similarity between sequences obtained and the 770bp from the template sequences for primer design.

The gene prediction from *F. vesca* genomic sequence enabled detection of 18 potential gene pairs. Of those, primers designed for 11 loci rendered amplification of at least the positive control DNA template of *F. vesca*.

Table 4-2 summarizes the results of PCR amplification using primers designed through both gene-pair detection approaches, as well as results on cloning amplicons and sequencing inserts. The “clone #” in the table is in most cases the PCR reaction number, followed by the colony that was determined to be a transformant by PCR and/or restriction enzyme digestion.

The hypothesis that a fingerprint for each allele belonging to the octoploid ‘Strawberry Festival’ would correspond to alleles from different diploids could be tested by GPHs 5, 23, 10, 27F10, 34D20, and 72E18. The full alignments for every GPH characterized in this dissertation can be found in appendix E.

Data of Individual Loci

GPH5

GPH5 was detected by the micro-colinearity search approach. The adjacent genes in *Arabidopsis* were At3g07320 (E value of $9\text{x}10^{-21}$, encoding a glycosyl hydrolase family 17 protein) and At3g07330 (E value of $2\text{x}10^{-60}$, encoding a glycosyl transferase family 2 protein). GPH5 is a particularly interesting locus, since amplification was observed for all diploids and 2 alleles of the octoploid were detected by *EcoRI* digestion. The following polymorphisms were identified in the 2.8kb analyzed: short indels of 4-12bp (9 bp insertion in *F. vesca*; 12 bp deletion in *F. iinumae*, shown in figure 4-4), 180 SNPs, of which 125 are ambiguous (may be sequencing or polymerase errors) and 55 likely true SNPs, because the base change occurs in more than a single clone. Most of the likely SNPs delineate the octoploid clones from the diploid ones. It is interesting to note that the octoploid alleles are grouped separately from diploid alleles not only for their SNPs, but also for small indels. Two SSR motifs were identified (AAG and AT), with 4 repeats each, for every clone. Therefore no polymorphism in the number of repeats was detected.

GPH23

At1g23740 (oxidoreductase, zinc-binding dehydrogenase family protein) and At1g23750 (DNA-binding protein) were similar to *F.× ananassa* with E values of $3x10^{-64}$ and $2x10^{-57}$, respectively.

Only *F. mandshurica*, *F. iinumae*, and *F.× ananassa* were amplified by the primers designed for this region. Larger deletions than those observed for other loci investigated, and different alleles from the diploids were observed for GPH23. Figure 4-5 illustrates the polymorphisms detected. After preliminary sequence alignment, the putative SNPs were verified by observation of unambiguous peaks in the chromatograms. Therefore, for this locus, a SNP is only an artifact if it was introduced during amplification by the polymerase. (CTC)4 SSRs were detected and occurred in equal number of repeats for every clone, in the same position when aligned. The implications of the polymorphisms are discussed below.

GPH10

Primers GPH10A and GPH10C were utilized to amplify a 4.4kb fragment from the octoploid ‘Strawberry Festival’. Four categories of polymorphic clones were detected by *EcoRI* restriction digestion (figure 4-6) and sequence was obtained for the full clones (figure 4-7). The primers flaking the most polymorphic region (10PPR1 and 10AB#22) were utilized to amplify that region from all six diploids included in this study. A cladogram based on this polymorphic region is shown in figure 4-9.

72E18

72E18 was the only GPH sequenced that presented polymorphism in the number of repeats in SSRs.

Estimations of Relatedness from Sequence Variation

Cladograms are branching diagrams assumed to be an estimate of a phylogeny where the branches are of equal length. Therefore, cladograms show common ancestry, but do not indicate the amount of evolutionary "time" separating taxa (information from the http://www.ebi.ac.uk/ website). In this study the use of cladograms generated from multiple sequence alignments provide an outstanding means to gauge the relatedness between strawberry genomes. When compared against each other, the use of cladograms depicts the relative divergence between similar sequences, and thus is a useful estimate of SNP frequency between the alleles in *F.* × *ananassa* and the putative diploid subgenome donors. The following cladograms derive from all GPH that contained at least one allele representing the octoploid strawberry compared to all cases where products could be amplified from diploids. The results indicate that octoploid alleles cluster together, as do diploid alleles. The most related diploid to octoploid alleles is consistently *F. iinumae,* and surprisingly, alleles closely matching *F. vesca* were not detected for any of these GPH loci.

Relatedness may also be assessed by studying the order of insertion-deletions and SSRs. Presumably, a set of similar indels or SSRs may be conserved between the diploid subgenome donors and the modern cultivated octoploid. The presence and order of these features provides evidence of relatedness. Table 4-3 represents the length and position of indels and SSRs identified in the sequenced clones. In this table, indels and SSRs are presented as variable “features” in genomic sequence as it is parsed from 5’ to 3’. With this method the size and position can be best described, presenting evidence of relatedness. In this table the variable features present in all genomes are revealed. When two or more values in the same column are shown, this represents indels present in the same region of a given locus, as the corresponding genotype deviates from a consensus sequence compiled from multiple sequence alignment of all sequences. A blank box indicates agreement with consensus in a given region. The corresponding genotype does not deviate from consensus. The sequence of the clones listed in the table is conserved with the sequence as they appear in the cladograms of figure 4-9 to facilitate the perception of relatedness.

Discussion

Synapsis between *F. vesca* and *F. virginiana* chromosomes has been shown to occur (Ichijima, 1926). This is regarded as the first evidence that *F. vesca* is a likely genome donor to *F.* × *ananassa*, since *F. virginiana* is the pollinating parent to *F.* × *ananassa*. Another study published a year later showed that the crosses between *F. vesca* × *F. chiloensis* and *F. vesca* × *F. virginiana* produced sterile hybrids (Mangelsdorf and East, 1927). The occurrence of natural hybrids between *F. chiloensis* and *F. vesca* (Bringhurst and Gill, 1970), the geographical predominance of *F. vesca*, and a recent study on chloroplast DNA showing that the *F. vesca* is closely related to *F.* × *ananassa* (Potter et al., 2000), support the hypothesis that *F. vesca* is a contributor to the genome of octoploid strawberries. In this study large intergenic regions were sequenced from a series of octoploid and diploid alleles to assess the relatedness between the cultivated strawberry and potential subgenome donors. Two central methods were used to detect relatedness, both based on multiple sequence alignments. The first used cladograms to display consolidation of single nucleotide polymorphisms (Table 4-2). The second method was as assessment of multinucleotide polymorphisms, detected as indels or SSRs that varied between accessions and a consensus sequence. The use of these complementary methods provides two levels of resolution that describe relatedness between alleles.

Contrary to the expected, however, data from five characterized loci and the inability to PCR-amplify *F. vesca* using primers designed for *F. × ananassa* (GPH23), display *F. vesca* 'Pawtuckaway' as the most unrelated sequence to any of the sequenced octoploid alleles. From the few loci studied, it does not appear that *F. vesca* is a more likely A-genome donor than any of the other diploids studied. This surprising finding contrasts directly with cytological evidence and suggests that *F. vesca* may not be a contributor to at least the 'Strawberry Festival' cultivar.

F. iinumae, on the other hand, was confirmed as one of the most distinct diploid. Table 4-3 shows that *F. viridis* and *F. iinumae* had the most dramatic changes in relationship to the other four diploids concerning size of their indels. *F. viridis* displays large indels: 44, 500, and 800bp in loci 11D02, 27F10, and 32L07, respectively. None of the deletions, however, corresponded to any of the *F. × ananassa* alleles sequenced. In the case of 32L07, no octoploid allele was sequenced because PCR amplification could not be detected for any of the following octoploids: 'Strawberry Festival', 'Carmine', 'Diamante', 'Rosa Linda', and 'Sweet Charlie'.

F. iinumae has a deletion greater than 500bp in the fragment 10PPR1AB22 of GPH10, five indels of approximately 30 bp (three in 11D02, and two in 34D20), and one of approximately 50 bp in 72E18. The indels in 34D20 and 72E18 from *F. iinumae* coincided with *F. × ananassa*, suggesting that *F. iinumae* is indeed a genome donor to the diploid. The cladograms from figure 4-9 suggest that in every locus studied, *F. iinumae* was the most similar diploid haplotype to the octoploid alleles. Phylogenetic analysis of the intron-containing region of the Adh gene of 20 *Fragaria* species identified two major clades, and pointed to *F. iinumae* as the best B clade candidate for Adh allele donor to octoploids (Davis and DiMeglio, 2004).

The data identified here provide further evidence to support the hypothesis that *F. iinumae* is a subgenome donor to the modern octoploid. In all comparisons herein where octoploid sequence was obtained, the octoploid related more closely to the *F. iinumae* haplotype. Thus, one conclusion that can be made is that *F. × ananassa* cv. Strawberry Festival contains clear evidence of the *F. iinumae* characters within its subgenome composition. But what about the A genome? The B genome donor has been disputed, but almost 100 years of evidence implicates *F. vesca* as an A-genome donor. In this data set, little evidence of the A-genome exists. There are several ways to reconcile this discrepancy, although all of them are speculative. The one important caveat is that 'Strawberry Festival' is only one octoploid accession and was used almost exclusively as the octoploid representative. 'Strawberry Festival' has a broad east-coast, west-coast lineage, so in many ways it is an excellent representative for this study. It is possible, albeit unlikely, that the allelic content of 'Strawberry Festival' is skewed to the B-genome *F. iinumae* components and somehow the A-genome is not being detected. This is surprising because the primers that detect the B genome variants were derived from the A genome donor. One alternative explanation is that perhaps the A genome underwent extensive modification, such as expansion, therefore preventing amplification of octoploid sequences by PCR. Alternatively, these regions could have been deleted from the octoploid genome, as the octoploid genome is smaller than four diploid genomes, indicating a loss of genetic material (Folta and Davis, 2006). A final explanation is that not all diploid species, including many *F. vesca* accessions, were tested, so the A genome may be represented by a genotype not tested in this study. There is no simple answer, and this finding may indicate that some higher-order mechanism is at work to limit the presence of subgenome sequence in the polyploid. Polysomic inheritance has been documented (Lerceteau-Köhler et al., 2003). If polysomic inheritance led to a trait of interest early on, it may have been selected as beneficial in breeding populations and fixed in subsequent lines.

Another unlikely explanation is that changes in *F. iinumae* paralled those in *F.* × *ananassa* in two separate and unrelated instances. Probability suggests that this cannot be the case, yet it remains a formal possibility, especially if the changes induced result in regulatory alterations that affect gene expression, biological function and possibly selection. It is also possible that cultivation and selection have important consequences in skewing subgenome representation. It has been demonstrated that *F. iinumae* is a robust plant, with more vigorous growth than *F. vesca* (Sargent et al., 2004). These characters may have lent themselves to the wild octoploids and were attractive to potential early breeders. These alleles may dominate certain selections, like 'Strawberry Festival'. Other cultivars need to be tested to assess allelic composition to further query this hypothesis.

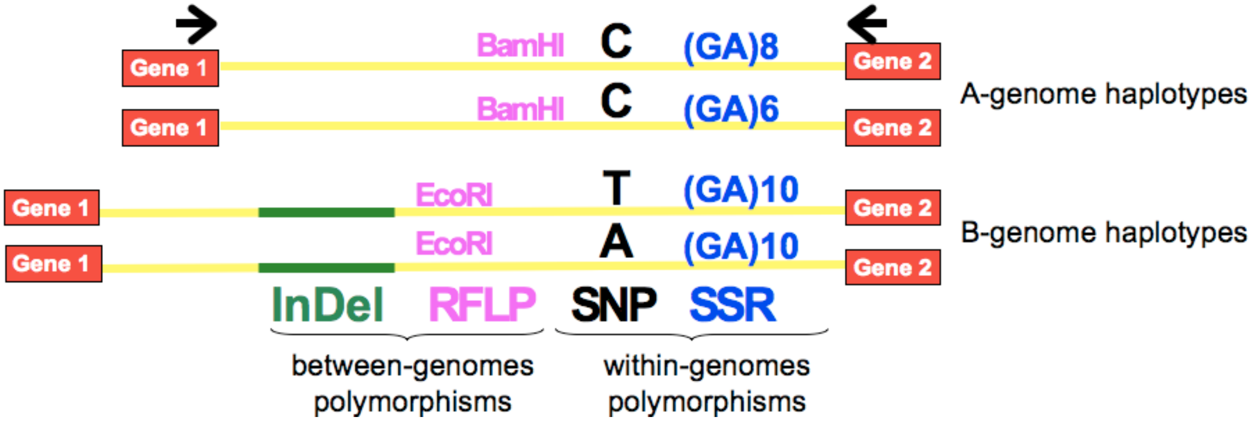

Figure 4-1. An idealized GPH locus. Arrows represent primers designed to amplify the intergenic spaces of a GPH. The combination of polymorphisms within (SSR, SNP) and between subgenomes (InDels, change in restriction sites) define each haplotype.

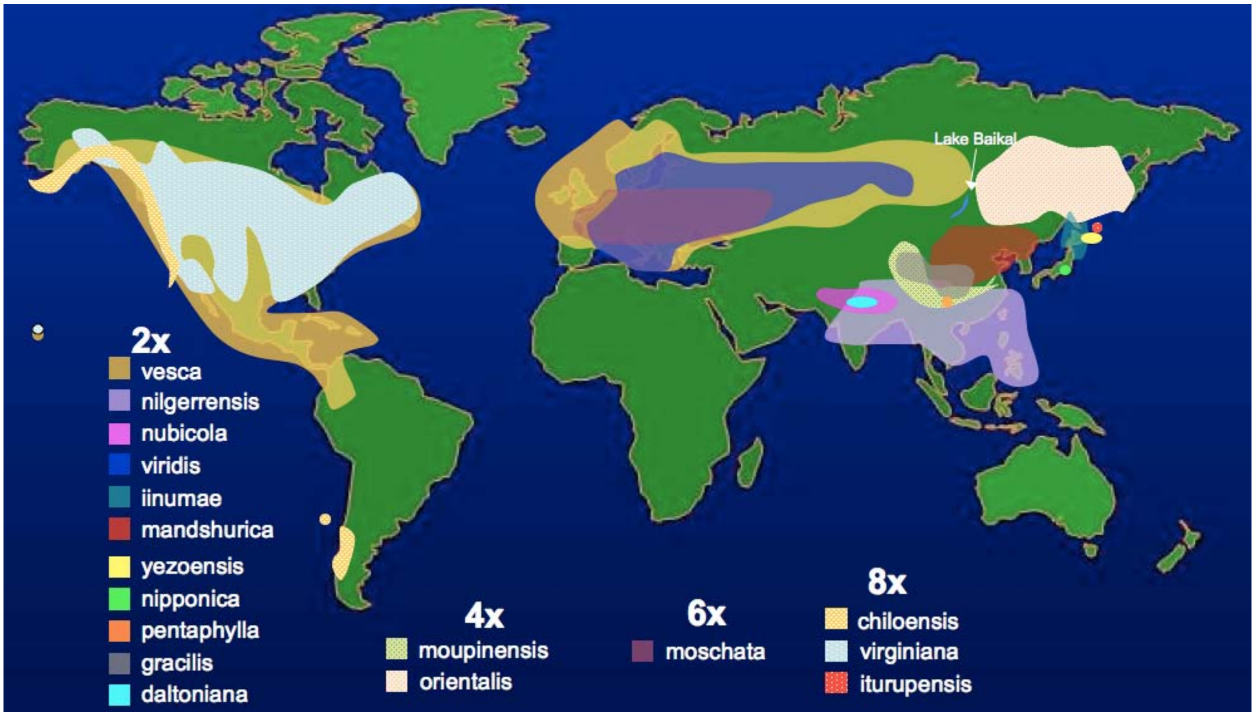

Figure 4-2. *Fragaria species* and their geographical locations

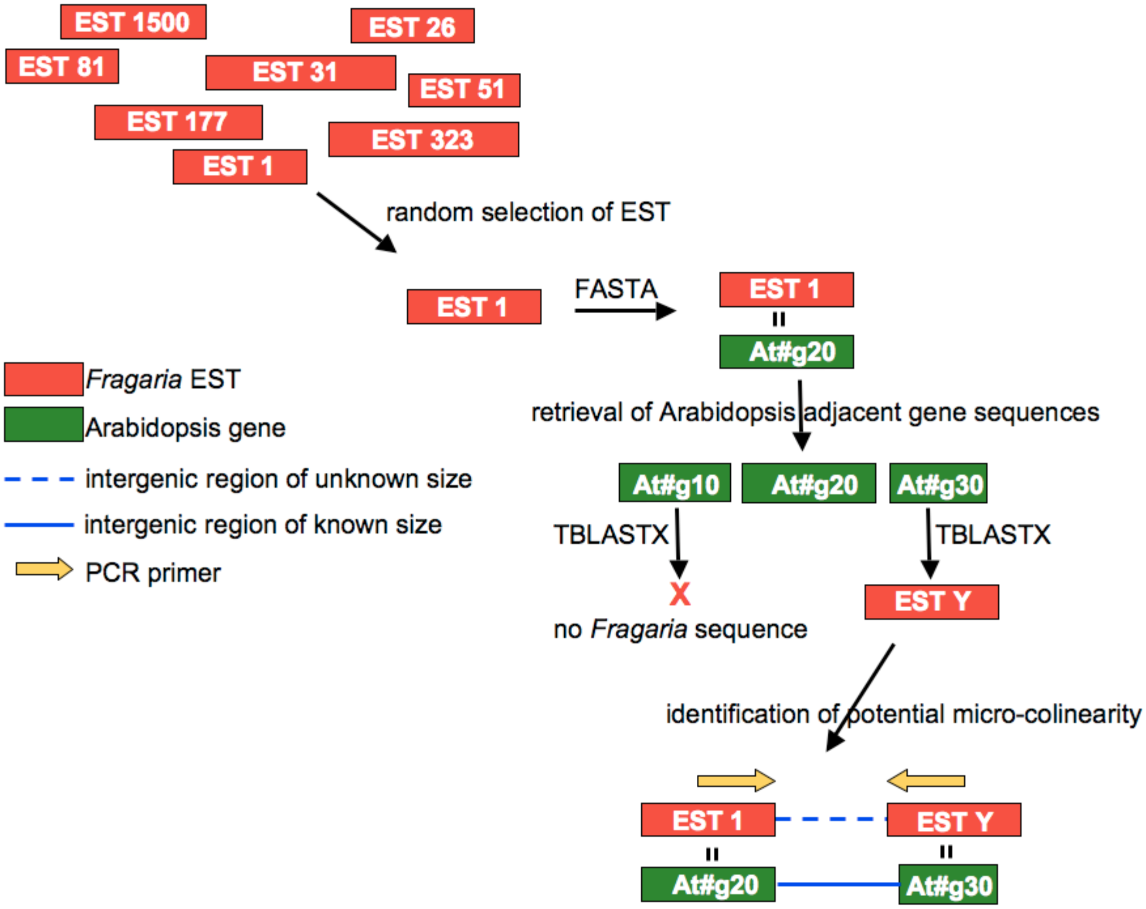

Figure 4-3. GPH design upon comparison between strawberry ESTs and *Arabidopsis* database. When the quest for homologies culminates in the detection of potentially neighboring genes, primers are designed in the strawberry EST and the intergenic region is amplified if the adjacency is true, the gene space is smaller than 4kb, and the gene orientations are conserved.

Table 4-1. PCR primers designed for amplification of micro-colinearity-inferred putative intergenic fragments

Primer name	*Fragaria* EST	*Arabidopsis* gene	Intergenic fragment size in *Arabidopsis* (bp)	Primer sequence
GPH4a	FA_SEa0007-G07	At2g20120	3,415	acgagggcttggaagaaagg
GPH4b	FA_SEa0015-B08	At2g20140		gcccaacaacagaaagacc
GPH5a#2	FA_SEa0002C03r	At3g07320	2,196	caatgccatggtctccggtc
GPH5b#2	FA_SEa0018E10r	At3g07330		tgccgttgcacacaccttcc
GPH20	FA_SEa0012D10r	At5g13440	1,043	gagggtaacgctcatggtt
GPH20	FA_SEa0012E08r	At5g13450		gtctccttcaattctttctcctc
GPH21a	AY679587	At5g06750	1,509	tgacatcccataagccatca
GPH21b	DQ011163	At5g06760		gggaggactacggcacataac
GPH21c	DQ011163	At5g06760		atcagatgtcggcactgc
GPH22 FaSCH6rgene	gi 48249442	At5g11250	1,632	tttcagctcagcaagcaagg
GPH22 FaHy5	FA_SEa0004E09r	At5g11260		gctcccaggaccaaacca
GPH23F	FA_SEa0013H07r	At1g23750	982	cttgagggccatcagcac
GPH23R	V01014C10_558132	At1g23740		tacacccacgccttcatctc
GPH27F	FA_SEa0014E11r	At1g74260	2,300	tgccgctgccatttctct
GPH27R	FA_SEa0011H07	At1g74280		ccatgctcttgataggccaaat
GPH31F	FA_SEa0014B05r	At2g30100	724	aatggagctgatggtttcgat
GPH31R	FA_SEa0016G12r	At2g30110		aaggatgatgacacgaactatca
GPH51F	CX662192	At3g176600	1,867	ggacacatggctcccaga
GPH51R	AY961594	At3g17670		caagacagcgggagcagt
GPH56F	AB211167	At4g38970	3,017	ccagggacgatgttttgctc
GPH56R	AJ414709	At4g38990		ggtggattacattttgggtgaca
GPH56R2	AJ414709	At4g38990		ttcaagctttggacaactaacg

Table 4-2. PCR primers that allowed amplicon generation. A “minus” sign in the “amplification” and “clone #” columns signify, respectively, no amplification and no transformants were observed.

Primer name	Template	PCR product size (kb)	Clone #	Vector	*E. coli* strain	Sequence obtained from forward end (bp)	Sequence obtained from reverse end (bp)
GPH4	‘Strawberry Festival’	2.0, 1.0, 0.5	13	TOPO	TOP10	1,109	-
			15	TOPO	TOP10	638	-
GPH5	*F. vesca*	2.8	21	TOPO	TOP10	1,268	1,299
	F. viridis	2.8	5	TOPO	TOP10	1,249	1,287
	F. iinumae	2.8	5	TOPO	TOP10	1,256	1,298
	F. nubicola	2.8	7	TOPO	TOP10	1,261	1,263
	F. nilgerrensis	2.8	19	TOPO	TOP10	1,262	1,295
	F. mandshurica	2.8	1	TOPO	TOP10	Full clone	
	‘Strawberry Festival’	2.8	2	TOPO	TOP10	1,257	1,306
			6	TOPO	TOP10	749	755
			7	TOPO	TOP10	Full clone	
GPH23	*F. vesca*	-					
	F. viridis	-					
	F. iinumae	2.0	2	TOPO	TOP10	Full clone	
			5	TOPO	TOP10	Full clone	
	F. nubicola	-					
	F. nilgerrensis	-					
	F. mandshurica	2.0	3	TOPO	TOP10	Full clone (2,024)	
	‘Strawberry Festival’	2.0	3	TOPO	TOP10	Full clone (2,081)	
			4	TOPO	TOP10	Full clone (2,111)	
GPH10	‘Strawberry Festival’	4.4	2	TOPO	TOP10	Full clone	
		4.4	7	TOPO	TOP10	Full clone	
		4.4	18	TOPO	TOP10	Full clone	
		4.4	19	TOPO	TOP10	Full clone	
		4.4	20	TOPO	TOP10	Full clone	
10PPR1/10AB22[*]	*F. vesca*	0.728		TOPO	TOP10	Full clone	
	F. viridis	0.726		TOPO	TOP10	Full clone	
	F. iinumae	0.266		TOPO	TOP10	Full clone	
	F. nubicola	0.722		TOPO	TOP10	Full clone	
	F. nilgerrensis	0.652		TOPO	TOP10	Full clone	
	F. mandshurica	0.724		TOPO	TOP10	Full clone	
	‘Strawberry Festival’	528	2			Subset of GPH10 sequence	
		644	7			Subset of GPH10 sequence	
		584	18			Subset of GPH10 sequence	
		584	19			Subset of GPH10 sequence	
		643	20			Subset of GPH10 sequence	
11D02	*F. vesca*	1.6	library				
	F. viridis	1.6	2031-1	pJET1	XL1-Blue	1,403	
	F. iinumae	1.6	2032-1	pJET1	XL1-Blue	Full clone	
	F. nubicola	1.6	2033-1	pJET1	XL1-Blue	1,299	
	F. nilgerrensis	1.6	-				
	F. mandshurica	1.6	-				
	‘Strawberry Festival’	1.6	-				

---

[*] 10PPR1/10AB22 is a locus within GPH10

Table 4-2. continued

Primer name	Template	PCR product size (kb)	Clone #	Vector	*E. coli* strain	Sequence obtained from forward end (bp)	Sequence obtained from reverse end (bp)
17O22	*F. vesca*	1.4	library				
	F. viridis	1.4	678	pJET1	XL1-Blue	Full clone	
	F. iinumae	1.4 & 1.0	668	pJET1	XL1-Blue	Full clone	
	F. nubicola	1.4	653	pJET1	XL1-Blue	Full clone	
	F. nilgerrensis	-					
	F. mandshurica	1.4	655	pJET1	XL1-Blue	Full clone	
	'Strawberry Festival'	1.5 & 1.4	-				
27F10	*F. vesca*	1.0	library				
	F. viridis	1.5	2039-1	pJET1	XL1-Blue	537	582
	F. iinumae	1.0	2040-1	pJET1	XL1-Blue	Full clone	
	F. nubicola	1.0	2041-1	pJET1	XL1-Blue	Full clone	
	F. nilgerrensis	1.8	-				
	F. mandshurica	1.0	2043-1	pJET1	XL1-Blue	Full clone	
	'Strawberry Festival'	1.0	2046-1	pJET1	XL1-Blue	-	
			2046-2	pJET1	XL1-Blue	Full clone	
29G10	*F. vesca*		library				
	F. viridis	-					
	F. iinumae	-					
	F. nubicola	0.7	2049-1	pJET1	XL1-Blue	Full clone	
	F. nilgerrensis	0.7	2050-1	pJET1	XL1-Blue	Full clone	
	F. mandshurica	0.7	2051-1	pJET1	XL1-Blue	Full clone	
	'Strawberry Festival'	-					
32L07	*F. vesca*	2.7	library				
	F. viridis	1.9	640 1090-11	pJET1 TOPO	XL1-Blue TOP10	Full clone	
	F. iinumae	-					
	F. nubicola	2.7	647	pJET1	XL1-Blue	No seq	
	F. nilgerrensis	2.7	993	TOPO	TOP10	No seq	
	F. mandshurica	2.7	1000	TOPO	TOP10	No seq	
	'Strawberry Festival'	I attempted to amplify fragment from the octoploids 'Carmine', 'Diamante', 'Rosa Linda', and 'Sweet Charlie', but amplification was not observed for any of them					
34D20	*F. vesca*	2.0	1826-3	pJET1	XL1-Blue	library	
	F. viridis	2.0	1827-3	pJET1	XL1-Blue	Full clone	
	F. iinumae	2.0	1828-4	pJET1	XL1-Blue	Full clone	
	F. nubicola	2.0	1829-1	pJET1	XL1-Blue	Full clone	
	F. nilgerrensis	2.0	1830-3	pJET1	XL1-Blue	Full clone	
	F. mandshurica	2.0	1831-5	pJET1	XL1-Blue	Full clone	
	'Strawberry Festival'	2.0	1832-2	pJET1	XL1-Blue	Full clone	
40M11	*F. vesca*	3.1	library				
	F. viridis	3.1	-				
	F. iinumae	2.9	-				
	F. nubicola	3.1	-				
	F. nilgerrensis	3.1	-				
	F. mandshurica	3.1	1088-1	TOPO	TOP10	884	
	'Strawberry Festival'	2.9	-				

Table 4-2. continued

Primer name	Template	PCR product size (kb)	Clone #	Vector	*E. coli* strain	Sequence obtained from forward end (bp)	Sequence obtained from reverse end (bp)
63F17	*F. vesca*	1.2	library				
	F. viridis	1.2	2	pJET1	XL1-Blue	Full clone	
	F. iinumae	1.2?	3	pJET1	XL1-Blue		
	F. nubicola		-				
	F. nilgerrensis		-				
	F. mandshurica	1.2	3	pJET1	XL1-Blue	Full clone	
	'Strawberry Festival'	1.2	1	pJET1	XL1-Blue		570
72E18	*F. vesca*	2.6	library	TOPO	TOP10		
	F. viridis	2.6	1096-4	TOPO	TOP10	1,052	1,515
	F. iinumae	2.6	1097-1	TOPO	TOP10	1,000	973
	F. nubicola	?	-				
	F. nilgerrensis	2.5	1099-1	TOPO	TOP10		1,958
	F. mandshurica	2.6	1100-6	TOPO	TOP10	Full clone	
	'Strawberry Festival'	2.5	1101-10	TOPO	TOP10	1,170	1,245
73I22	*F. vesca*	-					
	F. viridis	3.0	1834-16	pJET1	XL1-Blue	No seq	
	F. iinumae	-					
	F. nubicola	3.0	1836-1	pJET1	XL1-Blue	No seq	
	F. nilgerrensis	-					
	F. mandshurica	3.0	1838-5	pJET1	XL1-Blue	No seq	
	'Strawberry Festival'	-					

```
GPH5_ananassa_clone2      CAGAAGGTAATATGCATGATATAAATATCAAGTTAATTGTACAATGATATTATTTGTAATA 582
GPH5_ananassa_clone7      CAGAAGGTAATATGCATGATATAAATATCAAGTTAATTGTACAATGATATTATTTGTAATA 582
GPH5_viridis              TAGAAGGTAATATGCATGATATAAATATCAAGTTAATTGTACAGTGATAT---TTGTAACC 576
GPH5_iinumae              TAGAAGGTAATAT-------------ATCAAGTTAATTGTACAATAATAT---TTGTAATC 566
GPH5_nilgerrensis         CAGAAGGTAATATGCATGATATAAATACCAAGTTAATTGTACAATGATAT---TTGTAATC 579
GPH5_mandshurica          TAGAAGGTAATACGCATGATATAAATATCAAGTTAATTGTACAATGATAT---TTGTAATC 578
GPH5_nubicola             TAGAAGGTAATACGCATGATATAAATATCAAGTTAATTGTACAATGATAT---TTATAATC 583
GPH5_vesca                TAGAAGGTAATATGCATGATATAAATATCTAGTTAATTGTACAATGATAT---TTGTAACC 579
                           ************             * * ************* * ****   ** ***


GPH5_ananassa_clone2      GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_ananassa_clone7      GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_viridis              GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_iinumae              GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_nilgerrensis         GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAACT 240
GPH5_mandshurica          GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_nubicola             GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAACT 240
GPH5_vesca                GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
                          ********************************************************* **
```

Figure 4-4. Subset of the alignment of GPH5 octoploid and diploid clones. Single Nucleotide Polymophisms (SNPs) are in bold font. Same base changes that appear in a determinate position for more than one clone are likely to reflect real differences and are colored red. Hyphens signify indels whereas SSRs are magenta-colored.

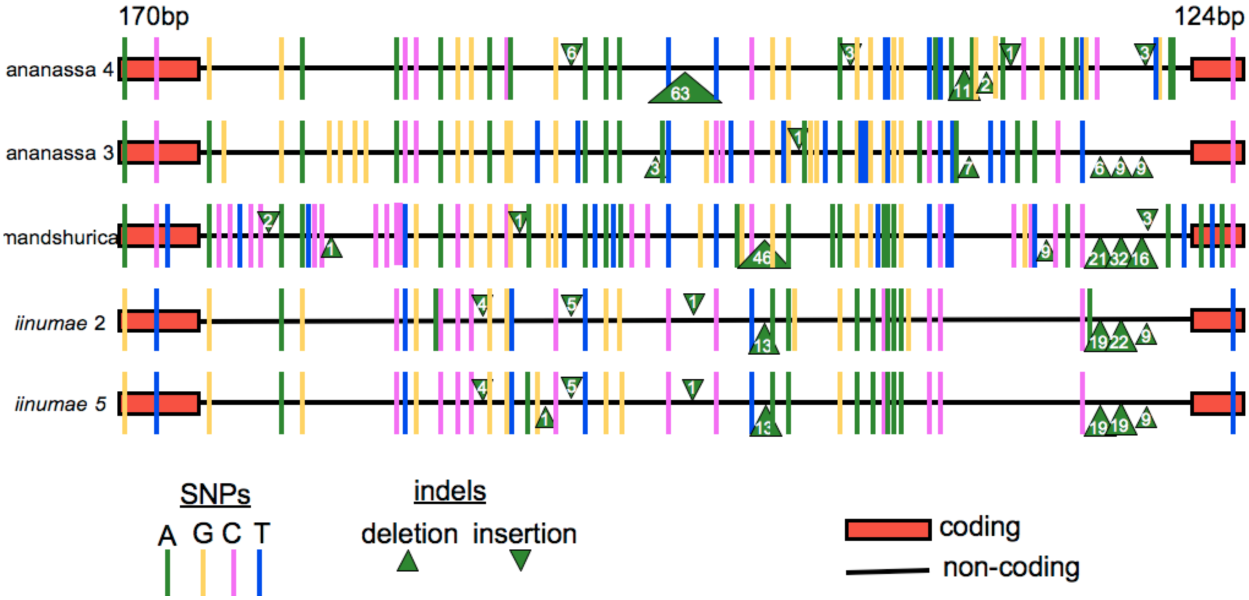

Figure 4-5. Diagrammatic representation of alignment of full GPH23 clones, depicting all polymorphisms identified, such as Single Nucleotide Polymorphisms, insertions, and deletions. Numbers in triangles indicate the length of indels.

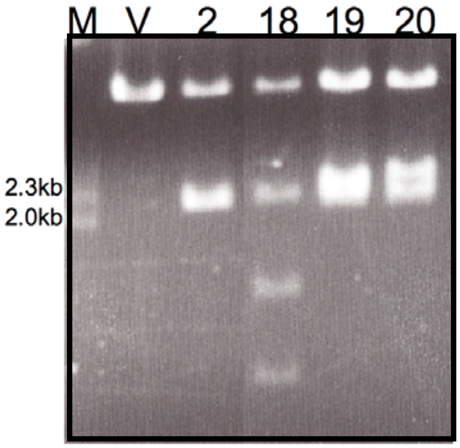

Figure 4-6. *EcoRI* Restriction patterns observed for GPH10 clones from the octoploid 'Strawberry Festival', indicating four different allele classes. M: molecular weight marker; V: empty vector; 2-20: polymorphic clones

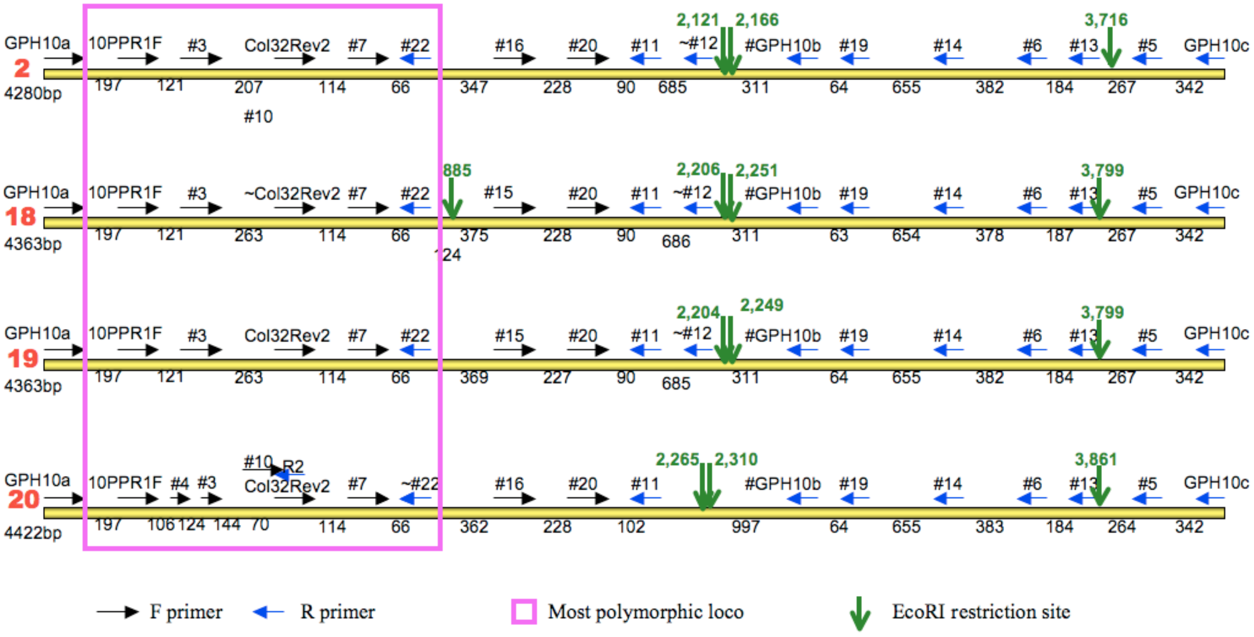

Figure 4-7. GPH10 clones, 4 alleles from the octoploid *Fragaria* × *ananassa*, detected by distinct *EcoRI* (green, vertical arrows) restriction patterns. The primers designed to amplify and sequence all 4.4kb clones are represented by black and blue arrows. The numbers between primers are the distances (in bp) between primers. The boxed region contained most of the polymorphism observed for GPH10, and it is comprehended between primers 10PPR1 and 10AB#22.

```
72E18_vesca             GAAAA-AAAGAGAGAGA--AAATTACAGATTTAAAGCGACGAACAA-TGAAAAGGAATGA 601
72E18_mandshurica       GAAAA-AAAGAGAGAGA--AAATTACAGATCTAAAGCGACGAACAG-TGAGAAGGAATGA 594
72E18_nilgerrensis      NNAAA-AAAGAGAGAGAGA---TTACAGATCTAN-GCGACGAACAA-TGAGAAGGAATGA 593
72E18_viridis           AGAAATAAAGAGAGAGA--AAATTACAGATCTAAAGTGACGAACAA-TGAGAAGGAATGA 604
72E18_iinumae           GAAAAAAAGAAGAGAGA--AAATTACAGATCTAAAGCGACGAACAAATGAGAAGGAATGA 658
72E18_ananassa          GAAAAAAAAAAGAGAGAGAAAATTACAGATCTAAAGCGACGAACAA-TGAGAAGGAATGA 638
                          *** **  ** ****   * ******** **  * ********  *** *********

72E18_vesca             GAGGCAAAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGGGAGAGA 661
72E18_mandshurica       GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGGGAGAGA 654
72E18_nilgerrensis      GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGT--------GA 645
72E18_viridis           GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGG--GAGA 662
72E18_iinumae           GAGACAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGT--------GAGAGA 710
72E18_ananassa          GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGA----------GTGAGGGAGAGA 688
                        *** ** ******************************* ****   *** ***

72E18_vesca             GAGAGAGAGATCGACGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 721
72E18_mandshurica       GAGAGAGAGATCGACGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 714
72E18_nilgerrensis      GAGAGAGAGATCGAAGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 706
72E18_viridis           GAGAGAGAGATCGAAGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 722
72E18_iinumae           GAGAGAGAGATCGAAGACGAGGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 771
72E18_ananassa          GAGAGAGAGATCGAAGACGAAGCTGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 749
                         ************* ***** ** ************************************
```

Figure 4-8. Subset of GPH72E18 alignment displaying SSR polymorphisms.

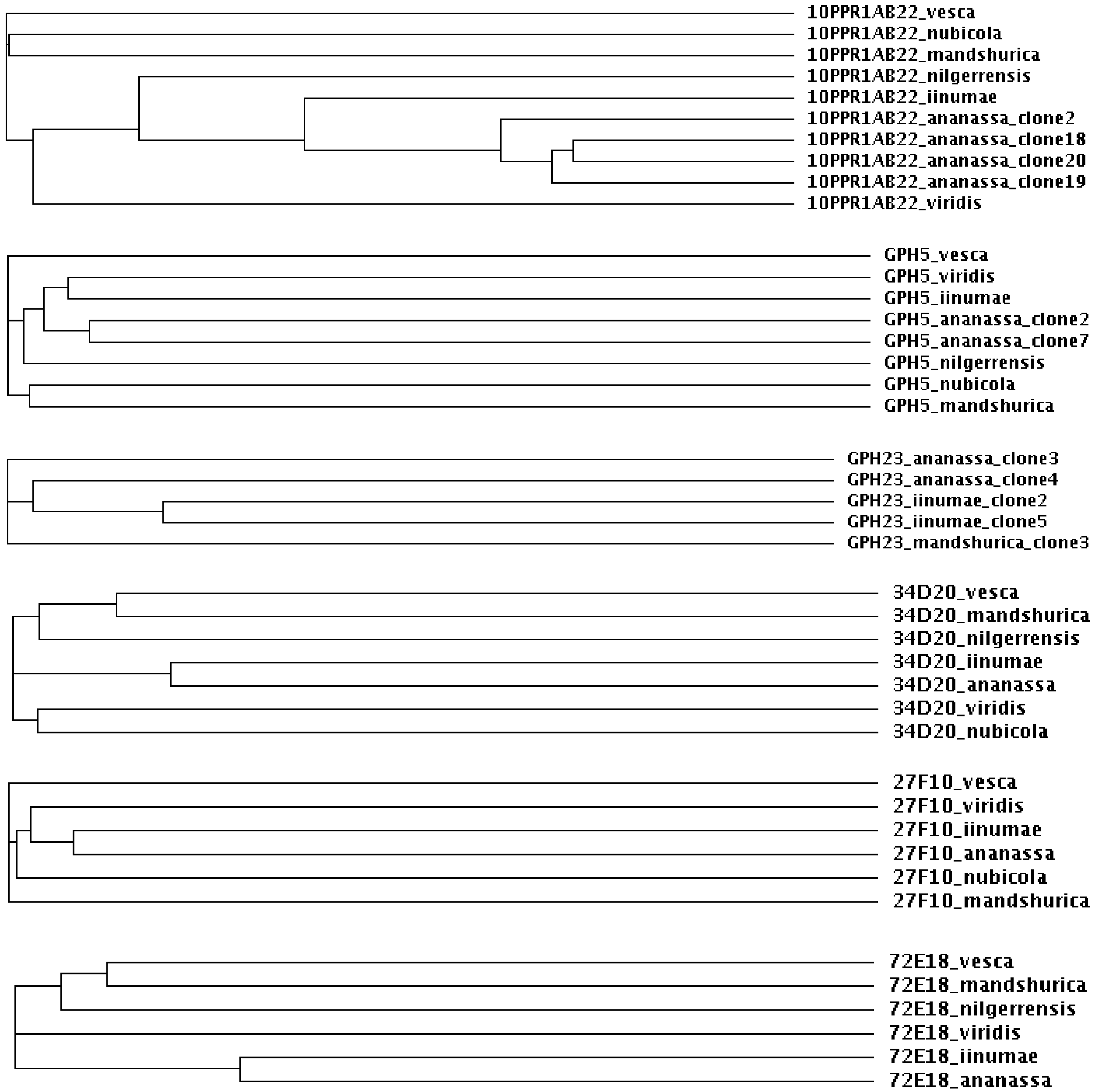

Figure 4-9. Cladograms of *F.* × *ananassa* and diploid alleles for six independent GPH loci. Amplified loci were sequenced, aligned with ClustalW, and their relatedness represented through cladograms. The *F. iinumae* clones were the most related diploid to *F.* × *ananassa* clones in every locus analyzed. *F. vesca* clones, on the other hand, were the furthest from the octoploid, contrary to prediction based on data of other author's previous studies.

Table 4-3. Overview of insertions and deletions detected through alignment of all sequenced clones. Each column represents an aligned region within haplotypes of a specific locus. The aligned regions where an indel or SSR were identified were named with Roman numbers. No relationship between clones of different loci is implied by the utilization of the same Roman number, as each locus was analyzed independently from the others. The Arabic numbers signify the number of bases in the deletions or insertions (*minus* or *plus* signs, respectively) in relationship to the consensus observed. White boxes represent accordance to the consensus sequence for the region in focus.

Clone	Indels and SSRs in Polymorphic Loci													
10PPR1AB22	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV
nubicola		-5						6 TA						
mandshurica		-5						5 TA						
vesca		-5						8 TA						
viridis		-5						7 TA						
nilgerrensis			-4						+6			+4	-44	
ananassa_18			-4				+36				-15		-176	
ananassa_20			-4	+19	+3	+7	+71		+6	-12	-15		-181	
ananassa_19			-4				+36				-15		-181	
ananassa_2			-4						+6	-20	-15		-181	
iinumae	+5		-4				-566							-8

11D02	I	II	III	IV	V	VI
viridis	-4		+8	+44		
nubicola					+5	
vesca					+5	
iinumae		-28		+32		+27

17O22	I	II	III	IV
vesca			-36	
mandshurica				
viridis		+5		-5
nubicola	-5		-16	
iinumae				

27F10	I	II	III	IV	V	VI	VII	VIII	IX	X
vesca	+2							-2		
mandshurica	+2				-2			-2		
nubicola					-9			-2		-3
iinumae		-7	+6	+12	-5					-3
ananassa		-14	+6	+12	-26	+8			+3	-3
viridis							+505			

29G10	I	II	III	IV	V
vesca					
mandshurica					
nubicola	-1	+2	-13	+7	
nilgerrensis					-1

Table 4-3. continued

Clone	Indels and SSRs in Polymorphic Loci

32L07	I	II	III	IV	V	VI
vesca				-9		
viridis	-792		-4		-9	-5

34D20	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
vesca		+4					+38				
mandshurica		+4					+38				
nilgerrensis			-7		-13			-11	-3	+3	+15
iinumae	-28		-2	-30		-3					
ananassa			-2	-30		-3			-3		+15
viridis			-2								
nubicola			-2		-13						

63F17	I	II	III	IV	V	VI
vesca						-2
mandshurica				-4	-6	
viridis	-6	+2	-4		-12	
ananassa						-2

72E18	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
vesca				4 GA	8 GA						
mandshurica	-8			4 GA	8 GA						
nilgerrensis				5 GA	6 GA	-11	-13	-18	-10	-44	
viridis				4 GA	7 GA						+3
iinumae			+53	3 GA	8 GA						
ananassa		-11	+45	4 GA	8 GA	-96					

CHAPTER 5

GENE-PAIR HAPLOTYPES: FUNCTIONAL AND TRANSFERABLE MARKERS AS NOVEL ADDITIONS TO THE DIPLOID *Fragaria* GENETIC LINKAGE REFERENCE MAP

Introduction

Strawberry (*Fragaria* × *ananassa* Duch.) is an economically valuable fruit crop, with average consumption of over 7.3 pounds *per capita* in 2005 in the United States (FAO STAT). The demand tends to increase due to public awareness of the potential health benefits of strawberry: small fruits have been shown to have high content of antioxidants (Wang, 2006), polyphenols and micronutrients that may play a role in human health.

Despite of the great importance of strawberry, knowledge of its genetic composition is very modest. The cultivated strawberry is octoploid, complicating development of molecular markers and construction of genetic linkage maps. Researchers have resorted to utilizing wild diploid strawberries to generate the first linkage relationships, in the hope of extending the findings to octoploid genomes. The first genetic linkages identified showed relationships between fruit color (Williamson et al., 1995) and runnering (Yu and Davis, 1995) to the shikimate dehydrogenase and phosphoglucoisomerase loci, respectively. These associations were shown in *Fragaria vesca*, a diploid that has been proposed to be a possible "A type" genome donor to the cultivated strawberry (Potter et al., 2000). The first indirect evidence of *F. vesca* as a genome contributor to the cultivated octoploid comes from cytological studies by Ichijima in 1926, where he showed the formation of 21 bivalents and 7 univalents during the pairing between *F. vesca* (then called *F. bracteata*) and *F. virginiana*, the pistillate parent to *F.* × *ananassa.*

The first genetic linkage map developed for strawberry was constructed using Randomly Amplified Polymorphic DNA (RAPD) markers developed for an F2 population derived from a cross between two subspecies of the diploid *F. vesca*: ssp. *vesca* 'Baron Solemacher' (redfruited, runnerless) and ssp. *americana* wild accession WC6 (Davis and Yu, 1997). The map was populated with 3 isozymes and 75 RAPD markers, of which 11 were codominant. This was possible due to a novel approach to Polymerase Chain Reactions (PCR), using mixed DNA templates for formation of heteroduplex bands (Davis et al., 1995). The locations of six genes involved in the anthocyanin pathway were assigned into this map later (Deng and Davis, 2001).

The second strawberry linkage map was developed for *F.* × *ananassa*, to increase the knowledge of the octoploid genome and to address questions on inheritance patterns in strawberry (if disomic or polysomic) (Lerceteau-Köhler et al., 2003). Amplified Fragment Length Polymorphism (AFLP) markers were used to generate separate maps for the male and the female parents, with 235 markers in 30 linkage groups, and 280 markers in 28 linkage groups, respectively. Though the study generated very detailed maps, with a total of 789 markers, AFLP markers are not easily transferable between species or even populations. The density of markers did add evidence of polysomic inheritance, since genes apparently crossed between subgenomes with some frequency.

A third map was constructed for strawberry (Sargent et al., 2004) addressing RAPD and AFLP transferability issues through the use of microsatellite markers or polymorphic Simple Sequence Repeats (SSRs). The map was based on a polymorphic F2 population generated from a wide inter-specific cross between the diploids *F. vesca* ssp. *vesca* f. *semperflorens* FDP815 (pistillate parent) and *F. nubicola* FDP601 (pollinating parent). These diploids have been shown to be the most closely related diploid relatives to the cultivated octoploid species (Potter et al., 2000). The creation of a reference map using a diploid relative is an approach commonly used to map genetically complex polyploids. Examples of polyploids for which reference maps have been constructed are wheat (Kam-Morgan et al., 1989), alfalfa (Diwan et al., 2000), and potato

(Milbourne et al., 1998). The map published in 2004 had 78 markers and new microsatellite loci were added later, totaling 182 markers (Sargent et al., 2006).

Strawberry belongs to the Rosaceae family, to which the horticulturally important peach, cherry, apple, raspberry, and rose also belong. Although SSRs are markers transferable between mapping progenies within and between species (Dirlewanger et al., 2002) (Hadonou et al., 2004), they are generally not transferable between genera. The challenge in developing transferable markers resides in the fact that markers are, by definition, placed on polymorphic regions of the DNA and, to be transferable, such markers are must be located on conserved regions. A recent study (Sargent et al., 2007) explored intron length polymorphisms, having PCR primers anchored in flanking exons that were conserved across *Prunus* and *Malus*, and thus generated highly transferable markers. In addition, because these markers were gene-linked, they also provided functional information.

A new approach to development of transferable and functional markers was explored by this research. The innovative mapping tool, named “Gene-Pair Haplotype” (GPH) consists of a stretch of intergenic space and takes advantage of its rich polymorphism for the development of markers. GPHs are PCR-amplifiable, with PCR primers anchored to exons of adjacent genes, making these makers transferable between species where microcolinearity is maintained. A significant degree of conservation between *Fragaria, Medicago* and *Arabidopsis* has been demonstrated (T. M. Davis, *personal communication*) suggesting that these same intervals might be easily transferable between rosaceous crops.

This investigation aimed to introduce the gene-pair haplotype concept as an innovative mapping tool, thereby increasing the number of transferable and functional markers genetically linked to the existing *F. vesca* x *F. nubicola* diploid reference map.

Materials and Methods

The diploid mapping population generated by Sargent *et al.* (Sargent et al., 2004) (a cross between *F. vesca* ssp. *vesca* f. *semperflorens* FDP815 and *F. nubicola* FDP601) was used in this study. Lyophilized tissue was received from the rosaceous genomics research group in East Malling Research station, in Kent, England. DNA was extracted as described in protocol #29, appendix A. Approximately 7 mg of lyophilized tissue was frozen in liquid nitrogen and ground in a mortar to a fine powder. After addition of 1ml of extraction buffer (2% CTAB, 1.4M NaCl, 100mM Tris-HCl pH 8.0, 20mM EDTA pH 8.0, 1% 2-mercaptoethanol), the tissue was further macerated until no defined leaf particles were observed. The volume was split into two 1.5-µl tubes, samples incubated at 65ºC for 1h, and 1 vol of 24:1 chloroform:octanol was added to each tube. After mixing the organic solvents with the extraction buffer and plant tissue, the samples were centrifuged at 13,000 rpm for 5 min. The upper phases were transferred to new tubes, and the nucleic acids precipitated by equal volume of isopropanol, centrifuged, the supernatant discarded, the pellet air-dried, and resuspended in 50µl TE pH 8.0. The DNA concentration varied from 40ng/µl to 4,543ng/µl.

Target regions for marker development were derived from the *F. vesca* ‘Pawtuckaway’ sequence annotation described in Chapter 3. Similarities between the *F. vesca* genomic sequence and either proteins or ESTs were sought for each of the 26 fosmid insert sequences. Within each fosmid clone, the most suitable pair of genes for PCR amplification was determined according to the following criteria:

i, The putative intergenic space should be large enough to permit detection of polymorphisms, but not larger than 3.5 kb due to technical limitations of amplification by PCR. Putative genes in fosmids 15B13 and 22L05 were separated by > 4kb, therefore these clones were excluded from the potential GPH pool.

ii, Tandem and non-tandem duplications were avoided as potential targets for PCR primer design. Target sequences should be unique to yield locus-specific amplification, since the assessment of more than one locus at a time would complicate data scoring. Tandem duplications were detected when adjacent *F. vesca* query sequences that had the same BLASTX hit, which appeared to indicate gene family clusters (e.g.: putative genes in fosmids 05N03, 13I03, and 18A19). An exception was made for the chalcone synthase (CHS) gene, which was included in the study although tandemly duplicated. The intergenic region is ~2,300 bp in *F. vesca* 'Pawtuckaway', but varies from 2 kb to over 8 kb in different rosaceous species tested, making this marker transferable across genera (T. M. Davis, *personal communication*).

Non-tandem duplications required indirect evidence, since only 1% of the *F. vesca* genome's sequence was available for analyses. If nucleotide identity was detected through BLASTN between a *F. vesca* sequence and more than one locus belonging to a single organism, that was regarded as evidence of potential duplication in *F. vesca*.

iii, Some fosmid clones did not appear to contain gene pairs when similarity to database-deposited protein sequence was the criterion adopted to classify a sequence as a putative gene. In those cases, a potential gene pair was inferred by two sequences displaying similarities, one to proteins and the other to ESTs.

Once apparent single copy, PCR-amplifiable putative gene pairs were identified, the software Primer3 (Rozen and Skaletsky, 2000) was used to facilitate design of PCR primers. For each primer pair, the forward primer was designed on the 3' end of a putative exon sequence of a gene, whereas the reverse primer was designed in the 5' end of the putative exon sequence of the downstream gene. In some cases, more than one primer pair was designed to generate a single band product, polymorphic between the parents—before or after restriction digest. In those cases, a primer with the fosmid clone name and orientation (F or R) had a suffix added to indicate another set. Figure 5-1 illustrates the case of primers designed for fosmid 40M11.

The PCR amplification 50µl-reaction components and conditions for the parental DNA and for the 94 F2 samples were: 1x buffer (at 10x concentration, composition was 35mM $MgCl_2$, 37.5µg/ml BSA, 160mM KCl, 400mM Tricine-KOH pH 8.0), 0.2mM each dNTP, 0.2µM each primer, 0.05unit Taq polymerase, 1µl DNA template, at variable concentrations (40ng/µl to 4.5µg/µl). Initial denaturation: 94ºC, 2 min, followed by 35 cycles of: denaturation at 94ºC for 15 sec, annealing for 45 sec, and extension at 72ºC. A last extension of 5min at 72ºC after the 35 cycles was executed. Table 5-1 contains functional information of the gene pair amplified, as well as annealing temperatures and the extension times dependent on the primer pair used. In general, extension was carried out for 1 minute per kb amplified, and the annealing temperature was primarily based on the primer melting temperature (T_m) calculated by Primer3 (Rozen and Skaletsky, 2000) using the formula described in (Rychlik et al., 1990) (though in many cases a range of temperatures had to be tested). Positive control primers FvLFYintron2F/ FvLeafy3' are anchored to the exons that flank Leafy gene's second intron (P. J. Stewart, *personal communication*). This intron size is variable among diploid species, being 770bp-long in *F. vesca* 'Pawtuckaway'. The annealing temperature and extension time were variable, since the primer pair under investigation was their determinant.

Following successful PCR amplification, 10µl of each single-band amplicons were digested with 1unit of different restriction enzymes (table 5-2) in a total volume of 20µl. Amplicon polymorphisms were resolved in 2% agarose gel, 1x TAE buffer, 0.5µg/ml ethidium bromide, at 80V, during variable times that were a function of the size of the digested fragments. The gel was exposed to 300-nm UV light for visualization of DNA fragments.

In order to obtain the most precise linkage, analyses were performed against the data set presented in Sargent *et al.* (Sargent et al., 2006), including new information available since the last publication. The novel GPH markers were assigned into linkage groups utilizing the software JoinMap 3.0 (Van Ooijen and Voorrips, 2001) with the application of the Kosambi mapping function and a minimum LOD score threshold of 3.0. The maps presented were constructed using MapChart software (Voorrips, 2002).

Results

Amplification was observed for all primer pairs, though not all were suitable for mapping purposes. Eight GPH primer pairs produced single-band amplicons that were scorable after restriction digest. The remaining primer pairs were not scored for the population for a variety of reasons. Primer pairs 01L02Fb/Rb, 01L02Fb/Rc, 22H18F/R, 22H18F/Rb, 30I24F/R, 32A10F/R, 32A10Fb/R, and 38H02F/R amplified multiple bands even at stringent annealing temperatures and restrictive extension times. The banding pattern for 01L02Fb/Rb appears to be due to a duplication, since two major bands are detected, one of the expected size, the other with higher molecular weight. Amplification by the other primer pairs displayed multiple bands, similar to non-specific amplification. There were primer pairs for which amplifications were observed, but they were not polymorphic (e.g. 10B08FbRb). For others, amplicons were polymorphic, but only a few members of the F2 population were amplifiable. This was the case of both 10B08 (GPHleafy/GPHacs, which amplified a 3.8 kb region that was polymorphic when digested with *EcoRI*) and 32L07F/Rb, polymorphic after treatment with *HaeIII*.

Both parents, when amplified by primers for 34D20, produced amplicons that were the same size. Restriction digestion revealed a rather complicated banding pattern. All of the digested amplicon fragment sizes < 700bp observed for 34D20 were expected, according to the predicted restriction pattern for *F. vesca* 'Pawtuckaway". An unexpected fragment of 1249 bp was observed for *F. vesca*, raising a concern that the putative single locus was in fact two loci. The other possibility was that the higher molecular weight band was a different *F. vesca* allele from the same locus. Had that been the case, a heterozygote should have been observed containing the female allele (1249, 300, 251bp) and the male allele (702, 335, 308, 251 bp). Such an individual was not observed, as 1249bp band cosegregated with the 758 and 429 bp bands. The presence of the 1.25 kb band was attributed to partial restriction digestion and the scoring was therefore carried out based solely on the expected 758 and 429bp bands *versus* the 702 and 335bp bands. Figure 5-2 shows banding pattern for digested amplicons of 34D20 and 72E18.

GPH40M11 is a dominant marker and amplifies a band only for the pistillate parent, *F. vesca*. Since the PCR amplification was precluded for half of the F2 population for some reason, this raised a concern about wrongly scoring individuals as homozygous *F. nubicola*. Thus, amplification patterns for all other 7 loci were compared, using a primer pair as positive control. Individuals for which amplification was observed in all those primer pairs but not observed for 40M11 were scored as homozygous for the *F. nubicola* allele.

The majority of the GPHs investigated were assigned to linkage group VII, as shown in figure 5-3.

Discussion

Gene pair haplotypes are intergenic, multiple character signatures that define suites of variability between two genomes. The purpose for these markers is to provide a complex field of discrete variation that can be related to a specific subgenome donor with the goal of eventually mapping genes to specific subgenomes of the octoploid strawberry. This chapter outlines the first step in this process, that is, to test if intergenic variability could be used to assign GPH loci to the diploid linkage map.

In all cases the GPH loci were assigned to the linkage map using a CAPS marker approach. Here amplicons were digested with a restriction enzyme that corresponded to sequence variation in the parental lines. A mapping population was treated with identical conditions to reveal the genotype of the specific F2 plant. Analysis of segregation with isozyme, morphological and molecular markers allowed assignment of these GPH loci to the diploid linkage map.

The assignment of these loci to the current map is important for two reasons. First, it demonstrates that the GPH is a viable marker- in this case based on a single restriction site. Other variable characters certainly exist in these regions that will complement the detection noted by this restriction site. In the future, these GPH loci will likely serve as anchors for the octoploid linkage map, because their likely variability supercedes that which is possible from a simple SSR or other marker used for diploid mapping.

This study places markers on linkage groups I, VI, and VII, with several independent markers in the latter. The next step is to translate these markers to an octoploid mapping population. This will immediately bring relevance to the endeavor because GPH loci stem from or are located near genes of known function. In this study GPH 17O22 is localized near F3H whereas 73I22 associates with chalcone synthase, two genes necessary for fruit color production and protective leaf pigments. A breeder with an interest in improving fruit color or possibly increasing plant survival in high light environments may find such loci useful in breeding selections.

The localization of the CHS gene determined by the GPH approach was different from the linkage group to which the gene was assigned when intron length was used to map it in a *F.*

vesca ssp. *bracteata* DN1C x *F. vesca* ssp. *vesca* 'Yellow Wonder' F2 population (Deng and Davis, 2001). This may be evidence of multiple copies of the CHS gene in the *Fragaria* genome. While described as a single-copy gene in Brassicaceae (Koch et al., 2000), CHS is a multigene family in many plant species (Jin-Xia et al., 2004). The CHS gene family is comprised of at least seven members, which, at least in petunia and poplar, are mapped to different linkage groups: II and V (Koes et al., 1987), and I and III (Tsai et al., 2006), respectively. It is possible that the different localizations in the genome correlate with different gene functions. In common morning glory (*Ipomoea purpurea*, Convolvulaceae) (Durbin et al., 1995) as well as in *Gerbera hybrida* (Asteraceae) (Helariutta et al., 1996), different family members have shown to have functional divergence.

The experimental outcomes of this chapter validate the use of GPH loci for mapping in the diploid strawberry and suggest great utility in application to octoploid mapping and breeding populations. Their complex characters, ease of detection, coupled to apparent disomic inheritance within octoploid subgenomes, indicate that these may be implemented in practical breeding scenarios.

Conclusions

The experimental trials outlined in this work test various aspects of strawberry structural genomics. From difficult honing of protocols to hasten DNA preparation from recalcitrant tissue, to computational analyses, development and proof-of-concept assessment of a novel molecular marker, these trials present new facets of understanding the complicated genome of the cultivated strawberry.

Recalcitrance to DNA extraction from plants is commonly attributed to their polyphenol and carbohydrate contents. Strawberry appears to be recalcitrant not only due to high sugars and phenols, but also because of strong physical barriers that guard the DNA. The results of over 103 systematic tests of various experimental conditions indicate that the most important consideration is complete disruption of the tissue via maceration, and that this process may be greatly enhanced by co-application of chemical lysis to disrupt tissue. My study provides a comprehensive evaluation of all published techniques and provides a unified protocol that works to some degree in all strawberry cultivars and species tested.

The importance of sequence information as a foundation for functional genomics studies in strawberry has been revealed by the discovery of enzymes associated with flavor (Wein et al., 2002) and fruit firmness (Llop-Tous et al., 1999) (Benitez-Burraco et al., 2003). This project represents the first efforts to examine the genome structure of *F. vesca*. The data indicate that the small genome of *F. vesca* maintains a character and compostion similar to other model plant species, suggesting that this species will have utility in answering questions within the Rosaceae family.

Annotation of fosmid inserts leads to the understanding of gene content and distribution, and permits marker generation for linkage mapping. More importantly, this initial survey of the strawberry genome is the first opportunity to compare strawberry to sequences to those of other organisms. Here relationships between the general properties of the genome have been deciphered. Strawberry is a gene-dense organism that maintains a significant content of mobile elements, and microcolinearity with other known genomes (T. M. Davis, *personal communication*).

Detection of gene pairs by searching for micro-colinearity between *F.* × *ananassa* and *Arabidopsis* is a clever approach, but it needs to be automated to increase the chances of finding adjacent genes. This approach has the advantage that it is not based on *F. vesca* sequence. Therefore, amplification of haplotypes is not biased towards *F. vesca*-like alleles. In addition, because sequences utilized for similarity search were from *F.* × *ananassa*, this method is better than the annotation of *F. vesca* genome method to address questions of diploid subgenome contributions to the octoploid. Primer pairs designed for gene pairs detected through this method amplified the octoploid, whereas most (8 out of 11) of the primer pairs generated through *F. vesca* genomic sequence did not amplify alleles from the cultivated strawberry. This study further supports the likelihood of *F. iinumae* as the B genome donor to the octoploid.

The approach based on gene prediction to identify gene pairs, had a higher amplification success rate and it is useful to characterize intergenic regions, serving as a tool to detect polymorphisms between diploids. Chapter 5 showed how this approach was successfully employed to create molecular markers in the *Fragaria* diploid reference map.

We have described the development and mapping of 8 markers, linked to at least one gene of known function. Therefore, this investigation proved the concept that putative intergenic regions may be used as functional markers. In addition, because the markers are designed for conserved sequences across different taxa in Viridiplantae, there is great potential for transferability and use on comparative mapping to appreciate Rosaceae structural genomics.

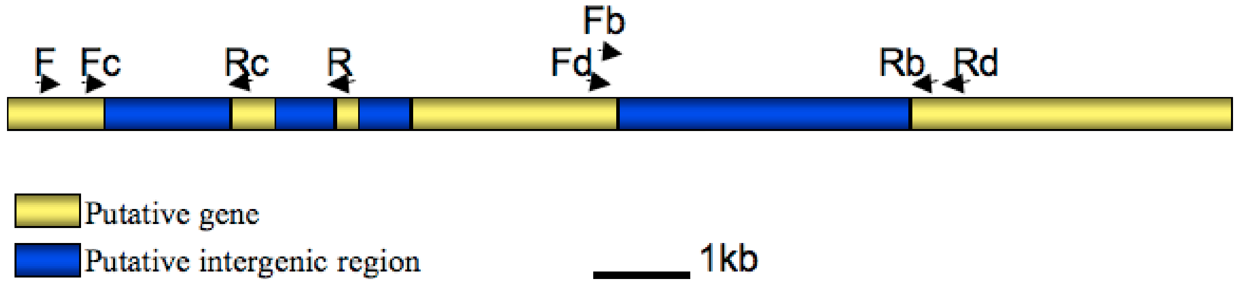

Figure 5-1. Fosmid 40M11 with primers designed on exons of FGENESH-predicted genic regions.

Table 5-1. PCR primer pairs and amplification conditions used in this study

Primer	Putative Gene Function or EST gb number	Sequence 5' to 3'	$T_{annealing}$ (ºC)	Extension Time
Control F FvLFYintron2F	Leafy	CACTGCCAAGGAGCGTGGTG	variable	variable
Control R FvLeafy3'	Leafy	TCAGTAGGGCAGCTGATG		
01L02Fb	EST AY573376	GAACCGTTCAAGTTCATAATTGG	54-65	1'30''-2'30''
01L02Rb	unknown protein	AAGGGAGGACGTTCAATGTG		
01L02Rc	unknown protein	ACGGAGATCGGGGACTTGT	54-58	2'30''
10B08F	Leafy protein	GGGCCAACTACATCAACAAGC	58-63	3'-4'
10B08R	ACC synthase	TGTTCTGTTGGGTGGACATGA		
10B08Fb	ACC synthase	TGCCATCGTTTCCATCAGTA	52	1'
10B08Rb	ribosomal protein	CGCGAAGATCATGAAGAACA		
11D02F	EST BQ105541	GAGCTGCTGTGTGAACCAAA	56-60	2'30''
11D02R	heat shock binding protein	GTTCAACTCCAGATGAAGTGAGG		
17O22F	Oligopeptidase	AAAATGGGTTGCACGAGTTC	60	2'
17O22Rb	Putative protein	GGGTTTCCTCACAAACTTCG		
17O22Fb	Oligopeptidase	GGTACCTCCAATGCAAGGAA	53-60	1'30'' - 2''
17O22R	Putative protein	TTCATCAGAGAAGGCGGACT		
22H18F	EST DY646954	ACCAATGCTTGGACACACAC	52-65	2'30''
22H18R	unknown protein	GATGAAATTCCATGCTTGTGAC		
22H18Rb	unknown protein	GGACTCCATGTAACACGGCTA	56-65	2'30''
27F10F	kinase	CCTGCAGGGTTTTTCATCAT	59	1'
27F10R	hypothetical protein	TGGAAATGTATTCTGGTTCTCC		
29G10F	phenylacetaldehyde synthase	TGGCCTTGTTTCCTAAACTCTT	59	1'
29G10R	unknown protein	AGAAGAAGGCAGCACCCAAT		
30I24F	transferase	TTGAGAGAGGTCTCCAAGCTC	54, 59	4'
30I24R	chromating remodeling factor	CGGAAGATGGCAAGCTATTG		
32A10F	isomerase (E > 10^{-12})	CGGAGAGAACGATGGAGTTG	52-62	1
32A10Fb		CCAAATGAATCAAGCTCAAGTG		1'30''
32A10R	pathogenesis-related protein	ATTGTCGACCAGTGCAGCAA		
32L02F	SMC2 (Structural maintenance of chromosomes)	GAGTTGAAAAACGGGTCGAA	53-61	2'
32L03Fb		CCTTCCAAGGTCACCTCCTT		
32L02Fc		TTAGCCCGGTTATGGAGTTG		
32L02R		GAAGGTTCAAGGAGCATGGA		
32L02Rb	Exostosin	AGGAAAATGCGGGAGAAAGT		
32L02Rc		GAACGATTTCCGAGGTGTGT		

Table 5-1. continued

Primer	Putative Gene Function or EST gb number	Sequence 5' to 3'	$T_{annealing}$ (ºC)	Extension Time
34D20Fb	RNA recognition motif	GCAGAAAGAAACTGATGTGCTT	60	3'30''
34D20Rc	cysteine-type peptidase	CGCAGTCGTAAAAATTCGTCT		
38H02F	serine/threonine kinase	CCAGGCCTAAGCTTGTCATC	53, 54, 60	4'
38H02R	exportin	AAGGCATTGAAATCATTCTACCA		
40M11F	F-box protein	ACACAGGTCATTGGGTCCAT	60	4'
40M11Fc		TTGACCCGGATAACATGGAT		
40M11R	transposase (E > 10^{-13})	GTGTTGCACAAGTCCATTCG		
40M11Rc	expressed protein (E > 10^{-9})	CTGACAGCGAATCAATCTGC		
40M11Fb	secretory protein SEC14	GGCCTTCTTGACATTCCAGT		
40M11Fd		CAACATTTTGGTGGCCTTCT		
40M11Rd	ATPase	CGGCCTATGAAACCACAGTT		
40M11Rb		TGGGGTTGTTGGAAAGAGAG		
63F17F	phospholipase D	CGCTCTATGGAAGGGACAAG	59	1'
63F17R	unknown protein	TTAAGGGGTCTGTTGATGTGC		
72E18Fb	actin	GCTAGGGAAAACAGCTCGTG	60	2'30''
72E18Rb	elongase	TGGGTTTGGTTTTGGGATAA		
73I22F	chalcone synthase A	CAAGCCTGAGAAGTTAGAAGC	62	5'
73I22R	chalcone synthase B	GAAAGTAGTAGTCGGGGTATGT		
GPH10a	unknown protein	GGCTTCTTCTTGTCCGGCAGC		
GPH10b	unknown protein	GAACTCCAGGTCAGATCTTCG		2'30''
GPH10c	unknown protein	CTCGCTGCAAATCAGCTACC		4'

Table 5-2. Fragment sizes of parental amplicons digested with restriction enzymes

Locus	Restriction Enzyme	Amplicon estimate fragment sizes (bp)		
		Non-digested	Digested	
			F. vesca	*F. nubicola*
17O22FRb	*RsaI*	1,374	486, 413, 292, 83, 67, 28, 5	511, 414, 293, 83, 67, 28, 5
34D20FbRc	*AluI*	2,050	1249, 758, 429, 300, 251, 107, 69, 48, 26	702, 335, 308, 251, 107, 69, 48, 45, 41, 26
40M11FdRd	Dominant marker	3,100	present	absent
63F17	*HaeIII*	1,266	992, 234, 40	840, 234, 40
72E18FbRb	*HhaI*	2,620	1,400, 800, 300	2,300, 300
73I22	*PvuII*	3,000	2,200, 1,000, 600	2,200, 1,500

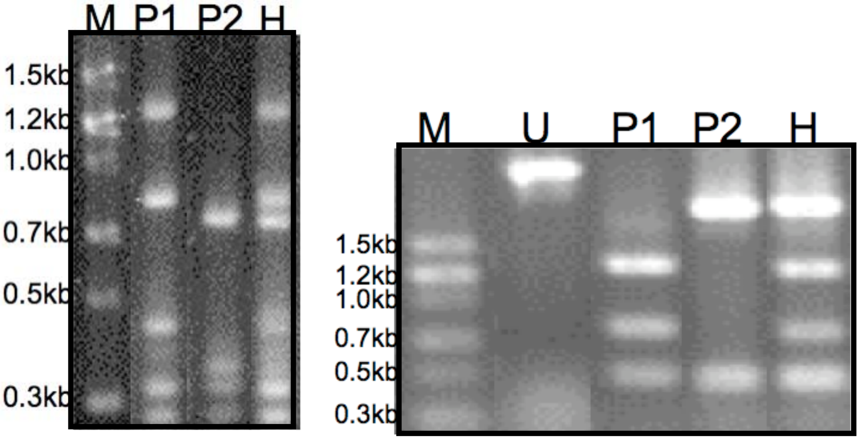

Figure 5-2. Amplicon restriction patterns for GPHs 34D20 and 72E18. M: molecular weight marker; U: uncut amplicon; P1: female parent, *F. vesca*, P2: male parent, *F. nubicola*; H: heterozygote.

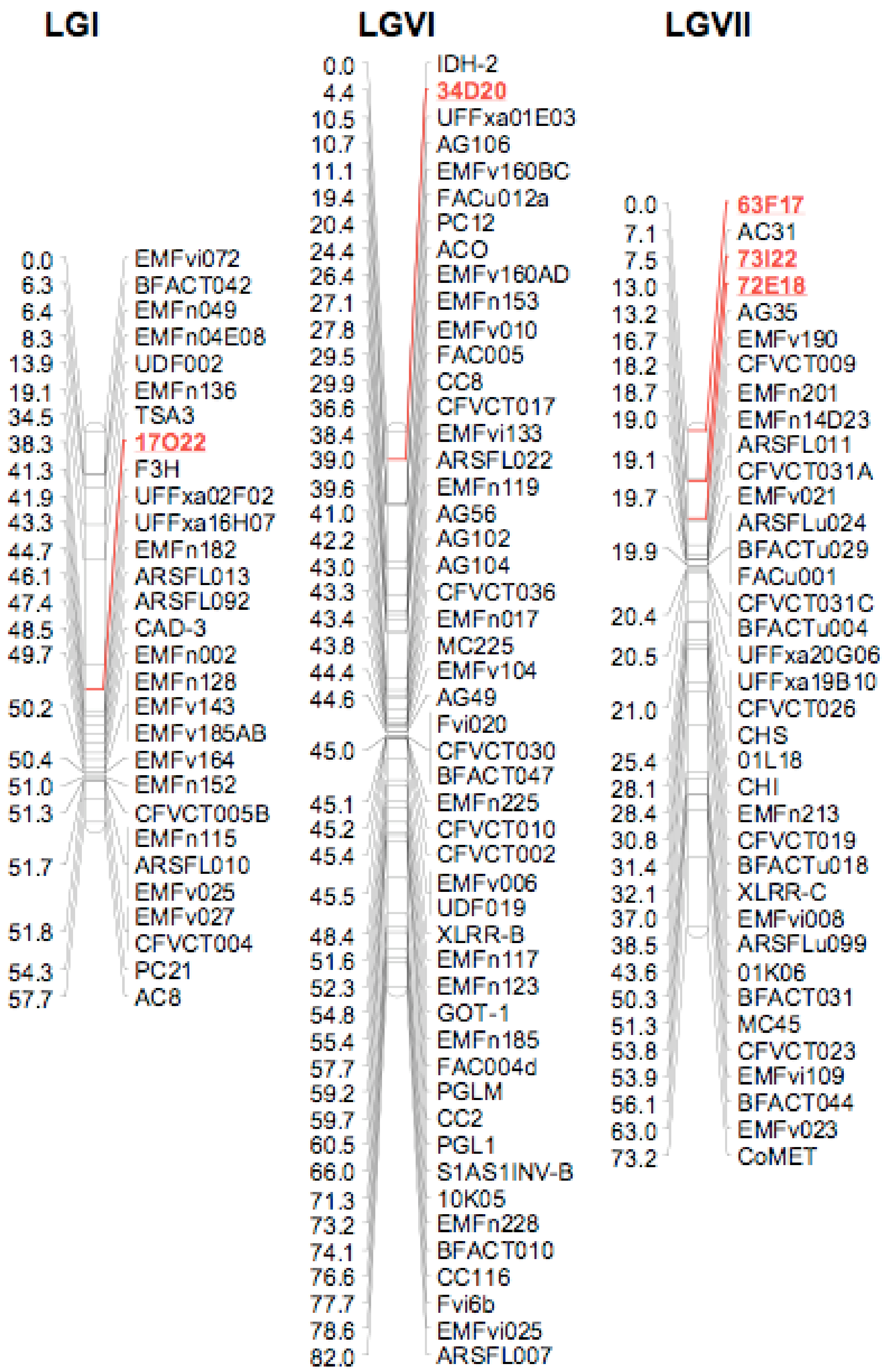

Figure 5-3. Gene-Pair Haplotypes assigned to linkage groups of the reference *Fragaria* map.

APPENDIX A
DNA EXTRACTION PROTOCOLS

The numbered items bellow represent different protocols, whereas numbers preceded by a "T" signify treatment number and correlate with the treatment numbers used in Table 2-1.

In all protocols that used either 2-mercaptoethanol, sodium bisulfite or sulfite, these reducing agents were added just prior to use of buffers. Most procedures included at least one 25:24:1 phenol:chloroform:isoamyl alcohol deproteination step followed by one 24:1 chloroform:octanol extraction. When RNAse-treated, the enzymatic reaction was carried out at 50µg/ml. Precipitation of DNA was executed by adding 0.7 to 1 volume of isopropanol or by sodium acetate to reach final concentration of 0.3M plus two volumes of absolute ethanol, then washed with 70% ethanol, dried, and resuspended in sterile, deionized water. Except for buffers that involved guanidine thiocyanate, which were kept at room temperature, plant material in buffer was incubated 30-60 minutes at 65°C, unless otherwise stated. When product was obtained, 5-10 µg of DNA were digested with 2-4 restriction enzymes. Below is a brief description of each protocol.

DNA Extraction from Leaves

1. Tomato [Fulton, 1995]: utilizes a combination of a DNA extraction buffer (0.35M sorbitol, 0.1M Tris-base, 5mM ethylenediaminetetraacetic acid, EDTA, pH 7.5) and a nuclei lysis buffer (0.2M Tris, 0.05M EDTA, 2M NaCl, 2% CTAB) to make the micro prep buffer (42% extraction buffer, 42% nuclei lysis buffer, 16% sarkosyl 5%, and 0.02% sodium bisulfite). Used 0.5g (T1), 1g (T2), and 2g (T3) of 'Strawberry Festival' fresh mature leaf tissue, extracted by 5ml buffer.

2. Woody plants [Kobayashi, 1998], modified by A. M. Hadonou. Two extraction buffers are consecutively used, buffer 1 being used twice and the buffer 2 only once. Following centrifugation with buffer 1 (50mM Tris-HCl pH 8.0, 5mM EDTA, 0.35M sorbitol, 0.3% 2-mercaptoethanol, 10% polyetheleneglycol, PEG), the supernatant is discarded before adding buffer 2 (50mM Tris-HCl pH 8.0, 5mM EDTA, 0.35M sorbitol, 0.3% 2-mercaptoethanol, 1% sarkosyl, 0.7M NaCl, 0.1% CTAB). Used 0.1g of two cultivars of *F. vesca* ssp. *vesca* f. *semperflorens*: Yellow Wonder (T4, T6) and Alexandria (T5, T7); fresh expanded leaf tissue, extracted by 1ml (T4, T5) or 10ml (T6, T7) of buffer.

3. Guanidine thiocyanate [Chomczynski, 1987]: The incubation was carried out for 5-15 minutes only and at room temperature instead of 65°C. Buffer composition: 4M guanidine thiocyanate, 100mM Tris-HCl, 10mM EDTA, 0.5M NaCl, 1% sarkosyl, 1% sodium sulfite. Newly expanded (T8) and unexpanded (T9) leaves of 'Sweet Charlie' were used for extraction from 100mg tissue in 100µl buffer. Further treatments to aliquots of the product of this prep were performed, aiming removal of contaminants: adsorption to a column from the DNeasy Plant Mini kit (T10) or dialyses into TE pH 7.0 at 4°C (T11). Dialyses was performed overnight, TE buffer replaced by fresh buffer, and dialyzed again for another day. Sample was 50µg/ml RNAse- and 150µg/ml proteinase K-treated. DNA isolation was continued with phenol extraction and standard downstream steps.

4. Guanidine thiocyanate and CTAB utilized consecutively (T12): DNA isolation according to Chomczynski [Chomczynski, 1987], and resuspension of the ethanol-precipitated DNA in CTAB buffer described in Chang [Chang, 1993] for re-extraction, an attempt to rid

DNA prep of polysaccharides. The incubations were carried out at room temperature and 65°C with guanidine thiocyanate and CTAB, respectively.

5. Guanidine thiocyanate and CTAB used simultaneously: extraction buffer kept at room temperature, 15 minutes: 4M guanidine thiocyanate, 100mM Tris-HCl, 10mM EDTA pH 8.0, 0.5M NaCl, 1% sodium sulfite, 1% sarkosyl, 2% CTAB, 1% PVP, 2% 2-mercaptoethanol. Treatments included extraction from 10mg (T13, T15) and 100mg (T14, T16) of lyophilized (T13, T14) or fresh (T15, T16) tissues.

6. DNAzol® Extra Strength kit [Chomczynski, 1997]: incubation at room temperature, as suggested by manufacturer. Exact composition of buffers is cryptic, though it is known to contain a guanidine detergent. Tested extraction from 100mg (T17, T19) and 500mg (T18, T20) of lyophilized (T17, T18) or fresh (T19, T20) tissues.

7. Pine tree [Chang, 1993]: this protocol was originally designed for RNA extraction and was adapted here to DNA extraction by omitting the lithium chloride step. Buffer: 2% CTAB, 2% polyvinyl pyrrolidone (PVP), 100mM Tris-HCl, 25mM EDTA, 2M NaCl, 0.5g/L spermidine, 2% 2-mercaptoethanol. After the addition of equal volume of chloroform, samples were homogenized using a Polytron for 1 minute, at 9/10 of maximum speed. Tissue: 0.5g in 7ml buffer (T21).

8. Urea [Settles, 2004]: Phenol deproteination step was done together with incubation with extraction buffer, at room temperature for 20 minutes in 8M urea, 0.4M NaCl, 60mM Tris-HCl pH 8.0, 25mM EDTA pH 8.0, 1.5% sarkosine (T22). A variant of the buffer was also experimented, which consisted of supplementation with 1% sodium sulfite and 1% PVP to prevent oxidation of phenols (T23).

9. Strawberry (Manning, 1991): Buffer: 0.2M Tris, pH adjusted to 7.6 using boric acid (which forms complexes with polyphenols at pH 7.5 (King, 1971) and with carbohydrates (Gauch and Dugger Jr., 1953)), 10mM Na2EDTA, 0.5% SDS, 2% 2-mercaptoethanol. After a 10-minute incubation at room temperature, equal volume of 25:24:1 of phenol:chloroform:isoamyl alcohol was added, mixed, and centrifuged for 10 min at 3,500rpm. Upper phase was transferred to a new tube (called "Tube A" here). "Tube B" contained inter- and lower phases from this first round of chloroform extraction. A second volume of extraction buffer was added to Tube B and a second round of chloroform extraction took place. The new upper phase from Tube B was combined with Tube A and split into 6 aliquots. Two aliquots (T24, T27) had polysaccharides precipitated by addition of 0.4 volume of 2-butoxyethanol, iced for 30 minutes, and centrifuged at 3,500rpm for 10 minutes. The other four aliquots were diluted by 2.5 (T25, T28) and 4 volumes (T26, T29) of a combination of 1M Na acetate buffer (pH adjusted to 4.5 by acetic acid) and water. The relative volumes of water and 1M Na acetate/acetic acid buffer were calculated to raise the Na concentration to 80mM. Considering that at this point each treatment had a volume of 3.3ml, the dilution by 2.5 volumes brought the volume to 8.3ml. Therefore, 664µl of the 1M Na acetate/acetic acid buffer and 4.3ml of water were required to reach the desired concentration of 80mM Na^+. In the case of the dilution by 4 volumes, and still considering initial volume as 3.3ml, the final volume was 13.2ml. The sample received 10ml of (water+ sodium buffer), of which 9.2ml were water and 800µl were the 1M Na acetate/acetic acid buffer. After dilutions were made, T25, T26, T28, and T29 were precipitated as before: 2-butoxyethanol, were iced, and centrifuged. The goal of the centrifugation here is to precipitate polysaccharides, not nucleic acids. The six supernatants were transferred to new tubes and equal volumes of 2-butoxyethanol were added to precipitate nucleic acids. After icing for 30 minutes, the tubes were centrifuged for 10 min at 3,500rpm, the supernatant discarded, and the pellet

washed with a 1:1 solution of 0.2M boric acid/Tris, 10mM Na2EDTA (pH 7.6) : 2-butoxyethanol. Pellets were washed with 70% ethanol, 0.1Kacetate/acetic acid (pH 6.0), then with absolute ethanol. After dry, pellets were resuspended in 1ml water, and 10µg DNA digested with restriction enzymes. An aliquot of one of the treatments (T28) was EcoRI-digested before and after treatment with 150µg/ml Proteinase K and with phenol:chloroform. A second attempt to isolate digestible DNA using the strawberry protocol was made, adding antioxidants 4% PVP and 5mM ascorbic acid to the extraction buffer (T32-T35).

10. Several attempts were made to determine which isolated variable in the strawberry protocol plays the major role in DNA yield. The possibilities raised were: i, the SDS, rather than CTAB, nature of the protocol. Treatment numbers T16, T18, and T24 used SDS, therefore testing this variable; ii, the boric acid, instead of HCl, used to adjust the pH of Tris; iii, the re-extraction of interphase formed after chloroform treatment; iv, the dilution that raised Na concentration to 80mM prior to DNA precipitation; v, precipitation by 2-butoxyethanol in place of isopropanol or ethanol. The isolated roles of boric acid and 2-butoxyethanol in DNA isolation were addressed by using a buffer similar to the one proposed by Murray and Thompson, but adjusting the pH of Tris to 7.6 with boric acid, rather than HCl (buffer: 200mM Tris/borate, 200mM EDTA, 2.2M NaCl, 2% CTAB, 2% 2-mercaptoethanol, 2% PVP), and precipitating one treatment with isopropanol (T30) and the other with 2-butoxyethanol (T31).

11. The strawberry protocol suggests two different dilutions (2.5 volumes or 4 volumes) to elevate the Na^+ concentration to 80mM. The chosen dilution here was the 2.5vol. An experiment was set up to test the merits of the combinations of two factors: i, re-extraction of the interphase by extraction buffer and chloroform; and ii, DNA precipitation by 2-butoxyethanol. The former factor was tested by keeping each, the first and the second extraction rounds, as separate treatments, therefore determining the gain in DNA yield given by the second extraction. The latter factor contrasted the use of isopropanol *versus* 2-butoxyethanol, where T32=first extraction round/isopropanol; T33=first extraction round/2-butoxyethanol; T34=second extraction round/isopropanol; T35=second extraction round/2-butoxyethanol.

12. Finally, 2-butoxyethanol was used in the guanidine thiocyanate protocol (number 3). The treatments were essentially the same as described for T8 in protocol number 3, except that 2% 2-mercaptoethanol was added to the extraction buffer and the Tris was adjusted by boric acid, not HCl. 100mg of tissue processed by 6ml buffer. Nucleic acids precipitations were done by isopropanol (control, T36), and 2-butoxyethanol (T37).

13. According to an article that proposes a method to isolate DNA from cashew (Rout et al., 2002), boric acid can be used in replacement of Tris, instead of assuming the role of simply adjusting the pH of a Tris solution. The buffer composition used in treatment T38 was 1M boric acid pH 8.0, 2mM EDTA, 1.4M NaCl, 4% CTAB, 0.2% 2-mercaptoethanol.

14. Epicentre kit. Used 10mg (T39), 30mg (T40), 100mg (T41) of 'Strawberry Festival' leaf tissue with 300µl buffer.

15. PowerPlant DNA Isolation kit from MO BIO (T42). A leaflet (350mg) of fresh FRA520 (*F. nubicola*) was ground with liquid nitrogen in microfuge tube. The remaining steps were carried out according to manufacturer's directions.

16. Qiagen DNeasy Plant Mini kit (T43). Followed company's directions for fresh tissue.

17. Silica-based DNA extraction. Nucleic acids tend to adsorb to silica in the presence of chaotropic salts, such as sodium iodide (NaI) (Vogelstein and Gillespie, 1979), guanidine thiocyanate, and guanidine hydrochloride. The binding capacity depends on the solution's ionic strength and pH, being higher at concentrated solutions and pH<7.5 (GeneClean Manual). Silica

columns have been used elsewhere to eliminate polysaccharide contaminants, which is verified by increase of the ratio A260/230 (Abdulova et al., 2002). The protocol used here was based on Rogstad's article (Rogstad, 2003), which uses a CTAB extraction buffer and describes the preparation of the silica binder. CTAB extraction buffer: 2% CTAB, 1.4M NaCl, 100mM Tris-HCl pH 8.0, 20mM EDTA pH 8.0, 1% 2-mercaptoethanol. 'Strawberry Festival' leaves were ground (10mg—T44 and 100mg—T45) and 5 ml of extraction buffer were added. Incubation was carried out at room temperature for 30 minutes. Equal volume of chloroform:octanol was added, samples were centrifuged, the upper phase was transferred to a new tube, and 2.5ml of silica binder were added. The mixture was agitated thoroughly for 5 min, then centrifuged. The supernatant was discarded, and 4ml of silica wash (25% isopropanol, 25% ethanol, 100mM NaCl, 10mM Tris-HCl pH 7.4, 2mM EDTA pH 8.0) were added, vortexed to resuspend the silica. Samples were centrifuged, supernatant discarded, and a second wash took place. The silica pellet was dried for 2 hours at 37ºC, and the DNA was eluted by 1ml of ultra pure water, vortexed, and incubated at 65ºC for 5 min. After centrifugation, the upper phase was transferred to a new tube, RNAse-treated, then DNA was precipitated by isopropanol.

The following protocols (18-22) attempted to extract DNA from nuclei isolated from leaf tissue. Protocols 23-33 consist of variations of the protocol by Murray and Thompson and utilized leaves (rather than isolated nuclei) for DNA extraction.

DNA Extraction from Isolated Nuclei

Nuclei were purified according to the procedure described by Folta and Kaufman [Folta, 2000] and nuclei were recovered from the 35/80 interphase of percoll gradients. Nuclei were incubated with each extraction buffer at 65°C for at least 10 minutes. The following buffers were mixed to 50-150μl of purified nuclei in storage buffer as an attempt to extract DNA:

18. Qiagen DNeasy Plant Mini kit. Different volumes (50μl—T46 and 150μl—T47) of isolated nuclei were processed according to manufacturer's directions.

19. Fulton's nuclei lysis buffer [Fulton, 1995], supplemented with 0.5% sodium bisulfite: 200mM Tris pH 7.5, 50mM EDTA pH 8.0, 2M NaCl, 2% CTAB. Two tubes, one 50μl nuclei (T48) and the other containing 75μl nuclei (T49), were incubated with 200 and 75μl of nuclei lysis buffer at 65ºC for 45min. Phenol:chloroform followed by chloroform extractions took place, the upper phase transferred to a new tube, and DNA precipitated by isopropanol.

20. Peterson's procedure [Peterson, 1997]: 20% SDS was added to a final concentration of 2% and mixed with 50μl nuclei (T50) or 150μl (T51) by gentle inversion to lyse the nuclei. The mixture was incubated in water bath at 65°C for 10 minutes, cooled to room temperature, then 5M sodium perchlorate was added to reach final concentration of 1M. Sodium perchlorate is used to dissociate nucleic acid-protein complexes [Wilcockson, 1973]. Following centrifugation, the upper phase was transferred to a new tube using a large-bore tip. After a phenol deproteinization step, the aqueous phase was dialyzed twice, the first overnight and the second for an entire day, both into TE pH 7.0 at 4°C. Samples were consecutively treated with 50μg/ml RNAse for 1 hour and with 150μg/ml proteinase K. After extractions with phenol:chloroform/isoamyl alcohol and chloroform/isoamyl alcohol, DNA was precipitated and resuspended.

21. Guanidine thiocyanate buffer (4M guanidine thiocyanate, 100mM Tris-HCl, 10mM EDTA, 0.5M NaCl, 1% sarkosyl, 1% sodium bisulfite) was used (750μl) to extract DNA from 50μl nuclei (T52). The buffer/nuclei were incubated at room temperature for 10min and were

followed by phenol:chloroform and chloroform extractions. DNA was precipitated by isopropanol.

22. Use of triisopropylnaphthalenesulfonic acid (TIPS) as a hydrotrope in a surfactant system (Bies and Folta, 2004). Hydrotropes stabilize surfactants (e.g. SDS) to allow them to remain soluble. Nuclei (150µl) were incubated with 1200µl of extraction buffer 1 (10mM EDTA, 10mM Tris, 1%SDS) at 65ºC for 20min (T53). The sample was treated with Proteinase K for 1h at 37ºC. After a phenol:chloroform extraction and centrifugation, the interphase was re-extracted with 5 volumes of extraction buffer 2 (50mM Tris-HCl pH 8.0, 5% SDS, 1%TIPS, 2% 2-mercaptoethanol, 4% PAS—p-aminosalicylic acid). The supernatants of both extractions were combined and nucleic acids precipitated by isopropanol.

Modifications of Murray and Thompson DNA Isolation Protocol

A series of modifications of the protocol proposed by Murray and Thompson were tested. Though the original protocol included cesium chloride gradient, this step was suppressed for all variations tested.

23. Extraction buffer: 200mM Tris, 2M NaCl, 50mM EDTA, 2% CTAB, 2% PVP, 2% 2-mercaptoethanol. After initial 45min incubation at 65ºC, solid CTAB was added to extraction buffer, raising CTAB concentration to 6%. Further incubation was necessary dissolve the CTAB. Both fresh (T54, T55) and lyophilized (T56, T57) were used, in 100mg (T54, T56) and 500mg (T55, T57) amounts. A chloroform:octanol deproteination step takes place, then the upper phase receives 0.1 volume of 10% CTAB. After a second chloroform:octanol extraction and transfer of the upper phase to a new tube, 3 volumes of 50mM Tris-HCl pH 8.0, 10mM EDTA, 1% CTAB were added to the aqueous phase. The concentration of CTAB here is maintained, but, since not salt was added, the ionic strength of the solution decreases from 2M NaCl to 0.5M. In low ionic strength, CTAB precipitates nucleic acids during a 30-minute incubation. The pellet formed after the incubation and successive centrifugation, the supernatant is discarded and the pellet dissolved in 0.5 volume of 1M NaCl. Prep was treated with RNAse and downstream stages of DNA precipitation by alcohol followed as the standard procedure cited above.

24. Increase in CTAB concentration to 6% as above, with the difference that here CTAB was not added as powder, instead as equal volume of 10% CTAB, 2 M NaCl. For DNA precipitation, 3 volumes of 6% CTAB, 100mM Tris-Hcl, 25mM EDTA were used, decreasing concentration of NaCl to 0.5M. Both fresh (T58) and lyophilized (T59) leaves were used.

25. Pea: extraction buffer: 0.7M NaCl, 1% CTAB, 50mM Tris-HCl pH 8.0, 10mM EDTA pH 8.0, 1% 2-mercaptoethanol, 0.01% sodium bisulfite. Departs from Murray and Thompson protocol in that DNA precipitation is achieved only by addition of ethanol, and not by decreasing salt concentration. All protocols bellow counted with precipitation methods that differ from the first proposed by Murray and Thompson. Different tissue-to-buffer rations were tested by extracting DNA from 10mg (T60), 50mg (T61), and 100mg (T62) of tissue, keeping the extraction buffer volume constant at 7ml.

26. Sugarcane [Aljanabi, 1999]: 200mM Tris-HCl, 50mM EDTA, 2.2M NaCl, 2% CTAB, 0.06% sodium sulfite, pH 8.0; after homogenization of the tissue and buffer, 0.5 volume of each 5% sarkosyl, 10% PVP, and 20% CTAB were added, elevating the CTAB concentration from 2% to 5% and decreasing NaCl concentration to 0.8M. Plant tissue: ‘Strawberry Festival’, fresh, mature, leaves of ‘Strawberry Festival’ (T63) or ‘Sweet Charlie’ (T64), 3.5g, 4ml buffer/g tissue.

27. Cacti (de la Cruz et al., 1997). Combination of CTAB and SDS extraction buffers. CTAB buffer: 100mM Tris-HCl pH 8.0, 20mM EDTA pH 8.0, 4% CTAB, 1.7M NaCl, 4% PVP,

5mM ascorbic acid, 10mM 2-mercaptoethanol. STE buffer: 100mM Tris-HCl pH 8.0, 50mM EDTA pH 8.0, 100mM NaCl, 10mM 2-mercaptoethanol. Fresh 100mg of 'Strawberry Festival' leaf tissue were ground in liquid nitrogen and subsequently incubated for 10 min at 65ºC with 1ml CTAB buffer. STE buffer (4ml) and SDS (to final concentration of 2%) were added and shaken vigorously for 7 minutes. A second 10-min incubation at 65ºC was carried out; 1.25ml of cold 5M KOAc was added and incubated in ice for 40 min, centrifuged at 3,500rpm for 10 min T65). The upper phase was transferred to a new tube and the nucleic acids precipitated by isopropanol. An alternative method (T66) substituted the addition of KOAc and ice incubation by addition of equal volume of 24:1 choloform:octanol, keeping steps after centrifugation the same.

28. Extraction buffer/plant material Incubation temperatures and durations were tested: 4ºC (T67-T70), 20ºC (T71-T74), 42ºC (T75-T78), 65ºC (T79-T82), and 0min (T67, T71, T795 T79), 5min (T68, T72, T76, T80), 30min (T69, T73, T77, T81), 60min (T70, T74, T78, T82). CTAB buffer (2% CTAB, 1.4M NaCl, 100mM Tris-HCl pH 8.0, 20mM EDTA pH 8.0, 1% 2-mercaptoethanol) was incubated in water baths with the various temperature treatments. When the buffer reached temperature equilibrium with the water baths, each tube received 1.6g of liquid nitrogen-ground strawberry leaves and the mixture was incubated at the various duration treatments. When incubation duration was reached, an aliquot of 10ml of the temperature treatment was mixed with chloroform. For incubation time "zero", an aliquot was taken right after buffer and ground tissue were mixed and chloroform was added. Samples were centrifuged at 4,000rpm for 10min. The upper phase was transferred to a new tube and nucleic acids were precipitated by isopropanol. After centrifugation and discard of the supernatant, the dry pellet was resuspended in water and treated with RNAse. The precipitation steps were repeated to obtain virtually RNA-free DNA. DNA was quantified with aid of a NanoDrop ND-1000 spectrophotometer. This experiment was repeated 3 times.

29. Tissue was ground in liquid nitrogen, and an aliquot of the extraction buffer (2% CTAB, 1.4M NaCl, 100mM Tris-HCl pH 8.0, 20mM EDTA pH 8.0, 1% 2-mercaptoethanol) was combined to the ground tissue to undergo further grinding and formation of slurry. The tissues tested were unexpanded (T83) and expanded (T84) leaves from the *F. nubicola* FRA520. Following formation of slurry, equal volume of 24:1 chloroform:octanol was added, vortexed, samples were centrifuged, and upper phase transferred to a new tube. Nucleic acids were precipitated by addition of 70% isopropanol, the alcohol was decanted, and the dry pellet resuspended in water.

30. An experiment was designed to contrast the traditional method of grinding tissue in liquid nitrogen, then adding the powder to buffer (T85) *versus* grinding tissue in liquid nitrogen, then adding the buffer (described immediately above) to the tissue and further grind until slurry is formed (T86). The 100mg per treatment of FRA520 plant material was mixed before nucleic acid isolation to eliminate the leaf age factor. A NanoDrop was used to quantify the nucleic acid content.

31. A factorial experiment tested interactions between formation (T87, T89) or not (T88, T90) of slurry and incubation temperatures of 4ºC (T87, T88) and 60ºC (T89, T90). After grinding the tissue (50mg per grinding method) in one of the two fashions tested, the material was split to be incubated for 1 hour in the two different temperatures. The downstream steps were followed as described above, including quantification of nucleic acids and absorbance at 230nm and 280nm by a NanoDrop ND-1000 spectrophotometer.

32. CTAB buffer concentrations of 2% (T91), 6% (T92), and 20% (T93) were tested. The slurry was formed by breaking down 400mg of tissue in liquid nitrogen first, then adding 2 ml of buffer for further grinding. Once homogenization was achieved, another 8ml of buffer were added and the mixture was incubated at 65ºC for 30min. The 10ml of buffer were split into 2 tubes (treatment replications) and 5 ml of chloroform:octanol were added to each tube. Nucleic acids from centrifugation upper phase were precipitated by isopropanol and the dry pellet resuspended in 1ml TE pH 8.0. Samples were quantified by NanoDrop.

33. Because homogenizing tissue in buffer seemed to have a positive effect on DNA recovery, an experimented was set up to test Polytron homogenizer speeds (half maximum speed—T95-T98; full speed—T99-T103) and duration of homogenizing treatment (no polytron—T94; 5 seconds—T95, T99; 15 seconds—T96, T100; 30 seconds—T97, T101; 60 seconds—T98, T102; 120 seconds—T103). Enough Laboratory Festival #9 tissue for all treatments (2g) was ground in liquid nitrogen and, by adding an aliquot of the buffer, ground to a paste consistency. The paste was divided into 10 tubes and enough buffer to reach 5ml was added to each tube just prior to treatment with Polytron. Samples were incubated at 65ºC for 30min. Downstream steps from incubation were as described above.

The final strawberry DNA extraction protocol is listed bellow.

Strawberry DNA Extraction Protocol

CTAB extraction buffer – 100ml

2% CTAB	2g
1.4M NaCl	28ml of 5M NaCl
100mM Tris-HCl, pH 8	10ml of 1M Tris
20mM EDTA pH8	4ml of 0.5M EDTA
1% BME	1ml
diWater	to 100ml

Tissue-to-buffer ratio = 40 mg/ml. For 12-ml tubes, maximum tissue processed is 200 mg.

- Grind 200 mg of liquid-nitrogen frozen leaves (young or unexpanded) in mortar-and-pestle
- Add 2 ml extraction buffer to ground sample, macerate in mortar until consistency of paste is achieved. Transfer the paste to a 12-ml tube, and add 3 ml buffer
- Homogenize utilizing a Polytron at full speed for 2 min
- Incubate for 1h at 65ºC, with intermittent agitation
- Add equal volume (5ml) of 24:1 chloroform:octanol
- Mix by shaking vigorously
- Centrifuge at 4,000 rpm, 5 min
- Transfer the upper, aqueous phase to a new 12-ml tube
- Precipitate DNA with equal volume of 70% isopropanol
- Mix by inverting the tube several times
- Centrifuge at 4,000 rpm, 5 min
- Discard the supernatant
- Air-dry nucleic acids pellet

- Resuspend pellet in 500ul to 1 ml (depending on the amount required to dissolve the pellet) of deionized water or TE pH 8.0.

APPENDIX B
In silico ANNOTATION AND DISTRIBUTION OF *Fragaria vesca* GENES

Under each fosmid name is a list of numbered potential genes predicted by FGENESH. The nucleotide intervals that had protein hits by BLASTP were used for a similarity search against the non-redundant Viridiplantae, protein database using BLASTX. The best matches identified by the algorithm are listed under “Protein Hit”. Threshold value was 10^{-15}. Letter “X” under “Protein Hit” denotes no similarity was detected in the protein database. Under “Orientation”, “+” signs signify that the query sequence is translated in the same direction it was input, where negative orientation signifies that the complement strand is translated. “EST Hits” are sequences of DNA for which an EST was detected within Rosaceae, with a minimum length of 100 nucleotides and 95% identity.

Gene distributions were calculated by dividing each fosmid insert size by the number of genes either predicted by FGENESH or identified by similarity to the non-redundant Viridiplantae protein database.

Simple Sequence Repeats (SSRs) with at least 5 repeats of a motif are represented by color-coded triangles:

△ in FGENESH-defined genic sequence; ◣in FGENESH-defined intergenic sequence

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
01L02	13	7				40,302	3.1	5.8
1			unknown	-				
2 ◣			X					
3 △			pectin lyase	+				
4			unknown	-				
5			beta-glucan binding	-				
6			enolase	-				
7 ◣			X					
8 △△△ ◣			X					
9 △△			unknown	+	DV440436.1			
10 ◣◣			X					
11 ◣			X					
12 ◣			X					
13			unknown	-	DY670952.1			
05N03	8	5				34,611	4.3	6.9
1 △ ◣			ATP binding/adenylate cyclase	+				
2			X					
3 ◣			Senescence-associated	+	CX661421.1			
4			hypothetical	-	DW248990.1			

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
5			X		CX661657.1			
6			X		CO816931.1			
7 ◣			peroxidase	-				
8			unknown	-				
11D02	9	3				37,961	4.2	12.7
1			ATP synthase, mitochondrial	-	DY668653.1			
2			X					
Not predicted			X		DW342667.1			
3			X		DW344738.1			
4			Release Factor 2, chloroplast	+	DW346600.1			
5			X					
6			X					
7			X					
8			X		BQ105541.1			
9			heat shock binding	-				
13I03	8	8				37,707	4.7	4.7
1			hydrolase	+				
2			leucyl-tRNA synthetase	-				
◣ 3 △△			leucyl-tRNA synthetase	-				
4 △△			leucyl-tRNA synthetase	-				
◣ 5			leucyl-tRNA synthetase	-				
6			zinc finger family	+				
7			2OG-Fe(II) oxygenase	-	DY670360.1			
8			integrase	+	DY671649.1			
15B13	7	2				23,212	3.3	11.6
1			senescence-associated	-	CX661347.1			
Not predicted			26S ribosomal RNA (not a protein)		CA854088.1			
2			X					
3			senescence-associated	-	CX661347.1			
4			X					
5			X		CX661421.1			

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
6			X					
7			X		CX661657.1			
17O22	9	6				34,090	3.8	5.7
1			homeodomain	+				
2△△			X					
3△ ◣			X					
4			oligopeptidase	+	BQ104655.1			
5			hypothetical	-	DY675330.1			
6			**X**					
7△ ◣			unknown	-				
8△ ◣			lectin protein kinase	+				
9△			hypothetical	-				
18A19	7	6				40,908	5.8	6.8
1 ◣◣			cytochrome P450	-				
2 ◣			X					
3			integrase	-				
4			integrase	-				
5			integrase	+				
6			integrase	-				
7			transferase	+				
22H18	8	4				37,851	4.7	9.5
1 ◣◣◣			X					
2			X					
3			polyprotein	-				
4△ △			X					
5 ◣			hypothetical	+				
6			X					
7△ ◣◣◣			unknown	+				
8 △			pre-mRNA processing factor 38	+				
22L05	8	3				35,112	4.4	11.7
1			X					
2			X					
Not predicted			X		DY674519.1 EST starts upstream of predicted gene 3, and spans oxidoreductase			
3			oxidoreductase	+	DY671565.1			
4△			oxidoreductase	+				
5 ◣			sulfate transporter	-				
6 ◣			X		CO380067.1			
7			X					

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
8			X					
27F10 ◣	11	8				37,110	3.4	4.6
1			kinase	-	DY675883.1			
2 △			hypothetical	-	CX661386.1			
3			unknown	-	DV438706.1			
4			integrase	-				
5			integrase	-				
6			integrase	-				
7 △ ◣			unknown	+				
8			X					
Not predicted			X		CO378700.1			
9			unknown	-				
10			X					
11			X					
29G10	10	4				31,681	3.2	7.9
1 ◣◣◣◣◣			transposase	-				
2			X					
3			flavin-binding monooxygenase-like	+	DY673408.1			
4			X					
5			X					
6			X					
7 △			X					
8			phenylacetaldehyde synthase	-				
9			unknown	+				
10 ◣◣			X					
30I24	7	5				37,599	5.4	7.5
1			X					
2 ◣◣◣◣			wall-associated kinase	+				
△△△ 3			X (E value=1e-10) arabidopsis response regulator 12					
4 ◣			chitinase	+	CX661529.1			
5			arabidopsis response regulator 12	+	DY671913.1			
6 △			transferase	-				
7			PICKLE chromating remodeling factor	-				

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
32A10	15	4				33,577	2.2	8.4
1			catalytic/ hydrolase	-	DY667800.1			
2 △△			X					
3			X					
4 ◣			X					
5			X					
6			X					
7			copper ion binding	+				
8 △△			X					
9			MADS-box	-				
10			X					
11			X					
12			X					
13			pathogenesis-related	-				
14 ◣			X					
15			X					
32L07	6	4				32,951	5.5	8.2
Not predicted			X		DY668002.1			
1			hypothetical	-				
2			SMC2	-				
3			disease resistance	-	DY666677.1			
4			X					
5			exostosin-like	-	CX662049.1			
6			X					
34D20	8	6				30,034	3.8	5.0
1 ◣ △△			RNA recognition motif	+				
2			cysteine-type peptidase	+				
3			X					
4			transposase	+				
5			anthocyanin 5-aromatic	-				
6			X (E value = 8e-14) anthocyanin malonyltransferase *FGENESH missed EST*	+				
7			NAC domain NAM	-				
Not predicted			X		DV438498.1			
8			X					

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
38H02	7	6				31,669	4.5	5.3
1 ◣◣◣◣			X					
2 △△△			transposon protein	+				
3 ◣◣			cytochrome P450	+				
4			cytochrome P450	+				
5			integrase	+				
6			serine/threonine kinase	-				
7 △			exportin	-				
38H05	11	1				32,050	2.9	32.1
1			X					
2 ◣			X					
Not predicted ◣			X		dbj|AB208565.1			
3			retrotransposon polyprotein	-				
Not predicted			X		dbj|AB208565.1			
4			X					
5			X					
6 ◣			X					
7			X					
8 ◣◣			X					
9			X					
10			X					
11 ◣◣			X					
40B22	9	8				36,230	4.0	4.5
1			unknown	+				
2			cyclin-like F-box	-				
3 △			X					
4			cyclin-like F-box	+				
5			cyclin-like F-box	-				
6 △			cyclin-like F-box	-				
7			Arf GTPase activating	-				
8 ◣◣			heavy metal transport/detoxification	+				
9			MuDR family transposase	-				
40M11	9	5				31,718	3.5	6.3
1			X					
2			cyclin 1-like F box	+				
3			X					
4			X					

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
5			X					
6			Secretory Protein SEC14	-	DY675900.1			
Not predicted			X		DY672841.1			
7			ATPase	-				
8			unknown	+				
9 ◣◣			glycosyl hydrolase	-				
Not predicted			X		CX662188.1			
43P07	10	4				43,641	4.4	10.9
1			X					
2			retrotransposon polyprotein	+				
3 ◣◣			X					
4 ◣◣◣			X					
5			DNA cytosine-5-methyltransferase	-	DY668476.1			
6			unknown	+	DY668476.1			
7 ◣◣◣◣◣◣◣			methyltransferase small domain	+				
8			X					
9			X					
10			X					
44J07 ◣	11	2				29,636	2.7	14.8
1			X					
2 △△			X					
3 ◣			X		DY672792.1			
4			X					
5 △△◣			X					
6			X					
7 △◣			disease resistance	-				
8 △			unknown	-	DY671343.1			
9			X					
10 △			X		DY650877.1			
11			X					
47H15	9	4				34,817	3.9	8.7
1 △◣			X		DY669025.1			
2 △◣			X					
3			polyprotein	-				
4			integrase	-				
5			retrotransposon protein	-				

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
6			X					
7 ◣			X					
8 ◣			X					
9 ◣			heat shock	+				
52E09	9	8				36,230	4.0	4.5
1			unknown	+				
2			cyclin-like F-box	-				
3 △			X					
4			cyclin-like F-box	+				
5			cyclin-like F-box	-				
6 △			cyclin-like F-box	-				
7			Arf GTPase activating	-				
8 ◣			heavy metal transport/detoxificati	+				
9			transposase	-				
63F17	6	3				28,318	4.7	9.4
1 △△			phospholipase D	+	DY672511.1			
2 ◣			unknown	-				
3 ◣			binding	+				
4 ◣◣◣			X					
5 △			X					
6 ◣			X					
72E18	12	11				36,293	3.0	3.3
1			hydrolase	+				
2 △			hydrolase	+				
3			reverse transcriptase	-				
4			hydrolase	+	DV438212.1			
5 △			X					
6			hydrolase	+	DY669358.1			
7			unknown	+	DY675437.1			
8			transferase	+				
9			spliceosome-associated	+				
10			unknown	+				
11			actin 7, actin 11	-				
12			glycoprotein-like	-	DY670963.1			
84N10 ◣	8	2				40,183	5.0	20.1
1 △△			ribosomal L24/L26	+				
2 △△△△△			X					
3			X					

Fosmid	Predicted Number of Genes		Protein Hit	Orientation	EST Hits (gb no.)	Fosmid Insert Size (bp)	Putative Gene Distr (kb between genes)	
	ab initio	Similarity					*ab initio*	Similarity-based
4			ATP binding	-				
5 ◣			X					
6 △			X					
7 △			X					
8 ◣			X					
Totals	235	129				905,491		
Means	9	5				34,827	4.0	9.1
Sample Standard Deviation	2.1	2.4				4,426	0.9	6.0

APPENDIX C
PCR PRIMERS USED TO AMPLIFY AND SEQUENCE GENE-PAIR HAPLOTYPES

GPH name	Primer name	Primer sequence
GPH4	GPH4a	ACGAGGGCTTGGAAGAAAGG
	GPH4b	GCCCAACAACAGAAAGACC
GPH5	GPH5#2a	CAATGCCATGGTCTCCGGTC
	GPH5#2b	TGCCGTTGCACACACCTTCC
	GPH5A2	GCTCTTTGGTGTTCAAAGTTGGAT
	GPH5B#2	ATCCAGCCAAACTGAAGGTG
	GPH5A3	CAGCCATGAAGTCAAGGTCA
GPH10	GPH10a	GGCTTCTTCTTGTCCGGCAGC
	GPH10b	GAACTCCAGGTCAGATCTTCG
	GPH10c	CTCGCTGCAAATCAGCTACC
	10ABCol7Rev2	GAGTTTGTCGAGCTGATC
	10ABCol32Rev2	ATAGAGGCGATGTTGTAG
	10AB#3	GGCCCTGATCACTCGACA
	10AB#4	GGTTTGGTTGGTTAAGGTG
	10AB#5	GACAGTACCTGAAAATTTGG
	10AB#6	AAGTATCATTAACAGGC
	10AB#7	ATCATATATGCGGGTGTG
	10AB#8	TAACGAGCAGTGGCGG
	10AB#9	ATCACCTCTACTCCCACGC
	10AB#10	CACCGTAACAGCTGAGCAAG
	10AB#11	ACACAAATGCCTCATCCACA
	10AB#12	ACTAAAGCCCAGCAACCCTC
	10AB#13	TTCTCTGTCAACCCTGCCTT
	10AB#14	GGGGCAAAGTTTACATAGCA
	10AB#15	AACTCGCCGGAAGACACTTA
	10AB#16	GCCGGAAGACACATATCGAT
	10AB#17	GCATCCCCTTTACATCCAAA
	10AB#18	GTTAGAGACGACGACGGGAG
	10AB#19	TGCCTGGCAAAGTAAACCTC
	10AB#20	GGCGTGTCAATTTGTGAATG
	10AB#21	TCATCTTCCTCTGTATGCGACT
	10AB#22	GGTTTTGTTTTTGGTGGGAA
	10AB#23	GTCGAGTGATCAGGCCGTA
	10PPR1F	AACGGAGAAGAAGACTGTCG
	PPR1R1	GATCGAACGGCTGATATTAAA
	10PPR1R2	TGTAGCTCATACTTTTGTTCTC
GPH23	GPH23F	CTTGAGGGCCATCAGCAC
	GPH23R	TACACCCACGCCTTCATCTC
	GPH23F2	GAACTGCGAAGATCTATCTGA
11D02	11D02F	GAGCTGCTGTGTGAACCAAA
	11D02R	GTTCAACTCCAGATGAAGTGAGG

GPH name	Primer name	Primer sequence
17O22	17O22F	AAAATGGGTTGCACGAGTTC
	17O22Rb	GGGTTTCCTCACAAACTTCG
27F10	27F10F	CCTGCAGGGTTTTTCATCAT
	27F10R	TGGAAATGTATTCTGGTTCTCC
29G10	29G10F	TGGCCTTGTTTCCTAAACTCTT
	29G10R	AGAAGAAGGCAGCACCCAAT
32L07	32L07F	GAGTTGAAAAACGGGTCGAA
	32L07Rb	AGGAAAATGCGGGAGAAAGT
34D20	34D20Fb	GCAGAAAGAAACTGATGTGCTT
	34D20Rc	CGCAGTCGTAAAAATTCGTCT
	34D20Fb2	TGGGTGTGGATGAACTATACG
40M11	40M11Fd	CAACATTTTGGTGGCCTTCT
	40M11Rd	CGGCCTATGAAACCACAGTT
63F17	63F17F	GCAGAAAGAAACTGATGTGCTT
	63F17R	CGCAGTCGTAAAAATTCGTCT
72E18	72E18Fb	GCAGAAAGAAACTGATGTGCTT
	72E18Rb	CGCAGTCGTAAAAATTCGTCT
	72E18Fb2	GCAGCAATCAAATCATTCCA

APPENDIX D
SEQUENCES GENERATED DURING CHARACTERIZATION OF “GENENPAIR HAPLOTYPES”

>GPH4_ananassa_13
ACGAGGGCTTGGAAGAAAGGAGGTCAATTTGGTTAAGGTGTGTTGGAGTCGCCAAGTTGAGGGTGATGCATTCTTGG
GAGTTAGAGTCGGATATGAGGGCTAAGTACCCCAAGTTGTTTCTTTTTGAGTTAGTATCTTAAAATTTCGGGGACGA
AATTTCTTTAAAGAGGGTAGAGTGTAATACCCCAGAAATTTGATATTAGTTTCTAATTTTATTTAGGAATTTTTGAG
TTAGAAGTTAGCGTGTTTTGAAGTTTGAAGGAAGAACGGAAGGGTTCGGATGCATAAATTGCTGAACCGGTTTTATG
GTTCTGAAAGGTCAAGAGTTGACTTTCTAATCCGTTGGGTTTCTCGAGAAACTTCCTTCACGGAAGTTGTAGAGCAC
GACGATACGAGTTCGTAGACACGTGGCACGCGTAAAACGGACTTCGTATGAGAAAGTTATGGTCAGCAGAAGTTGTG
GCTTTTCGGGAATATTTAGGTTAAATAGGAAATTTTCGTTTTGGGTTCTATTATTTTTCAGAAATTCCTTTCTTCCC
CTTCTTCTCTCTCCCCGACCCCGAGAGAACCCAAGCTTCCCAGCCGACCCGACCCGGACCCGGTTGACCCCACCCGG
ATTTTCCGGCCATCTCCGGCCGACCCAGGCAACGGCACTGGTCGGGTTCTCTTCCTCTCCTCCGTCAGAGCTGACCT
GTGGCGGTGATGTGCGCCGTTTCGGTCCCGAGGTGGTGACCCGAAGCTCGGAAGTTCGGGTTAGAGACCCGGTCGGA
TTCCTTCATCCGGCGGCGGCGACAAGGTAGGAGACCGATCGGGATAGAAACCCCTTGGCGTCTTGGTCCGATTGCTG
GTGGCCTTGAAGTGCGACACACGGCGGAGAGTGGCAGTGGTGGATCCTAATCTTTCCGGCGAGTTTCCGGCAAATTC
CCGGCCGATTTGGTTTCGACCTCAGGTATGGAAGTTGCTCTCCTTGCTCTGAGCTATATTTTTGGTGTTGGAAGTTT
GTCCGTTTTCGAAGGTTAGTGGGGTGGCGCGTGGGACCCACGTGCAGTCGCTAAGGGCAGCGCGTAGCGGCGCGTCC
GTAGGTGGTTGTGGTCTTTCTGTTGTTGGGC

> GPH4_ananassa_15
ACGAGGGCTTGGAAGAAAGGAGCCTTTGTCTGTGAATATGTTGGGGAGATAATGACCTACAAGTACTTGTATAATCG
GGGAAGACACACATACTCAATCACACTGGATGCCGGTTGGGGCATTGCTTAGCCTTGTGTTTTTGGTTAAGCAGATC
ATATTTCCCAGAATTAGATCTGCATAATAATCCATTTTCACGCCAGTCATTTGGTGCCTGTGGTTTAGAACTTAGAT
TTCAGAAGTTCTATCAAGTTTGTCACTTCCTCACCTCTTGTGATGAGAGAAATTTTCAATTCGTTGATGTTGACAAA
GATCATCTGACTATAAATTTGCCCATGTAATCGTATTCTGTTTCTCCTAAAAATATTGTTCTTGTAAATTTGGGGAA
ATCCGGAAAAGGCTATACTGTCATTTGCTTCCTAACTTGTCTTGAGCAATGACCTAATGATTTTCCTATAGCTTTTG
TTGGTTTTTTTCTCGTTCTTTCTTTCTCTGACGTTATGTTTAATTCCCTCAACAACTCCAGATGCTATGATGAAAAC
TTGGTTGATATCCCAGTTCAAGTGGATACTCTTGCTCGCTATTATTACCATGTATGAATTTGGCTGCTTCCTTTCTA
AATGGTCTTTCTGTTGTTGGGC

>GPH5_vesca_clone21
CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATCAT
CAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAGTGAAGACAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATTGTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAAAATGGGATCTACAG
AAAAGAGCTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGTGAGCAGGTTAGCTAGAAGCTT
CAAACAAAGCGTCAATTGCCCACAGTTATTCTTTGATAGATATATGTTGAACTGTAAGAGACATATTTCAAGCTCTT
TGGTGTTCAAAGTTGGATTCAATTACATGTAGACACAGTTACCATTTTCCCATATGAAATAGAAGGTAATATGCATG
ATATAAATATCTAGTTAATTGTACAATGATATTTGTAACCAGTGAAAATAATGACAATCTTTATAACAAAATTTCAG
TTATCTTTCCATTGCTGTATGAACTGTTACCATTAGCCTCTCACACAAGAACAACAACACCAAACAAACAGAACCAG
ACCAAATCACACCAATATAAAACAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGAAGACAA
AGAATCCTTTTGTCATATGGATTGAATCTGAATTAGTTGGAGTGTTTCTGGCTGTCATATCTCATATGCAGGCATGT
TACATGTCTCATGATGTCTTCATTTGGTGACAAAAGCTAAATCTTAACCTGACCTAAGTATCAAGACATATTGGACA
ATTGGGCTTAATCATAGTCTAAGCCCAAATCTGTACTAGCCCATAATATGCTTTTTATAGAAAACACTCTGTGATCT
TCACCATTGAGGAGTCAAGTTACTCAGCCCTGAAGTAAAAGTCCAGTCAGTAGTGCAGTTGAGTTCAACTTGTTCTG
GGTTCTTCAAAGTTTGAAACTTTAAGCTTCGATGGAGGAAGAGAAGGATGCCTTTTATGTTGTTCGAAAGGGAGATG
TGGTTGGCATATATAAAAGCCTGAAAAGATTGCCCAAAACCCAAGCTGGGTTCATCCGAAAAAGTTTTTGAATCTTT
TTAAAGCCCTTTTTTAATAATTTGGAATNCCACCTTCCnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnGTATTAGATGAAAACTAGTTTT
TCTAAGAACTTGATGAGTTGATGGAGGATTACATATGAGGTTTGGTTATGTTTTTAGGTATGCAATCCTTCTGTAAG

```
TGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGAGTACCTTGTCTCGCATGGGCTTAAGAATGCTTCATATA
CTATCAGTGCCAGTGATGTGAAAGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATGAT
TTCATGTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAACTAGTTGCATATGATATTA
AATCTTTTGTTTGGTTCCTCATTGTATAATTTGGTTCTTTATTATAATCTTTCAGCAGCCAGCATCTTCTATGGTTA
ATTCAGGGTTCAAAATTGCACCTGATCAGCTTACACGAAAGAGAATGCCGGATGTGATAAATTCAGGTGTCAATGAC
CCTCCACAAAAGAGATCACTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTATGATAGAAGATCTTCTAT
TTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATTTTTCAGCAGCCAGCGTCTTCTATGGTTAATTCAGGCTT
CAAAATTGCACCTCATCAGCTTACGCGAATGAGATTGCCGGATGTGGTAAATTCAGGTGTCAATGACCCTCCACAGA
GGACATTGCCGGATGTAAGTATCTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGATACC
TTCAGTTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGATGTTTTTTGGCTT
CCAGTTTTTCTTCACATAAGCATTTTAAAGCTGATCATTGTAATCGAACTCGAATTATTCTACTACTGGTGTAAGTT
GCCTTGTGTCACCACCACTAAGATCACAATTTCGTATTTTATGATCAACACCGAAGACCTATGTCTAGTGTCGTGAT
TATGGTCATGTGAAGTGGATTTCTTAATATATGCCTCGTCTATGTCTTCATCCAGCAATCCTGCATACTTGAGTTTG
ATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGTGCAGGAGCTGTACTTCGTGCTGAAGATGGGAGTGTTGTATGT
GGAGTTCATGAAAACATTGTGAATTTTTTTGATATATATTTTTGTTTTTGTAAAAATGGATCTCTTCATAACATTGG
GGTTACTATAGTTGCACCGGCTGCGGGAAGGTGTGTGCAACGGCA
```

```
>GPH5_viridis_clone5
CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATCAT
CAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAGTGAAGACAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATTGTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAAAAATGGGATCTACAG
AAAAGAGCTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGTGAGCAGGTTAGCTAAAAGCTT
CAAACAAAGCGTCAATTGCCCACAGTTATTCTTTGATAGATATATGTTGAACTGTAAGAGACATATTTCAAGCTCTG
GTGTTCAAAGTTGGATTCATTTACATGTGACACAGTTACCATTTTCCCATATGAAATAGAAGGTAATATGCATGATA
TAAATATCAAGTTAATTGTACAGTGATATTTGTAACCAGTGAAAATAATGACAATCTTTATAACAAAATTTCAGTTA
TCTTTCCATTGCTGTATGAACTGTTACCATTAGCCTCTCACACAAGAACAGCAACACCAAACAAACAGAACCAGACC
AAATCACACCAATATAAAACATAATTGGATTTTCATGAAAGGCAGCAAGGCATGATCAATGAAGGAGAAGACAAAGA
ATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATCTCATATGCAGGCATGTTAC
ATGTCTGCATTTGGTGACAAAAGCTAAATCTTAAGAATTAAGACATATTGGACCATTGGGCTTAATCATAGTCTGAG
CCCAAATCTGTACTAGCCCATAATATGCTTTTTATAGAAAACACTCTCTGTGATCTTCGCCATTGAGGAGTCAAGTT
ACTCAGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAGTTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAAC
TTTAAGCTTCAATGGAGGAAGAGAAGGATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGC
TTGAAGGATTGCCAAAACCAAGCTGGTTCATTGGTAAAAGTTTTGATCTTTTAAGCCTTTTATAATTTGATCACTCT
CATTGTTTTATCAATTnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnGAACTTGATGTTATTTACTTGATGAGTTGACGGAGGATTACATATAAGGCTTGGTTATG
TTTTTAGGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGAGTACCTTGTCT
CGCATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAAAGATGGTCTGTTTGGAAGCCTTGTTGCT
TGTCCTTACCAGGTTTGAATTGATGATTTCATGTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTG
GCGTGGAGCTAGTTGCATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAAGTCGGTTTTTTATTATAATCT
TTCAGCAGCCAGCATCTTCCATGGTTAATTCAGGGTTCAAAATTGCACCTAATCAGCTTACACGAAAGAGAATGCCG
GATGTGGTCAATTCAGGCGTCAATGACCCTCCACAAAAGAGATCGCTGGATGTAAGTATCATATGCTACATGGAACT
TTTGTAGTTTGATAGAAGATCTTCTATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCTTTCGGCAGTC
AGCATCTTCTATGGTTAATTCAGGCTTCAAAATTGCACCTTACCAGCTTGCACGAATGAGATTGCCGGATGTGATAA
ATTCAGGTGTCAATGACCCTCCACAGAGGACATTGCCAGATGTAAGTATCTTATGCTACATGGAAATTTTGTTCTGG
CCAGATTGGCATGAAAATCCAGATACCTTCAGTTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATT
TATTCTGCAAATGATGTTTTCAGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGTAATCGAACTC
GAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGGTCACAATTTCGTATTTTATGATCAACAC
TGAATACCTATGTCTAGTGTCGTGATTATGGTCATGTGAAGTGGATTTCTTAATATATGCCTCATCTATGTCTTCAT
CCAGCAATCCTGCATACTTGAGTTTGATGGAGCTTCAAAAGGAAGTCCTGGATTATCTGGTGCAGGAGCTGTACTCC
GTGTTGAAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATATATTTGTTTTTG
TAAAAATGGATCTCTTTATAACATTGGGGTTACTGTAGTTGCACCGGCTGCGGGAAGGTGTGTGCAACGGCA
```

```
>GPH5_iinumae_clone5
```

CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATC**C**T
CAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAGTGAAGACAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATTGTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAG
AAAAGAGCTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGTGAGCAGGTTAGCTAAAAGCTT
CAAACAAAGC**A**TCAATTGCCCACAGTTATTCTTTGATAGATATATGTTGAACTGTAAGAGACATATTTCAAGCTCTT
TGGTGTTCAAAGTTGGATTCAATTACATGTAGACATAGTTACCATTTTCCCATTTGAAATAGAAGGTAATATATCAA
GTTAATTGTACAATAATATTTGTAATCAGTGAAAATAATGACAATCTTTATAACAAAATTTCAGTTATCTTTTCACT
GCTGTATGAACTGTCACCATTAGCCTCTCACACAAGAACAACAACACCAAACAAACAGAACCAGACCAAATCACACC
AATATAAAACAGAATTGGATTTCCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGAAGACAAAGAAACCTTTTGT
CATAGGGATTGAACCTGAATTATCTGGAGTGTTTCTGGCTGTCATATCTCATATGCAGGCATGTTACATGTCTGCAT
TTGGTGACAAAAGCTAAATCTTAAGAATTTAGACGTATTAGACCATTGGGCTTAATCATCGTCCGAGCCCAAATCTG
CACTAGCCCGTAATATGCTTTTTATAGAAAACAGACTCTCTGTGATCTTCGCCATTGAGGAGTCAGGTTACTCAGCT
CTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAGTTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAACTTTAAGCT
TCAATGGAGGAAGATAAGGATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGCTTGAAGGA
TTGCCAAAACCAAGCTGGTTCCTCGGTAAAGTTTTGATCTTTTAAGCCCTTTATAATTTGATTACTCTCATTGTTTT
ATCAATTTTTGATTTCCCATTTGATnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnACTA
GTTTTTCTAAGAACTTGATGTCATTTACTTGATGAGTTGATGGAGGATTACACATGTGGTTTGGTTTTGTTTTTAGG
TATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGAGTACCTTGTCTCACATGGG
CTTAAGAATGCTTCATATACTATCAGCGCCAGTGATGTGAAAGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTA
CCAGGTTTGAATTGATTTCATGTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAGCTA
GTTGCATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCGGTTTTTTATTATAATCTTTCAGCAGCCA
GCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCACCTAATCAGCTTACACGAAAGAGAATGCCGGATGTGGTAAA
GTCAGGCGTCAATGACCCTCCACAAAAGAGATCATTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTTTG
ATAGAAGATCTTCTATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCTTTCAGCAGCCAGCGTCTTCTA
TGGTTAATTCAGGCTTCAAAATTGCACCTTATCAGCTTACACGAATGAGATTGCCGGATGTGGTAAATTCAGGTGTC
AATGACCCTCCACAGAGGACATTGCCGGATGTAAGTATCTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCA
TGAAAATCCGGATACCTTCAGTTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAA
TGATGTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGTAATCGAACTCGAGTAATTCT
GCTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATTTCGTATTTTATGATCAACACTGAATACCTA
TGTCTAGTGTTGTGATTATGGTCATGTGAAGTGGATTTCTTAATATATGCCTCATCTATGTCTTCATCCAGCAATCC
TGCATACTTGAGTTTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGTGCAGGAGCTGTACTCCGTGCTGAAGA
TGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATATTTTTGTTTTTGTAAAAGTGGA
TCTCTTTATAACATTGGGTTTACTATAGTTGCACCGGCTGCGGGAAGGTGTGTGCAACGGCA

>GPH5_nubicola_clone7
CAATGCCATGGTCTCCGGTCTATTTCAAC**C**GGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATCAT
CAGAATCAGATTTGTTTGC**T**CTTGCAGAAAGAGAATCCTCGAGTGAAGACAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATTGTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAG
AAAAGAACTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGTGAGCAGGTTAGCTAAAAGCTT
CAAACAAAGCGTCAATTGCCCACAGTTATTCTTTGATAGATATATGTTGAATGAACTGTAAGAGACATATTTCAAGC
TCTTTGGTGTTCAAAGTTGGATTCAATTACATGTAGACACAGTTACCATTTTCCCATTTGAAATAGAAGGTAATACG
CATGATATAAATATCAAGTTAATTGTACAATGATATTTATAATCAGTGAAAATAATGACAATCTTTATAACAAAATT
TCAGTGATCTTTCCATTGCTGTATGAACTGTTACCATTAGCCTCTCACACAAGACCAACAACACCAAACAAACAGAA
CCAGACCAAATCACACCAATATAAACAGAACTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGAAGA
CAAAGAATCCTTTCGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATCTCGTATGCAGGCA
TGTTACATGTCTGCATTTGGTGACAAAAGCTAAATCTTAACATGACCTAAGAATTAAGACATATTGGACCATTGGGC
TTAATCATAGTCTAAGCCCAAATCTGTACTAGCCCATAATATTCTTTTTATAGAAAACAGAGATTCTCTGTGATCTT
CACCATTGAGGAGTCAAGTTACTCGGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAGTTCAACTTGTTCTG
GGTTCTTCAAAGTTTGAAGCTTTAAGCGTCAATGGAGGAAGAGAAGGATGCCTTTTATGTTGTTCGAAAGGGAGATG
TGGTTGGCATATATAAAAGCTTGAAGGATTGCCAAAACCAAGCTGGTTCATCGGTAAAGTTTCGATCTTTTAAGCCT
TTTATAATTTGATCACTCTCATTGTTTTAnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn

nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnTAAGAACTTGATGTTATTTGCTTGATGAGTTGAT
GGAGGATTACATATGAGGTTTGGTTATGTTTTTAGGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTG
CCTAAGGAGGCCGAGGAGTACCTTGTCTCGCATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAA
AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATGATTTCATGTGTTCTAGTTTCTGTT
TGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAACTAGTTGCATATGATATTAAATCTTTTGTTTGGTTCCTCAT
TGTATAATTCGGTTTTTTATTATAATCTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCACC
TGATCAGCTTACACGAAAGAGAATGCCGGATGTGATAAATTCAGGTGTCAATGACCCTCCACAAAAGAGATCATTGG
ATGTAAGTATCATATGCTACATGGAACTTTTGTAGTATGATAGAAGATCTTCTATTTGGTTACTCATTGTATGATTC
GGGTTTTTATTATAATTTTTCAGCAGCCAATGTCTTCTTTGGTTAATTCAGGCTTCAAAATTGCACCTCATCAGCTT
ACGCAAATGAGATTGCCGGATGTGGTAAATTCAGGTGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGTAT
CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGATACCTTCAGTTTGGCTGGATTATGGA
GTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGATTTTTGGCTTCCAGTTTTTCTTCACATAAGCATTT
TAAAGCTGGTCATTGTAATTGAACTCGAATAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATC
ACAATTTCGTATTTTATGATCAACACCGAATACCTATGTCTAGTGTCGTGATTATGGTCATGTGAAGTGGATTTCTT
AATATATGCCTTGTCTATGTCTTCATCCAGCAATCCTGCATACTTGAGTTTGATGGAGCTTCAAAAGGAAATCCTGG
ATTATCTGGTGCAGGAGCTGTACTCCGTGCTGAAGATGGGAGTGTTGTATGTGGAGTTCATGAAAGCATTGTGAATT
TTTTTTTATATATATTTTTGTTTTTGTAAAAATGGTTTTATAACATTGGGGTTACTATAGTTGCACCGGCTGCGGGA
AGGTGTGTGCAACGGCA

>GPH5_nilgerrensis_clone19
CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGAC**T**GGAAGATCAT
CAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAGTGAAGACAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATTGTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAAAAATGGGATCTACAG
AAAAGAACTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGTGAGCAGGTTAGCTAAAAGCTT
CAAACAAAGCGTCAATTGCCCACAGTTATTCTTTGATAGATATATGTTGAACTGTAAGAGACATATTTCAAGCTCTT
TGGTGTTCAAAGTCGGATTCAATTACATGTAGACACAGTTACCATTTTCCCATTTGAAACAGAAGGTAATATGCATG
ATATAAATACCAAGTTAATTGTACAATGATATTTGTAATCAGTGAAAATAATGAAAATCTTTATAACAAAATTTCAG
TTATCCTTCCATTGCTGTGTGAACTGTTACCATTAGCCTCTCACACAAGAACAACAACACCAAACAAACAGAACCAG
ACCAAATCACACCAATATAAAACAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGAGAAGAC
AAAGAATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATCTCATATGCAGGCAT
GTTACATGTCTGCATTTGGTGACAAAAGCTAAATCTTAACATGACCTAAGAATTAACACATATTGGACCATTGGGCT
TAATCATAGTCTAAGCCCAAATCTGTACTAGCCCATAATATGCTTTTTATAGAAACAGAGATTCTCTGTGATCTTCA
CCATTGAGGAGTCAAGTTACACAGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAGTTCAACTTGTTCTGGG
TTCTTCAAAGTTTGAAACTTTAAGCTTCAATGGAGGAAGAGAAGGATGCCTTTTATGTTGTTCGAAAGGGAGATGTG
GTTGGCATATATAAAAGCTTGAAGGATTGCCAAACCAAGCTGGTTCCTCGGTAAAGCTTTGATCTTTTAAGCCTTTT
ATAATTTGATTACCCTTATTGTTTTATCAAnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnGTTTTTCTAAGAACTTGATGTTATTTACTTGATGAGTTGATGGAGGATTAC
ATATGTGGTTTGGTTTTGTTTTTAGGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGG
CCGAGGAGTACCTTGTCTCACATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAAAGATGGTCTG
TTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATTTCATGTGTTCTAGTTTCTGTTTGGGTATCTGTTA
TTTTCATGGCATGTGGCGTGGAGCTAGTTGCATATGGTATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCGGT
TTTTTATTATAATCTTTCAGCAGCCAGCATCTTCTATGGTTGATTCAGGGTTCAAAATTGCACCTAATCAGCTTACA
CGAAAGAGAATGCCGGATGTGGTAAATTCAGGCGTCAATGACCCTCCACAAAAGAGATCGCTGGATGTAAGTATCAT
ATGCTACATGGAACTTTTGTAGTTTGATAGAAGATCTTCTATTTGGTTACTCATTGTATGATTCAGGTTTTTATTAT
AATCTCTCAGCAGCCAGCGTCTTCTATGGTTAATTCAGGCTTCAAAATTGCACCTTATCAGCTTACACGAATGAGAT
TGCCGGATGTGGTAAATTCAGNGTGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGCATCTTATGCTACAT
GGAAATTTTATTCTGGCCAGATTGGTATGAAAATCCAGATACCTTCAGTTTGGCTGGATTATGGAGTTGCGTTGATC
ACTTGTTTATTGTATTTATTCCGCAAATGATGTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGA
TAATTGTAATCGAACTCGAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATTTCG
TATTTTATGATCAACACTGAATACCTATGTCTAGTGTCGTGATTATGGTCATGTGAAGTGGATTTCTTAATATATGA
CTCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGTTTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGG
TGCAGGAGCTGTACTCCGTGCTGAAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGAT

ATATATTTTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGTTACTATAGTTGCACCGGCTGAGGGAAGGTG
TGTGCAACGGCA

>GPH5_mandshurica_clone1
CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATCAT
CAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAGTGAAGACAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATTGTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAG
AAAAGAGCTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGGAGCAGGTTAGCTAAAAGCTTC
AAACAAAGCGTCAATTGCCCACAGTTATTCTTTGATAGATATATGTTGAACTGTAAGAGACATATTTCAAGCTCTTT
GGTGTTCAAAGTTGGATTCAATTACATGTAGACACAGTTACCATTTTCCCATTTGAAATAGAAGGTAATACGCATGA
TATAAATATCAAGTTAATTGTACAATGATATTTGTAATCAGTGAAAATAATGACAATCTTTATAACAAAATTTCAGT
GATCTTTCCATTGCTGTATGAACTGTTACCATTAGCCTCTCACACAAGAACAACAACACCAAACAAACAGAACCAGA
CCAAATCACACCAATATAAACAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGAAGACAAAG
AATCCTTTCGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATCTCATATGCAGGCATGTTA
CATGTCTGCATTTGGTGACAAAAGCTAAATCTTAACATGACCTAAGAATTAAGACATATTGGACCATTGGGCTTAAT
CATAGTCTAAGCCCAAATCTGTACTAGCCCATAATATTCTTTTTATAGAAAACAGAGGTTCTCTGTGATCTTCACCA
TTGAGGAGTCAAGTTACTCGGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAGTTCAACTTGTTCTGGGTTC
TTCAAAGTTTGAAACTTTAAGCGTCAATGGAGGAAGAGAAGGATGCCTTTTATGTTGCTCCAACGGGAGATGTGGTT
GGCATATATACAAGCTTGAAGGATTGCCAAAACCAAGCTGGTTCATCGGTAAAGTTTTGATCTTTTAAGCCTTTTGT
AATTTGATCACTCCCATTGTTTTATCAATTTTAGATTTCCCATTTGATTACATTACTGGCTCTTGTTTATTTTGTTG
AACTAACTATGCCCTTTCGTTCTAACATGCAACTGAAAATAACTGCTAGATTGTATAGCTGAGCCTTTATGGTGTTC
ATTATGTAAAAGAGAATGAATTCTGGTGGTGGGTATAAAGCACCTCCCTGAGATTATATGAGATACTATGCTTCTGG
AAAATGTTATAAGATGAAAACAACTTTTTCTAACAACTTGATGTTATTTACTTGATGAGTTGATGGAAGATTACGTA
TGTGGTTTGGTTTTGTTTTTAGGTATTCAATCCTTCTGAAATTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCG
AGGAGTACCTTGTCTCGCATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAAAGATGGTCTGTTT
GGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATGATTTCATGTGTTCTAGTTTCTGTTTGGGTATCTGTTA
TTTTCATGGCATGTCGCGTGGAACTAGTTGCATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCGGT
TTTTTATTATAATCTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCACCTGATCAGCTTACA
CGAAAGAGAATGCCGGATGTGATAAATCCAGGTGTCAATGACCCTCCACAAAAGAGATCACTGGATGTAAGTATCAT
ATGCTACATGGAACTTTTGTAGTATGATAGAAGATCTTCTATTTGGTTACTCATTGTATGATTCGGGTTTTTATTAT
AATTTTTCAGCAGCCAGGGTCTTCTATGGTTAATTCAGGCTTCAAAATTGCACCTCATCAGCTTACGCGAATGAGAT
TGCCGGATGTGGTAAATCCAGGTGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGTATCTTATGCTACATG
GAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGATACCTTCAGTTTGGCTGGATTATGGAGTTGCGTTGATCA
CTTGTTTATTGTATTTATTCTGCAAATGATGTTTTTTGGCTTCCAGTTTTTCTTCACATAAGCATTTTAAAGCTGAT
CATTGTAATCAAACTCGAATAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATTTTGT
ATTTTATGATCAACACCGAATACCTATGTCTAGTGTCGTGATTGTGGTCATGTGAAGTGGATTTCTTAATATATGCC
TCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGTTTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGT
GCAGGAGCTGTACTCCGTGCTGAAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATTTTTTTGATAT
ATATTTTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGTTGCTATAGTTGCACCGGCTGCGGGAAGGTGTG
TGCAACGGCA

>GPH5_ananassa_clone2
CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATCAT
CGGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAGTGAAGAGAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATCCTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAG
AAAAGAGCTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGTGAGCAGGTTAGCTAAAAGCTT
CAAACAAAGCGTCAATTGCCCACAGTTATTCTTTGATAGATATATGCTGAACTGTAAGAGACATATTTCAAGCTCTT
TGGTGTTCAAAGTTGGATTCAATTACATGTAGACACAGTTACCATTTTTCCATTTGAAACAGAAGGTAATATGCATG
ATATAAATATCAAGTTAATTGTACAATGATATTATTTGTAATAAGTGAGAATAATGACAATCTTCATAACAAAATTT
CAGTTATCTTTCCATTGCTGTATGAACTGTTACCATTAGCCTCTCACACAAGAACAACTACACCAAACAAACAGAAC
CAGACCAAATCACACCAATATAAAACAGAATTGGATTTTCATGAAAGGCGGCAAGGCACAATCAATGAAGGAGAAGA
CAAAGAATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATCTCATATGCAGGCA
TGTTACATGTCTGCATTTGGTGACAAAAGCTAAATCTTAACATGACCTAAGAATTAAGACATATTGGACCATTGGGC
TTAATCATAGTCTAAGCCCAAATCTGTACTAGCCCATAATATGCTTTTTATAGAAAATACTGTGATCTTCACCATTG

AGGAGTCAAGTTACTCAGCCATGAAGTCAAGGTCAAGCCAAGTAGTGCAGTTGAGTTCAACTTGTTCTGGGTTCTTC
AAAGTTCGAAACTTTAAGCTTCAATGGAGGAAGAGAAGGATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGC
ATGTATAAAAGCTTGAAGGATTGCCAAAACCAAGCTGGTTCATCGGTAAAGTTTTGATCTTTTAAGCCTTTTGTAAT
TTGATCACTCCCATTGTTTTATCAATTTTTGATTTCCCATTTGATTACATTACTGGGTCTTGTTTATTTTGTTGAAA
TAACTATGCCCTTTCGTTCTAGCATGCAACTGAAATTTACTGCTAGATTGTATTGTTGTGCCGTTATGGTGTTCATT
ATGTAAAAGAGAATGAATTCTGGTGGTGGGTATAGAGTACCTCCCTGATTTTTTATGAGATACTATGCTTCTGGAAA
ATGTTATAAAGATGAAAACTAGTTTTTCTAAGAACTTGATGTTATTTACTTGATGAGTTGATGGAGGATTACGTATG
TGGTTTGGTTTTGTTTTTAGGTATTCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAG
GAGTACCTTGTCTCACATGGGCTTAAGAATGCTTCATGTACTATCAGTGCCAGTGATGTGAAAGATGGTCTGTTTGG
AAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATTTCATGTGTTCTAGTTTCTGTTTGGGCATCTGTTATTTTC
ATGGCATGTGGCGTGGAGCTAGTTGCATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCAGTTTTTT
ACTATAATCTTTCAGCAGCCAGCATCATCTATGGTTAATTCAGGGTTCAAAATTGCGCCTAATCAGCTTACACCAAA
GAGAATGCCGGATGTGGTAAATTCAGGCGTCAATGACCCTCCACAAAAGAGATCGCTGGATGTAAGTATCATATGCT
ACATGGAACTTTTGTAGTTTGATAGAAGACCTTCTATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCT
TTCAGCAGCCAGCGTCTTCTATGGTTAATTCAGGCTTCAAAGTTGCACCTGATCAGTTTACACGAATGCGATTGCCG
GATGTGGTAAATTCAGGTGTCAATTACCCTCCACAGAGGACATTGCCGGATGTAAGTATCTTATGCTACATGGAAAT
TTTGTTCTGGCCAGATTGGCATGAAGATCCAGACACCTTCAGTCTGGCTGGATTATGGAGTTGCGTTGATCACTTGT
TTATTGTATTTATTCTGCAAATGATGTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTG
TAATCGAACTCAAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATTTCGTATTTT
ATGATCAACACTGAATACCTATGTCTAGTGTCATGATTATAGTCATGTGAAGTGGATTTCTTAATATATGCCTCATC
TATGTCTTCATCCAGCAATCCTGCATACTTGAGTTTGATGGTGCTTCAAAAGGAAATCCTGGACCATCTGGTGCAGG
AGCTGTACTCCGTGCTGAAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATAT
TTTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGCTACTATAGTTGCACCGGCTGCGGGAAGGTGTGTGCA
ACGGCA

>GPH5_ananassa_clone6
CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATCAT
CAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAGTGGAGAGAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATTGTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAAAAGAAAAATGGGATCTACAG
AAAAGAGCTTGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn

nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnACTGGTGTAAGTTGCCTTGTGTCACCTCCACTAAGATCACAATTTCGTATTTTATGATCAACACT
GAATACCTATGTCTAGTGTCATGATTATAGTCATGTGAAGTGGATTTCTTAATATATGCCTCATCTATGTCTTCATT
CAGCAATCCTGCATACTTGAGTTTGATGGTGCTTCAAAAGGAAATCCTGGACCATCTGGTGCAGGAGCTGTACTCCG
TGCTGAAGATGGGAGTGTTGTATGTGGATTTCATGAAAACATTGTGAATCTTTTAGGATATATATTTTTGTTTTTGT
AAAAATGGATCTCTTTATAACGTTGGGGTTACTATAGTTGCACCGGCTGCGGGAAGGTGTGTGCAACGGCA

>GPH5_ananassa_clone7
CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAAGAAGACCGGAAGATCAT
CGGAATCAGATTTGTT**C**GCGCTTGCAGAAAGAGAATCCTCGAGTGAAGAGAAGATCCTAAGGAGGAACTCCGAGTCT
GGTTTAGAATCCTTGAGCAAACTCAAGGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAG
AAAAGAGCTTGCTCTTGCTTTCC**A**CCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGTTCACTTCT
ATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAATAGGTGAGCAGGTTAGCTAAAAGCTT
CAAACAAAGCGTCAATTGCCCACAGTTATTCTTTGATAGATATATGTTGAACTGTAAGAGACATATTTCAAGCTCTT
TGGTGTTCAAAGTTGGATTCAATTACATGTAGACACAGTTACCATTTTTCCATTTGAAACAGAAGGTAATATGCATG
ATATAAATATCAAGTTAATTGTACAATGATATTATTTGTAATAAGTGAAAATAATGACAATCTTTATAACAAAATTT
CAGTTATCTTTCCATTGCTGTATGAACTGTTACCATTAGCCTCACACACAAGAGCAACAACACCAAACAAACAGAAC
CAGACCAAATCACACCAATATAAAACAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGAAGA
CAAAGAATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATCTCATATGCAGGCA
TGTTACATGTCTGCATTTGGTGACAAAAGCTAAATCTTAACATGACCTAAGAATTAAGACATATTGGACCATTGGGC
TTAATCATAGTCTAAGCCCAAATCTGTACTAGCCCATAATATGCTTTTTGTAGAAAATACTGTGATCTTCACCATTG
AGGAGTCAAGTTACTCAGCCATGAAGTCAAGGTCAAGCCAAGTAGTGCAGTTGAGTTCAACTTGTTCTGGGTTCTTC
AAAGTTCGAAGCTTTAAGCTTCAATGGAGGAAGAGAAGGATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGC
ATATATAAAAGCTTGAAGGATTGCCAAAACCAAGCTGGTTCATCGGTAAAGTTTTGATCTTTTAAGCCTTTTGTAAT
TTGATCACTCCCATTGTTTTATCAATTTTTGATTTCCCATTTGATTACATTACTGGGTCTTGTTTATTTTGTTGAAA
TAACTATGCCCTTTCGTTCTAGCATGCAACTGAAATTTACTGCTAGATTGTATTGTTGTGCCGTTATGGTGTTCATT
ATGTAAAAGAGAATGAATTCTGGTGGTGGGTATAGAGTACCTCCCTGATTTTTTATGAGATACTATGCTTCTGGAAA
ATGTTATAAGATGAAAACTAGTTTTTCTAAGAACTTGATGTTATTTACTTGATGAGTTGATGGAGGATTACGTATGT
GGTTTGGTTTTGTTTTTAGGTATTCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGG
AGTACCTTGTCTCACATGGGCTTAAGAATGCTTCATGTACTATCAGTGCCAGTGATGTGAAAGATGGTCTGTTTGGA
AGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATTTCATGTGTTCTAGTTTCTGTTTGGGCATCTGTTATTTTCA
TGGCATGTGGCGTGAAGCTAGTTGCATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCAGTTTTTTA
TTATAATCTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCGCCTAATCAGCTTACACCAAAG
AGAATGCCGGATGTGGTAAATTCAGGCGTCAATGACCCTCCACAAAAGAGATCGCTGGATGTAAGTATCATATGCTA
CATGGAACTTTTGTAGTTTGATAGAAGATCTTCTATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCTT
TCAGCAGCCAGCGTCTTCTATGGTTAATTCAGGCTTCAAAATTGCACCTGATCAGTTTACACGAATGCGATTGCCGG
ATGTGGTAAATTCAGGTGTCAATTACCCTCCACAGAGGACATTGCCGGATGTAAGTATCTTATGCTACATGGAAATT
TTGTTCTGGCCAGATTGGCATGAAAATCCAGACACCTTCAGTTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTT
TATTGTATTTATTCTGCAAATGATGTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGT
AATCGAACTCAAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACACTAAGATCACAATTTCGTATTTTAT
GATCAACACTGAATACCTATGTCTAGTGTCATGATTATAGTCATGTGAAGTGGATTTCTTAATATATGCCTCATCTA
TGTCTTCATCCAGCAATCCTGCATACTTGAGTTTGATGGTGCTTCAAAAGGAAATCCTGGACCATCTGGTGCAGGAG
CTGTACTCCGTGCTGAAGATGGAAGTGTTGTATGCGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATATTT
TTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGTTACTATAGTTGCACCGGCTGCGGGAAGGTGTGTGCAAC
GGCA

>GPH10_ananassa_clone2
GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGCCGATACCTCCTCCGCCTCCGACGACTT
CGAACACAGCGGAATCGCTAGCCTCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGTGAAT
AGCGAACTGAGTAGTTTGGATTTGAGAAGAGGATGTAATTGGTAACGGAGAAGAAGACTGTCGACATTTTTGGAGAA
AGCTTTCATCTTTGAAGTGGAGTGTAGGATAATAACAAACTCGTTATCTAAAAGGCAGGTTTAATATCAGCCGTTAG
ATCATATTACGGCCCTGATCACTCGACATATGTTGATATACGCCCAACTCAAATTCGATATATATTTTCGATATACA
TATATTTTATTTTTTTAAAGTAACTAAATGACTATGTACATCGTTTAACAAAAGAAACAATTGAAGTTAAATTAAGA
GCACCATAACAGCTGAGAAAGAGTACGAGAACAAAAGTATGAGCTAAAACAAATAGAGAAAATATAGAGGCGATGTT

GTAGAAATAATTGAACATTAGAAAATTAAATTACCTAAAAGCCGATGAGTAAAATAATAACGAACTCGTAACCTAAA
AGCGGCTTCATATCATCCGCTTGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATT
CCCAATTTTCATTTCCCACCAAAAACAAAACCCACACGACGCCGTTTTGCTCCAATCCCCCTTTCTTCTTCAACCCC
ATAGTCGCCTCAGCTCAGTTCCATTTGTCTCAGATGCGATGGCCTCCGGCGACCCAATCTCCGACTACACCCAAACA
CATCGCATTGTCCTTCTAATCGACCTCAACCCACTCCTCCATCTCCAAGATCCAACCCAATTCCTCACCTCTGTCCT
CTCCTCAATCAAAACCCTAACCTCCTTCCCTTCTCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCTCTCCTCT
CCTCTCCGCCTCCAAGCTCCCGTCTTCGTCTCTAACGATCTCTTTCAACTCGCCGGAAGACACATATCGATCCCTAT
CTCAAACCCTGGCGTCTCTCTCGTTTGACCGGAAGTTGACCGGGTCCGATTCGCCGCGGGGAACGCTTGTTGCGGCT
GCGATGCGGCAGCTGGTGCATGATTACGCTTGGGAGCAGGTGATCTGCGACGCCGTGGCGGCGGAGACAGATACGTT
TTCGAATTGCTGTGGTTTGAGGTCTAATTTGGCTGTTGTGTTTTCACCGGCGTGTCAATTTGTGAATGAGTTCTTGA
ATTGTGAGGGTTTGGAGGATTTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGAGGCATAT
GTGTATAGAGATATTCAATTGAGTCGGGTTGATGTGAGGTATGGATTCGATAGCGGTGAGGATGAGGTAGTTGGATT
GAAATGTGGTGTTTTCGAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGATTGTGCTTGGTT
CGGCTCTTGTTCCATTTGGTTTGATTTATCCAGAGATTGGGGTGTCATCTAGGATTTTCGGGTGTAATGATCGATAT
AAGAAGGTTAGAGCGCATTTGAGTCTTGAGATATCGGATGTAAAGGGGATGCCTTTGGAGTGCAAGTTTTGTGATCT
TGAGTTGGCTGATTTGAAAATGTTGTGTAGGAGTAGAGGTGATGATCGCTTGTTTTCGGTGGAAGGCATGAGCTCGC
AGACAAGAGGTCATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAATGGAGTGTCGAAGATTCAGGTTAAGGCT
TTGCAGAAGGATAGTGAGTTTGGGAAATTTAAGGGGGAATTGTCGGATCTGATTCTGGTCTATGAAGTTTCAGGAAA
AGATGGAAAAGAAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAATGCTATCAGTGGAATTGGGTGAGTTTG
TACCGAGGAAATTGCCACCTGTTTGGCAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTAGTTTCT
ATTTCAAATGATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTACAGTTTCTTCAGCTCTTATTTTTGTTAT
GGATGAAGGAATTCACCCTCATAAAAAAGGGCATGGCATTGGTGCAGTGAATAAGGGTCAGTCTCGTCCAAAGATGA
AGAATGAGATGTGCAAACCTGATGCTGATTTGAACGACTTTTGTGGGTCGCAAACTGGGCCTTCACCATCTAATAAG
CATTCTGCTGAGATTGATGGAAAGAAAAAAAGTAGCAAAAGAAGTTCACATTCACTCAAAGATCTCACCTGGAGTTC
TTTCTGTAAGGCAGCATTCGAATTTTCAGACTTACATTTGGAAGAGGTTTACTTTGCCAGGCAACGTAGCAGCTCAA
AAAAGTTGAAATTTCTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGAGGAATCTAAGGTG
CACCAGGAAAAACAAAAGGAGATGAGCAATAGGTTGGATTTGTTGCACCAAGAGAGCGAACAGCCAATGTCGTCATC
TGGTTCAGCTGGAGAAATTTCTTTCCCTGTGGCCTTTGGAGTACAGGATGAAGCTGCTCAGGAACATAGATTACAAA
CCTCAGAAGATTTTTTCTGTAATTTCTCTGATAAGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTGGGGGCA
TTCGCACATCGGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAACATAGCACAACAACCCCTTCAGAAGATCA
AACTCCTGTAAAATATGACAATCTTGATGATTTGGTTACTGCTGAGCTGTTAAAACTTTTACTCAGAGATCCCAAGG
ATATGGTTGCCAGGCACAAAAGCTATGATTCATCTTCTCAAGCATCTGATCCTGGATGTGAAGGCTTTACTTCAGAA
ATAATAGTTCGAGAGTATCCTTTCATTTCTCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAATTCTAC
TTTAATGCTATGTAAACTTTGCCCCTTGTTAGTGTTACACTTTTCCTTCACTAGCACAAAGATATGAATTACAGATA
CTTTTCCGGATGGAGATTTTACAATCAGAAGTTGGAGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGTGAAACA
TATTTGCACGCTTTTGGAGACCATTCGTGCTCGGTGTCATCTGGAGGGAGGCTTCTTTGGTGACTGGACCCTAGAAA
ATTATGCTGGAAAGATTATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAACTATTATTTTCT
AATGAAATTGTATTCATGAACACTGAAATGGTAGATACTCAGTTATTTACAATGAAACTCCAATATATGTTTATGGT
TTGCCTGTTAATGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTATGTGATTGATTTGTATG
CCCTCCCAAAAGGCCTTTGGGGGTAGTATGAAGAAGGGAGACATTGACAGTCAAAAATATTATCTCCTTATTTTACG
TACAAAATTGATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAAAAGCTAAATACCTCCTTTGC
ATAATTTCATATGACTTAAGTGACTTTCCTTATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGTATGCAT
ACAACTTTTGCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTCCAGACCCAAAAAAAACTGCCCAAA
TTTATGTGAAAACACTGCATTTATGTTTGAAGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCAAATTTT
CAGGTACTGTCAGACTCTTGAAGACGTGGTTCATAAAATCTACACAAAAATGGATTTGTTACTGTTTGATGATGAGG
AAGAACTCCCTAATAATTTATTCAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGATGAGGTG
GGTGAAAATAGTAGAATGAAGAAATTGGTATCAGCAGAAGATGAATCCCCTGATCCACAGAAACATTACAATGGAAG
GCCAAGTGCTCAAGTACTTAAACAAGAAGAGCATGCTCGCAAGTTGATGAAAGCTCAAGAGAGTAGAGAGAGGGCTT
GGAGAATTGCTTCTTTCACAAGTCGGGTAGCTGATTTGCAGCGAG

>GPH10_ananassa_clone7_(same_restriction_pattern_as_clone20)
GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGCCGATACCTCCTCCGCCTCCGACGACTT
CGAACACAGCGGAATCGCTAGCCTCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGTGAAT
AGCGAACTGAGTAGTTTGGATTTGAGAAGAGGATGTAATTGGTAACGGAGAAGAAGACTGTCGACATTTTTGGAGAA
AGCTTTCGGCTTTGAAGTGGAGTGTAGGGTAATAACAAACTCGTGATTAAAAGACAGGATTAATGTCAGTGAGGTTT
GGTTGGTTAAGGTGTTAACTGATAAATTTAAGGTCATAGGTTCAAACCTCACGACATATGTAGGGTGTATGAATTAT

```
TAATAAAAGACAAATTTAATATCAGCCGTTAGATCATATTACGGCCTGATCACTCGACATATGTTGATATACGCCCA
ACTCAAATTCGATATATATTTTCGATATACGTATATTTTATTTTTTTAAAATAATTAAATAACTATTTACGTTGTTT
AACAAAAGAAACAATTGAAGTTAAATTAAGAGCACCGTAACAGCTGAGCAAGAGTACGAGAACAAAAGTATGAGCTA
CATCATTTGTTCATATAGAGAAAATATAGAGGCGATGTTGTAGAAATAATTGAACATTAGAAAATTAAATTACCTAA
AAGCCGATGAGTAAAATAATAACGAACTCGTAACCTAAAAGCGGCTTCATATCATCCGCTTGATCATATATGCGGGT
GTGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAATAAAACCCACACG
ACGCCGTTTTGCTCCAATCCCCCTTTCTTCTTCAACCCCATAGTCGCCTCAGCTCAGTTCCATTTGTCTCAGATGCG
ATGGCCTCCGGCGACCCAATCTCCGACTACACCCAAACACATCGCATTGTCCTTCTAATCGACCTCAACCCACTCCT
CCATCTCCAAGATCCAACCCAATTCCTCACCTCTGTCCTCTCCTCAATCAAAACCCTAACCTCCTTCCCTTCTCTCT
CTTCCTCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCTCTCCTCTCCTCTCCTCTCCGCCTCCAAGCTCCCGT
CTTCGTCTCTAACGATCTCTTTCAACTCGCCGGAAGACACATATCGATCCCTATCTCAAACCCTGGCGTCTCTCTCG
TTTGACCGGAAGTTGACCGGGTCCGATTCGCCGCGGGGAACGCTTGTTGCGGCTGCGATGCGGCAGCTGGTACATGA
TTACGCTTGGGAGCAGGTGATCTGCGACGCCGTGGCGGCGGAGACAGGTACGTTTTCGAATTGCTGTGGTTTGAGGT
CTAATTTGGCTGTTGTGTTTTTACCGGCGTGTCAATTTGTGAATGAGTTCCTGAATTGTGAGTTGAATTGTGAGGGT
TTGGAGGATTTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGAGGCATATGTGTATAGAGA
TATTCAATTGAGTTGGGTTGATGTGAGGTATGGATTCGATAGCGGTGAGGATGAGGTAGTTGGATTGAAATGTGGTG
TTTTCGAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGATTGTGCTTGGTTCGGCTCTTGTT
CCATTTGGTTTGATTTATCCAGAGATTGGGGTGTCATCTAGGATTTTCGGGTGTAATGATCGATATAAGAAGGTTAG
AGCGCATTTGAGTCTTGAGATATCGGATGTAAAGGGGATGCCTTTGGAGTGCAAGTTTTGTGATCTTGAGTTGGCTG
ATTTGAAAATGTTGTGTAGGAGTAGAGGTGATGATCGCTTGTTTTCGGTGGAAGGCATGAACTCGCAGACAAGAGGT
CATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAATGGAGTGTCGAAGATTCAGGTTAAGGCTTTGCAGAAGGA
TAGTGAGTTTGGGAAATTTAAGGGGGAATTGTCGGATCTGATTCTGGTCTATGAAGTTTCAGGAAAAGATGGAAAAG
AAGTTTCTGGTGGTTTGTTTGTAGATAGGGTTCTTGAAATGCTATCAGTGGAATTGGGTGAGTTTGTACCGAGGAAA
TTGCCACCTGTTTGGCAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTAGTTTCTATTTCAAATGA
TAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTACAGTTTCTTCAGCTCTTATTTTTGTTATGGATGAAGGAA
TTCACCCTCATAAAAAAGGGCATGGCATTGGTGCAGTGAATAAGGGTCAGTCTCGTCCAAAGATGAAGAATGAGATG
TGCAAACCTGATGCTGATTTGAACGACTTTTGTGGGTCGCAAACTGGGCCTTCACCATCTAATAAGCATTCTGCTGA
GATTGATGGAAAGAAAAAAAGTAGCAAAAGAAGTTCACATTCACTCAAAGATCTCACCTGGAGTTCTTTCTGTAAGG
CAGCATTCGAATTTTCAGACTTACATTTGGAAGAGGTTTACTTTGCCAGGCAACGTAGCAGCTCAAAAAAGTTGAAA
TTTCTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGAGGAGTCTAAGGTGCACCAGGAAAA
ACAAAAGGAGATGAGCAATAGGTTGGATTTGTTGCACCAAGAGAGCGAACAGCCAATGTCGTCATCTGGTTCAGCTG
GAGAAATTTCTTTCCCTGTCGCCTTTGGAGTACAGGATGAAGCTGCTCAGGAACATAGATTACAAACCTCAGAAGAT
TTTTTCTGTAATTTCTCTGATAAGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTGGGGGCATTCACACATCG
GCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAACATAGCACAACAACCCCTTCAGAAGATCAAACTCCTGTAA
AATCTGACAATCTTGATGATTTGGTTACTGCTGAGCTGTTAAAACTTTTACTCAGAGATCCCAAGGATATGGTTGCC
AGGCACAAAAGCTATGATTCATCTTCTCAAGCATCTGATCCTGGATGTGAAGGCTTTACTTCAGAAATAATAGTTCG
AGAGTATCCTTTCATTTATCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAATTCTACTTTAATGCTAT
GTAAACTTTGCCCCTTGTTACTGTTACACTTTTCCTTCACTAGCACAAAGATATGAATTACAGATACTTTTCCGGAT
GGAGATTTTACAATCAGAAGTTGGAGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGTGAAACATATTTGCACGC
TTTTGGAGACCATTCGTGCTCGGTGTCATCTGGAGGGAGGCTTCTTTGGTGACTGGACCCTAGAAAATTATGCTGGA
AAGATTATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAACTATTATTTTCTAATGAAATTTG
TATTCATGAACACTGAAATGGTAGATACTCAGTTATTTACAATGAAACTCCAATATATGTTTATGGTTTGCCTGTTA
ATGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTATGTGATTGATTTGTATGCCCTCCCAAA
AGGCCTTTGGGGGTAGTATGAAGAAGGGAGACATTGACCGTCAAAAATATTATCTCCTTATTTTACGTACAAAATTG
ATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAAAAGCTAAATACCTCCTTTGCATAATTTCAT
ATGACTTAAGTGACTTTCCTTATTAATCTAGATTTGCAACCTTGTTTCTCTGACACTATGTATGCATACAACTTTTG
CAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTCCAGACCCAAAAAAAACTGCCCAAATTTATGTGAA
CACACTGCATTTATGTTTGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCAAATTTTCAGGTACTGTCAG
ACTCTTGAAGACGTGGTTCATAAAATCTACACAAAAATGGATTTGTTACTGTTTGATGATGAGGAAGAACTCCCTAA
TAATTTATTCAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGATGAGGTGGGTGAAAATAGTA
GAATGAAGAAATTGGTATCAGCAGAAGATGAATCCCCTGATCCACAGAAACATTACAATGGAAGGCCAAGTGCTCAA
GTAGTTAAACAAGAAGAGCATGCTCGCAAGTTGATGAAAGCTCAAGAGAGTAGAGAGAGGGCTAGGAGAATTGCTTC
TTTCACAAGTCGGGTAGCTGATTTGCAGCGAG
```

>GPH10_ananassa_clone18

```
GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGCCGATACCTCCTCCGCCTCCGACGACTT
CGAACACAGCGGAATCGCTAGCCTCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGTGAAT
AGCGAACTGAGTGGTTTGGATTTGAGAAGAGGATGAAATTGGTAACGGAGAAGAAGACTGTCGACATTTTTAGAGAA
AGCTTTCAGCTTTGAAGTGGAGTGTAGGATAATAACAAACTCGTTATCTAAAAGACAGGTTTAATATCGGCCGTTAG
ATCACATTACGGCCCTGATCACTCGACATATGTTGATATACGCCTAACTCAAATTCGATATATATTTTCGATATACA
TTTTTTTTTTAAGTAACTAAATGACTATTCGATATATATTTTCGATATACATTTTTTTTTTAAAGTAACTAAATGAC
TATTTACGTCGGTTAATAAAAGAAACAATTGAAGTTAAATTAAGAGCACCATGACAGAGTACGAGAACAAAAGTATG
AGCTACATTGTTTGCTCGTCGGTTTGTTCATATGGAGAAAATGTAGAGGCGATGTTGTAGAAATAATTGAACATTAG
AAAATTAAATTACCTAAAAGCCGATGAGTAAAATAATAACAAACTCGTAACCTAAAAGCGGCTTCATATCATCCACT
GGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCA
AAAACAAAACCCACACGACGCCGTTTTGCTCCAATCCCCCCCCCCTTTCTTCTTCAACCCCATAGTCGCCTCCTCAG
CTCAGTTCCATTTGTCTCATGCGATGGCTTCCGACTCGAATTCCGGCGACCCAATCTCCTCCTACACCCAAACCCAT
CGCATCGTCCTTCTAATCGACCTCAACCCACTCCTCAATCTCCAAGATCCAACCCAATTCCTCACCCCTGTCCTCTC
CTCAATCAAAACCCTAACCTCCTTCCCTTCTCTCTCTTCATCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCT
CTCCTCTCCTCTCCGCCTCCAAGCTCCCGTCTTCGTCTCTAACGATCTCTTTCAACTCGCCGGAAGACACTTATCGA
TCCCTATCTCAAACCCTGGCGTCTCTCTCTTTTGACCGCAAGTTGGCCGGGTCCGATTCGCCGCGGGGAACGCNTGT
TGCGGCGGCGATGCGGCAGCTGGTGCATGATTACGCTTGGGAGCCGGTGATCTGCGACGCCGCGGCGGCGGAGACCG
GTACGTTATCGAATTGCTGTGGTTTGAGGTCTAATTTGGCTGTTGTGTTTTCACCGGCGTGTCAATTTGTGAATGAG
TTCTTGAATTGTGAGGGTTTGGAGGATTTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGA
GGCATTTGTGTGTAGAGATATTCAATTGAGTTGGGTTGATGTGAGGTATGGATTCGATAGCGGTGAGGATGAGGTAG
TTGGATTGAAATGTGGTGTTTTCGAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGATTGTG
CTTGGTTCGGCTCTTGTTCCATTTGGTTTGATTTATCCAGAGATTGGGGTGTCATCTAGGATTTTCGGGTGTAATGA
TCGATATAAGAAGTTTAGAGCGCATTTGAGTCTTGAGATATCGGATGGAAAGGGGATGCCTTTGGAGTGCAAGTTTT
GTGATCTTGAGTTGGCTGATTTGAAAATGTTGTGTAGGAGTAGAGGTGATGATGGCTTGTTTTCGGTGGAAGGCATG
AACTCGCAGACAAGAGGTCATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAACGGAGTGTTGAAGATTCAGGT
TAAGGCTTTGCAGAAGGATAGTGAGTTTGGGAAATTTAAGGGGGAATTGTCGGATCCGATTCTGGTCTATGAAGTTT
CAGGAAAAGATGGAAAAGAAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAATGCTATCAAGTGGAATTGGG
TGAGTTTGTACCAAGGAAATTGCCACCTGTTTGGCAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCAT
TAGTGTCTATTTCAAATGATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTACAGTTTCTTCAGCTCTTATT
TTTGTTATGGATGAAGGAATTCACCCTCATAAAAAAGGGCATGTCATTGGTGCAGTGAATAAGGGTCAGTCTCGTCC
AAAGATGAAGAATGAGATGTGCAAACCTGATGCTGATTTGAACGACTTTTGTGGGTCGCAAACTGGGCCTTCACCAT
CTAATAAGCATTCTGCTGAGATTGATGGAAAGAAAAAAAGTAGCAAAAGAAGTTCACATTCACTCAAAGATCTCACC
TGGAGTTCTTTCTGTAAGGCAGCATTCGAATTTTCAGACTTACATTTGGAAGAGGCTTACTTTGCCAGGCAACGTAG
CAGCTCAAAAAAGTTGAAATTTCTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGAGGAGT
CTAAGGTGCACCAGGAAAAACAAAAGGAGATGAGCAATAGGTTGGATTTGTTGCACCAAGAGAGCGAACAGCCAATG
TCATCATCTGGTTCAGCTGGAGAAATTTCTTTCTCTGCGGCCTTTGGAGTACAGGATGAAGCTGCTCAGGAACATAG
ATTACAAACCTCAGAAGATTTTTTCTGTAATTTCTCTGATAAGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACT
TGGGGGCATTCGCACATCGGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAGCATAGCTCAACAACCCCTTCA
GAAGATCAAACTCCTGTAAAATCTGACAATCTTGATGATTTGGTTACTGCTGAGCTGTTAAAACTTTACTCAGAGAT
CCCAAGGATATGGTTGCCAGGCACAAAAGCTATGATCCATCTTCTCAAGCATCTGATCCTGGATGTGATGGCTTTAC
TTCAGAAATAATAGTTCGAGAGTATCCTTTCATTTATCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCA
ATTCTACTTTAATGCTATGTAAACTTTGCCCCCTGTTACTGTTACACTTCCTTCACTAGCACAAAGATATGAATTAC
AGATACTTTTCCGGATGGAGATTTTACAATCAGAAGTTGGAGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGTG
AAACATATTTGCACGCTTTTGGAGACCATTCGTGCTCAGTGTCATCTGGAGGGAGGCTTCTTTGGTGACTGGACCCT
AGAAAATTATGCTGGAAAGATTATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAACTATTAT
TTTCTAATGAAATTTGTATTCATGAACACTGAAATGGTAGATACTCAGTTATTTACAATGAAACTCCAGTATATGTT
TATGTTTTGCCTGTTAATGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTATGTGATTGATT
TGTATGCCCTCCCAAATGGCCTTTGGGGGTAGTAAGAAGAAGGGAGACATTGACAGTCAAAAATATTATCTCCTTAT
TTTACGTACAAAATTGATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAAAAGCTAAATACCTC
CTTTGCATAATTTCATATGACTTAAGTGACTTTCCTTATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGT
ATTCATACAACTTTTGCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTCCAGACCCAAAAAAAACGG
CCCAAATTTATGTGAAAACACTGCATTTATGTTTGAAGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCA
AATTTTCAGGTACTGTCAGACTCTTGAAGACGTGGTTCATAAAATCTACACAAAAATGGATTTGTTACTGTTTGATG
ATGAGGAAGAACTCCCTAATAATGTATTCAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGAT
GAGGTGGGTGAAAATAGTAGAATGAAGAAATTGGTATCAGCAGAAGATGAATCCCCTGATCCACAGAAACATTACAA
TGGAAGGCCAAGTGCTCAAGTAGTTAAACAAGAAGAGCATGCTCGCAAGTTGATGGAAGCTCAAGAGAGTAGAGAGA
GAGCTAGGAGAATTGCTTCTTTTACAAGTCGGGTAGCTGATTTGCAGCGAG
```

>GPH10_ananassa_clone19

```
GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGCCGATACCTCCTCCGCCTCCGACGACTT
CGAACACAGCGGAATCGCTAGCCTCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGTGAAT
AGCGAACTGAGTGGTTTGGATTTGAGAAGAGGATGAAATTGGTAACGGAGAAGAAGACTGTCGACATTTTTAGAGAA
AGCTTTTAGCTTTGAAGTGGAGTGTAGGATAATAACAAACTCGTTATCTAAAAGACAGGTTTAATATCAGCCGTTAG
ATCATATTACGGCCCTGATCACTCGACATATGTTGATATACGCCTAACTCAAATTCGATATATATTTTCGATATACA
TTTTTTTTTTAAGTAACTAAATGACTATTCGATATATATTTTCGATATACATTTTTTTTTTAAAGTAACTAAATGAC
TATTTACGTCGGTTAATAAAAGAAACAATTGAAGTTAAATTAAGAGCACCATGACAGAGTACGAGAACAAAAGTATG
AGCTACATTGTTTGCTCGTCGGTTTGTTCATATGGAGAAAATATAGAGGCGATGTTGTAGAAATAATAGAACATTAG
AAAATTAAATTACCTAAAAGCCGATGAGTAAAATAATAACAAACTCGTAACCTAAAAGCGGCTTCATATCATCCACT
GGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCA
AAAACAAAACCCACACGACGCCGTTTTGCTCCAATCCCCCCCCCTTTCTTCTTCAACCCCATAGTCGCCTCCTCAGC
TCAGTTCCATTTGTCTCATGCGATGGCTTCCGACTCAAATTCCGGCGACCCAATCTCCTCCTACACCCAAACCCATC
GCATCGTCCTTCTAATCGACCTCAACCCACTCCTCAATCTCCAAGATCCAACCCAATTCCTCACCCCTGTCCTCTCC
TCAATCAAAACCCTAACCTCCTTCCCTTCTCTCTCTTCATCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCTC
TCCTCTCCTCTCCGCCTCCAAGCTCCCGTCTTCGTCTCTAACGATCTCTTTCAACTCGCCGGAAGACACTTATCGAT
CCCTATCTCAAACCCTGGCGTCTCTCTCTTTTGACCGCAAGTTGGCCGGGTCCGATTCGCCGCGGGGAACGCTTGTT
GCGGCGGCGATGCGGCAGCTGGTGCATGATTACGCTTGGGAGCCGGTGATCTGCGACGCCGCGGCGGCGGAGACCGG
TACGTTATCGAATTGCTGTGGTTTGAGGTCTAATTTGGCTGTTGTGTTTTCACCGGCGTGTCAATTTGTGAATGAGT
TCTTGAATTGTGAGGGTTTGGAGGATTTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGAG
GCATTTGTGTGTAGAGATATTCAATTGAGTTGGGTTGATGTGAGGTATGGATTCGATAGCGGTGAGGATGAGGTAGT
TGGATTGAAATGTGGTGTTTTCGAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGATTGTGC
TTGGTTCGGCTCTTGTTCCATTTGGTTTGATTTATCCAGAGATTGGGGTGTCATCTAGGATTTTCGGGTGTAATGAT
CGATATAAGAAGTTTAGAGCGCATTTGAGTCTTGAGATATCGGATGGAAAGGGGATGCCTTTGGAGTGCAAGTTTTG
TGATCTTGAGTTGGCTGATTTGAAAATGTTGTGTAGGAGTAGAGGTGATGATGGCTTGTTTTCGGTGGAAGGCATGA
ACTCGCAGACAAGAGGTCATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAACGGAGTGTTGAAGATTCAGGTT
AAGGCTTTGCAGAAGGATAGTGAGTTTGGGAAATTTAAGGGGGAATTGTCGGATCCGATTCTGGTCTATGAAGTTTC
AGGAAAAGATGGAAAAGAAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAATGCTATCAGTGGAATTGGGTG
AGTTTGTACCAAGGAAATTGCCACCTGTTTGGCAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTA
GTGTCTATTTCAAATGATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTACAGTTTCTTCAGCTCTTATTTT
TGTTATGGATGAAGGAATTCACCCTCATAAAAAAGGGCATGTCATTGGTGCAGTGAATAAGGGTCAGTCTCGTCCAA
AGATGAAGAATGAGATGTGCAAACCTGATGCTGATTTGAACGACTTTTGTGGGTCGCAAACTGGGCCTTCACCATCT
AATAAGCATTCTGCTGAGATTGATGGAAAGAAAAAAAGTAGCAAAAGAAGTTCACATTCACTCAAAGATCTCACCTG
GAGTTCTTTCTGTAAGGCAGCATTCGAATTTTCAGACTTACATTTGGAAGAGGTTTACTTTGCCAGGCAACGTAGCA
GCTCAAAAAAGTTGAAATTTCTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGAGGAGTCT
AAGGTGCACCAGGAAAAACAAAAGGGGATGAGCAATAGGTTGGATTTGTTGCACCAAGAGAGCGAACAGCCAATGTC
ATCATCTGGTTCAGCTGGAGAAATTTCTTTCTCTGCGGCCTTTGGAGTACAGGATGAAGCTGCTCAGGAACATAGAT
TACAAACCTCAGAAGATTTTTTCTGTAATTTCTCTGATAAGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTG
GGGGCATTCGCACATCGGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAGCATAGCTCAACAGCCCCTTCAGA
AGATCAAACTCCTGTAAAATCTGACAATCTTGATGATTTGGTTACTGCTGAGCTGTTAAAACTTTTACTCAGAGATC
CCAAGGATATGGTTGCCAGGCACAAAAGCTATGATCCATCTTCTCAAGCATCTGATCCTGGATGTGATGGCTTTACT
TCAGAAATAATAGTTCGAGAGTATCCTTTCATTTATCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAA
TTCTACTTTAATGCTATGTAAACTTTGCCCCCTGTTACTGTTACACTTCCTTCACTAGCACAAAGATATGAATTACA
GATACTTTTCCGGATGGAGATTTTACAATCAGAAGTTGGGGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGTGA
AACATATTTGCACGCTTTTGGAGACCATTCGGTGCTCAGTGTCATCTGGAGGGAGGCTTCTTTGGTGACTGGACCCT
AGAAAATTATGCTGGAAAGATTATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAACTATTAT
TTTCTAATGAAATTTGTATTCATGAACACTGAAATGGTAGATACTCAGTTATTTACAATGAAACTCCAATATATGTT
TATGTTTTGCCTGTTAATGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTATGTGATTGATT
TGTATGCCCTCCCAAATGGCCTTTGGGGGTAGTAAGAAGAAGGGAGACATTGACAGTCAAAAATATTATCTCCTTAT
TTTACGTACAAAATTGATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAAAAGCTAAATACCTC
CTTTGCATAATTTCATATGACTTAAGTGACTTTCCTTATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGT
ATTCATACAACTTTTGCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTCCAGACCCAAAAAAAACGG
CCCAAATTTATGTGAAAACACTGCATTTATGTTTGAAGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCA
AATTTTCAGGTACTGTCAGACTCTTGAAGACGTGGTTCATAAAATCTACACAAAAATGGATTTGTTACTGTTTGATG
ATGAGGAAGAACTCCCTAATAATGTATTCAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGAT
```

GAGGTGGGTGAAAATAGTAGAATGAAGAAATTGGTATCAGCAGAAGATGAATCCCCTGATCCACAGAAGCATTACAA
TGGAAGGCCAAGTGCTCAAGTAGTTAAACAAGAAGAGCATGCTCGCAAGTTGATGGAAGCTCAAGAGAGTAGAGAGA
GAGCTAGGAGAATTGCTTCTTTTACAAGTCGGGTAGCTGATTTGCAGCGAG

>GPH10_ananassa_clone20
GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGCCGATACCTCCTCCGCCTCCGACGACTT
CGAACACAGCGGAATCGCTAGCCTCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGTGAAT
AGCGAACTGAGTAGTTTGGATTTGAGAAGAGGATGTAATTGGTAACGGAGAAGAAGACTGTCGACATTTTTGGAGAA
AGCTTTCGGCTTTGAAGTGGAGTGTAGGATAATAACAAACTCGTGATTAAAAGACAGGATTAATGTCAGTGAGGTTT
GGTTGGTTAAGGTGTTAACTGATAAATTTAAGGTCATAGGTTCAAACCTCACGACATATGTAGGGTGTATGAATTAT
TAATAAAAGACAAATTTAATATCAGCCGTTAGATCATATTACGGCCTGATCACTCGACATATGTTGATATACGCCCA
ACTCAAATTCGATATATATTTTCGATATACGTATATTTTATTTTTTTAAAATAATTAAATAACTATTTACGTTGTTT
AACAAAAGAAACAATTGAAGTTAAATTAAGAGCACCGTAACAGCTGAGCAAGAGTACGAGAACAAAAGTATGAGCTA
CATCATTTGTTCATATAGAGAAAATATAGAGGCGATGTTGTAGAAATAATTGAACATTAGAAAATTAAATTACCTAA
AAGCCGATGAGTAAAATAATAACGAACTCGTAACCTAAAAGCGGCTTCATATCATCCGCTTGATCATATATGCGGGT
GTGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAATAAAACCCACACG
ACGCCGTTTTGCTCCAATCCCCCTTTCTTCTTCAACCCCATAGTCGCCTCAGCTCAGTTCCATTTGTCTCAGATGCG
ATGGCCTCCGGCGACCCAATCTCCGACTACACCCAAACACATCGCATTGTCCTTCTAATCGACCTCAACCCACTCCT
CCATCTCCAAGATCCAACCCAATTCCTCACCTCTGTCCTCTCCTCAATCAAAACCCTAACCTCCTTCCCTTCTCTCT
CTTCCTCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCTCTCCTCTCCTCTCCTCTCCGCCTCCAAGCTCCCGT
CTTCGTCTCTAACGATCTCTTTCAACTCGCCGGAAGACACATATCGATCCCTATCTCAAACCCTGGCGTCTCTCTCG
TTTGACCGGAAGTTGACCGGGTCCGATTCGCCGCGGGGAACGCTTGTTGCGGCTGCGATGCGGCAGCTGGTACATGA
TTACGCTTGGGAGCAGGTGATCTGCGACGCCGTGGCGGCGGAGACAGGTACGTTTTCGAATTGCTGTGGTTTGAGGT
CTAATTTGGCTGTTGTGTTTTTACCGGCGTGTCAATTTGTGAATGAGTTCTTGAATTGTGAGTTGAATTGTGAGGGT
TTGGAGGATTTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGAGGCATATGTGTATAGAGA
TATTCAATTGAGTTGGGTTGATGTGAGGTATGGATTCGATAGCGGTGAGGATGAGGTAGTTGGATTGAAATGTGGTG
TTTTCGAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGATTGTGCTTGGTTCGGCTCTTGTT
CCATTTGGTTTGATTTATCCAGAGATTGGGGTGTCATCTAGGATTTTCGGGTGTAATGATCGATATAAGAAGGTTAG
AGCGCATTTGAGTCTTGAGATATCAGATGTAAAGGGGATGCCTTTGGAGTGCAAGTTTTGTGATCTTGAGTTGGCTG
ATTTGAAAATGTTGTGTAGGAGTAGAGGTGATGATCGCTTGTTTTCGGTGGAAGGCATGAACTCGCAGACAAGAGGT
CATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAATGGAGTGTCGAAGATTCAGGTTAAGGCTTTGCAGAAGGA
TAGTGAGTTTGGGAAATTTAAGGGGGAATTGTCGGATCTGATTCTGGTCTATGAAGTTTCAGGAAAAGATGGAAAAG
AAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAGTGCTATCAAGTGAAATTGGGTGAGTTTGTACCGAGGAA
ATTGCCACCTGTTTGGCAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTAGTTTCTATTTCAAATG
ATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTACAGTTTCTTCAGCTCTTATTTTTGTTATGGATGAAGGA
ATTCACCCTCATAAAAAAGGGCATGGCATTGGTGCAGAGAATAAGGGTCAGTCTCGTCCAAAGATGAAGAATGAGAT
GTGCAAACCTGATGCTGATTTGAACGACTTTTGTGGGTCGCAAACTGGGCCTTCACCATCTAATAAGCATTCTGCTG
AGATTGATGGAAAGAAAAAAAGTAGCGAAAGAAGTTCACATTCACTCAAAGATCTCACCCGGAGTTCTTTCTGTAAG
GCAGCATTCGAATTTTCAGACTTACATTTGGAAGAGGTTTACTTTGCCAGGCAACGTAGCAGCTCAAAAAAGTTGAA
ATTTCTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGAGGAGTCTAAGGTGCACCAGGAAA
AACAAAAGGAGATGAGCAATAGGTTGGATTTGTTGCACCAAGAGAGCGAACAGCCAATGTCGTCATCTGGTTCAGCT
GGAGAAATTTCTTTCCCTGTCGCCTTTGGAGTACAGGATGAAGCTGCTCAGGAACATAGATTACAAACCTCAGAAGA
TTTTTTCTGTAATTTCTCTGATAAGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTGGGGGCATTCACACATC
GGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAACATAGCACAACAACCCCTTCAGAAGATCAAACTCCTGTA
AAATCTGACAATCTTGATGATTTGGTTACTGCTGAGCTGTTAAAACTTTTACTCAGAGATCCCAAGGATATGGTTGC
CAGGCACAAAAGCTATGATTCATCTTCTCAAGCATCTGATCCTGGATGTGAAGGCTTTACTTCAGAAATAATAGTTC
GAGAGTATCCTTTCATTTATCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAATTCTACTTTAATGCTA
TGTAAACTTTGCCCCTTGTTACTGTTACACTTTTCCTTCACTAGCACAAAGATATGAATTACAGATACTTTTCCGGA
TGGAGATTTTACAATCAGAAGTTGGAGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGTGAAACATATTTGCACG
CTTTTGGAGACCATTCGTGCTCGGTGTCATCTGGAGGGAGGCTTCTTTGGTGACTGGACCCTAGAAAATTATGCTGG
AAAGATTATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAACTATTATTTTCTAATGAAATTT
GTATTCATGAACACTGAAATGGTAGATACTCAGTTATTTACAATGAAACTCCAATATATGTTTATGGTTTGCCTGTT
AATGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTATGTGATTGATTTGTATGCCCTCCCAA
AAGGCCTTTGGGGGTAGTATGAAGAAGGGAGACATTGACCGTCAAAACTATTATCTCCTTATTTTACGTACAAAATT
GATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAAAAGCTAAATACCTCCTTTGCATAATTTCA
TATGACTTAAGTGACTTTCCTTATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGTATGCATACAACTTTT

GCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTCCAGACCCAAAAAAAACTGCCCAAATTTATGTGA
ACACACTGCATTTATGTTTGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCAAATTTTCAGGTACTGTCA
GACTCTTGAAGACGTGGTTCATAAAATCTACACAAAAATGGATTTGTTACTGTTTGATGATGAGGAAGAACTCCCTA
ATAATTTATCCAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGATGAGGTGGGTGAAAATAGT
AGAATGAAGAAATTGGTATCAGCAGAAGATGAATCCCCTGATCCACAGAAACATTACAATGGAAGGCCAAGTGCTCA
AGTAGTTAAACAAGAAGAGCATGCTCGCAAGTTGATGAAAGCTCAAGAGAGTAGAGAGAGGGCTAGGAGAATTGCTT
CTTTCACAAGTCGGGTAGCTGATTTGCAGCGAG

Polymorphic segment of GPH10: fragment between primers 10PPR1 and 10AB22

>10PPR1AB22_vesca
AACGGAGAAGAAGACTGTCGACATTTTAAGAGAAAGCTTTCAGCTTTGAAGTGTAGGATAATAACAAAGAAACTCGT
TATCTGAAAGACAAGTTTAATATCAGCCGTTGGATCATATTACGGCCCTGATCGCTCGACATAATTCGATATATATA
TATATATATTTTTTTTTTTTCTAAAAAAAAAAATCGATATACAGTATATTTTTTTTGAATTAATTAAATGAGTATTT
AGATCGCTTAAAAAGATAAACAATCGAAATTGGTTTAAGAACACCATAGGAGCAAGAGTATGAGAACAAAAGTATGA
GCTACACTGTTTGCTCATCGGTTTATTTATATGGAGAAAATATCAAGGTGATGTTGTATAAACAATTAAACATTACA
AAATCAAATTACCTAACAATGAACCATTTTCAGACATGTAAAATCATAAAATTAAAAGGTTCGAGTCGCATATGAGT
TTGTCGAGCTGATCAAATACCACAGTTTACTTGACTGAACAAACTTACGTAACGAGTCAAACGAGCTAAAAACGAGT
CGAATAAAAATCGGGCACCATCAATATCGAGACTATGTAAGAGCCGAGGAGTAAAATAATAACAAACTCGTTATCTA
AAAGACAGGTTTAATATCAGCCGTTGGACCATATGTACAGGTGTGATTCGAAAACCGAAGTTAACCCGCCAAACCCT
CATTCCCAATTTTCATTCCCACCAAAAACAAAACC

>10PPR1AB22_nubicola
AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAAGTGTAGGATAATAACAAAGAAACTCGT
TATCTGAAAGACAGGTTTAATATCAGCCGTTGGATCATATTACGGCCCTGATCGCTCGACATAATTCGATATATATA
TATATTATTTTTTTTCTAAAAAAAAAAAATCGATATACAGTATATTTTTTTTGAATTAATTAAATAAGTATTTAGATCG
CTTAAAAAGATAAACAATTGAAGTTGGTTTAGAAGCATCATAGGAGCAAGAGTACGAGAACAAAAGTATGAGCTACA
CTGTTTGCTCGTCGGTTTATTTATATGGAGAAAATATCAAGGTGATGTTGTATAAACAATTAAACATTACAAAATCA
AATTACTTAACAATGAACCATCTTCAGACATGTAAAATCAGAAAGTTAAAAGGTTCGAGTCGCATATGAGTTTGTCG
AGCTGATCAAATACCACAGTTTACTTGACTGAACAAACTTACGTAACGAGTCAAACGAGCTAAAAACGAGTCGAATA
AAAATCGGGCACCATCTATATCGAGACTATGTAAGAGCCGAGGAGTAAAATAATAACAAACTCGTTATCTAAAAGAC
AGGTTTAATATCAGCCCTTGGACCATATGTACGGGTGTGATTCGAAAACCGAAGTTAACCCGCCAAACCCTCCTTCC
AATTTTCATTTCCCACCAAAAACAAAACC

>10PPR1AB22_mandshurica
AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAAGTGTAGGATAATAACAAAGAAACTCGT
TATCTGAAAGACAGGTTTAATATCAGCCGTTGGATCATATTACGGCCCTGATCGCTCGACATAATTCGATATATATA
TATTATTTTTTTCTAAAAAAAAAAAAAATCGATATACAGTATATTTTTTTTTGAATTAATTAAATGAGTATTTAGATC
GCTTAAAAAAAATAAACAATCGAAGTTGAATTAGGAGCACCATAGGAGCAAGAGTATGAGAACAAAAGTATGAGCTA
CATTGTTTGCTCGTCGGTTTATTTATATGGAGAAAATATCAAGGTGATGTTGTATAAACAATTAAACATTACAAAAT
CAAATTACTTAACAATAAACCATCTTCAGACATGTAAAATCAAAAAGTTAAAAGGTTCGAGTCGCATATGAGTTTGT
CGAGCTGATCAAATACCACAGTTTACTTGACTGAACAAACTTACGTAACGAGTCAAACGAGCTAAAAACGAGTCGAA
TAAAAATCGGGCACCATCAATATCGAGACTATGTAAAAGCCGAGGAGTAAAATAATAACAAACTCGTTATCTAAAAG
ACAGGTTTAATATCAGCCGTTGGACCATATGTACAGGTGTGATTCGAAAACCGAAGTTAACCCGCCAAACCCTCATT
CCCAATTTTCATTCCCACCAAAAACAAAACC

>10PPR1AB22_nilgerrensis
AACGGAGAAGAAGACTGTCGACATATCTAGAGAAAGCTTTCAGCTTTGAAGTGGAGTGTAGGATAATAACAAACTCG
TTATCTGAAAGACAGGTTTAATATCAGCCGTTGGATTATATTCCGGCCCTGATCTCTCGACATATGTTGATATACGC
CTGACTCAAATTCTATATACATTTTCGAAAGAATTTTTGTTGTTGAAGTAACTAAATGACTATACGATTGAAGATAG
ATTAAGAGAAACATAGCAACTGAGTAAAAAGTATGAGAACAAAAGTATGAGCTACATTGTTTGCTCCTCGGTCTGTT
TATATGGAGAAAATATAAATGTGATGTTGTAAAAATAATTGAACATTAAAAAAAATTAAATTATCTAATAACGAATC
ATCTTTAGACGTACGTACAAAATCAGAGAGTTAAGAGATTCGAGTTGCTCAAATACCATATTTTACTTGACTTAACA

ATACGAGTCAAACGAGCTAAAAACGAGTCGATTAAATCTAAGAGCCGAGGAGTAAAGTAATAACAAACTCGTTATCT
AAAATACAGGTTTAATATCAGCCGTTGGATCATATATACAGGTGTGATTCGAAAACCGAAGTTAACCCGCCAAACCC
TCATTCCCAATTTTCATTCCCACCAAAAACAAAACC

>10PPR1AB22_viridis
AACGGAGAAGAAGACTGTCGACATTTCTAGAGAAAGCTTTCAGCTTTGAAGTGTAGGATAATAACAAAGAAACTCGT
TATCTGAAAGACAAGTTTAATACCAGCCGTTGGATCATATTACTGCCCTGATCGCTCGACATAATTCGATATATATA
TATATATTATTTTTTTCTAAAAAAAATAAATCGATATACAGTATATTTTTTTTTGAAGTAATTAAATGATTATTTAA
ATCGCTTAAAAAGATAAACAAGAAGTTGGTTTAGGAGCACCATAGGAGCAAGAGTATGAGAACAAAAGTATGAGCCA
CACTGTTTGCTCTTCGGTTTGTTTATACAGAGAAAATATAAAAGTGATGTTGTAGAAACAATTGAACACTAAAAAAT
CAAATTACCTAACAACGAACCATCTTCAGACATACAAGATCAGAAAGTTAAGAGGTTCGAGTCGCACATGAGTTTCT
TGAGGCGATCAAATACCACAGTTTACTTGACTCAACAACTTTACGCATACGAGTCAAACGAGCTAAAAACGAGTCGA
ATAAAAATCGGGCACCATCAATATCGAGACTATGTAAGAGCCGAGGAGTAAAAAATAATAACAAACTCGTTATCTAA
AAGACAGGTTTAATATCAGCCGTTGGACCATATGTACAGGTGTGATTCGAAAACCGAGTTAACCCGCCAAACCCTCC
TTCCCAATTTTATTTCCCACCAAAAACAAAACC

>10PPR1AB22_iinumae
AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAACTTTGAAGTAGTGTAGGATAATAACAA
ACTCGTTATCTAAAAGACAGGTTTAATATCAGCCGTTAGATCCTATTAAGAGCCGAGGAGTAAAATAATAACAAAGT
CGTAACCTAAAAGCGGCTTCATATCATCTACTGGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCGCC
AAACCCAATTTTCATTTCCCACCAAAAACAAAACC

>10PPR1AB22_ananassa_clone2
AACGGAGAAGAAGACTGTCGACATTTTTGGAGAAAGCTTTCATCTTTGAAGTGGAGTGTAGGATAATAACAAACTCG
TTATCTAAAAGGCAGGTTTAATATCAGCCGTTAGATCATATTACGGCCCTGATCACTCGACATATGTTGATATACGC
CCAACTCAAATTCGATATATATTTTCGATATACATATATTTTATTTTTTTAAAGTAACTAAATGACTATGTACATCG
TTTAACAAAAGAAACAATTGAAGTTAAATTAAGAGCACCATAACAGCTGAGAAAGAGTACGAGAACAAAAGTATGAG
CTAAAACAAATAGAGAAAATATAGAGGCGATGTTGTAGAAATAATTGAACATTAGAAAATTAAATTACCTAAAAGCC
GATGAGTAAAATAATAACGAACTCGTAACCTAAAAGCGGCTTCATATCATCCGCTTGATCATATATGCGGGTGTGAT
TCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAACAAAACC

>10PPR1AB22_ananassa_clone7_(same_restriction_pattern_as_clone20)
AACGGAGAAGAAGACTGTCGACATTTTTGGAGAAAGCTTTCGGCTTTGAAGTGGAGTGTAGGGTAATAACAAACTCG
TGATTAAAAGACAGGATTAATGTCAGTGAGGTTTGGTTGGTTAAGGTGTTAACTGATAAATTTAAGGTCATAGGTTC
AAACCTCACGACATATGTAGGGTGTATGAATTATTAATAAAAGACAAATTTAATATCAGCCGTTAGATCATATTACG
GCCTGATCACTCGACATATGTTGATATACGCCCAACTCAAATTCGATATATATTTTCGATATACGTATATTTTATTT
TTTTAAAATAATTAAATAACTATTTACGTTGTTTAACAAAAGAAACAATTGAAGTTAAATTAAGAGCACCGTAACAG
CTGAGCAAGAGTACGAGAACAAAAGTATGAGCTACATCATTTGTTCATATAGAGAAAATATAGAGGCGATGTTGTAG
AAATAATTGAACATTAGAAAATTAAATTACCTAAAAGCCGATGAGTAAAATAATAACGAACTCGTAACCTAAAAGCG
GCTTCATATCATCCGCTTGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCA
ATTTTCATTTCCCACCAAAAATAAAACC

>10PPR1AB22_ananassa_clone18
AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAAGTGGAGTGTAGGATAATAACAAACTCG
TTATCTAAAAGACAGGTTTAATATCGGCCGTTAGATCACATTACGGCCCTGATCACTCGACATATGTTGATATACGC
CTAACTCAAATTCGATATATATTTTCGATATACATTTTTTTTTTAAGTAACTAAATGACTATTCGATATATATTTTC
GATATACATTTTTTTTTTAAAGTAACTAAATGACTATTTACGTCGGTTAATAAAAGAAACAATTGAAGTTAAATTAA
GAGCACCATGACAGAGTACGAGAACAAAAGTATGAGCTACATTGTTTGCTCGTCGGTTTGTTCATATGGAGAAAATG
TAGAGGCGATGTTGTAGAAATAATTGAACATTAGAAAATTAAATTACCTAAAAGCCGATGAGTAAAATAATAACAAA
CTCGTAACCTAAAAGCGGCTTCATATCATCCACTGGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCG
CCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAACAAAACC

>10PPR1AB22_ananassa_clone19
AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTTAGCTTTGAAGTGGAGTGTAGGATAATAACAAACTCG
TTATCTAAAAGACAGGTTTAATATCAGCCGTTAGATCATATTACGGCCCTGATCACTCGACATATGTTGATATACGC

```
CTAACTCAAATTCGATATATATTTTCGATATACATTTTTTTTTTAAGTAACTAAATGACTATTCGATATATATTTTC
GATATACATTTTTTTTTTAAAGTAACTAAATGACTATTTACGTCGGTTAATAAAAGAAACAATTGAAGTTAAATTAA
GAGCACCATGACAGAGTACGAGAACAAAAGTATGAGCTACATTGTTTGCTCGTCGGTTTGTTCATATGGAGAAAATA
TAGAGGCGATGTTGTAGAAATAATAGAACATTAGAAAATTAAATTACCTAAAAGCCGATGAGTAAAATAATAACAAA
CTCGTAACCTAAAAGCGGCTTCATATCATCCACTGGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCG
CCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAACAAAACC
```

```
>10PPR1AB22_ananassa_clone20
AACGGAGAAGAAGACTGTCGACATTTTTGGAGAAAGCTTTCGGCTTTGAAGTGGAGTGTAGGATAATAACAAACTCG
TGATTAAAAGACAGGATTAATGTCAGTGAGGTTTGGTTGGTTAAGGTGTTAACTGATAAATTTAAGGTCATAGGTTC
AAACCTCACGACATATGTAGGGTGTATGAATTATTAATAAAAGACAAATTTAATATCAGCCGTTAGATCATATTACG
GCCTGATCACTCGACATATGTTGATATACGCCCAACTCAAATTCGATATATATTTTCGATATACGTATATTTTATTT
TTTTAAAATAATTAAATAACTATTTACGTTGTTTAACAAAAGAAACAATTGAAGTTAAATTAAGAGCACCGTAACAG
CTGAGCAAGAGTACGAGAACAAAAGTATGAGCTACATCATTTGTTCATATAGAGAAAATATAGAGGCGATGTTGTAG
AAATAATTGAACATTAGAAAATTAAATTACCTAAAAGCCGATGAGTAAAATAATAACGAACTCGTAACCTAAAAGCG
GCTTCATATCATCCGCTTGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCA
ATTTTCATTTCCCACCAAAAATAAAAC
```

```
>GPH23_ananassa_clone3
CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTACTAACGACCTTTACAGTGAGAGTGTGACCAGAG
GTGCCTGGGCGGAGCTGCCCAACCTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTCGAGTCTGCCATTTGATAAAA
GACCTGCCAGAATCCACGCCACCAAACATCTTTAGCACTAATCCAATCCATAACAACTTCATAAAACACACATAGCA
TCAACATGCAATAATGTGGGTCCATAAGAACCATGTGCATGACATAAGATTCCTCAAGCTTCGATTTCCTAATTTGC
TTCAAAAGAAGTAAGTTCAGAGTCACTCAAACCCTAATATAGATCTCAAATTTAATGAAACATATTCCTAAGAGCCT
ACACAAATATAAAATCGTAACTGAACTTAATCTGAAACTGTCGTATAAATTGTAAATCGATCAAAACCAAACTTCAT
GTTCAGATTCACAGACCGTATCAGAGATAGCATACAAGTGACCTCTGAAACAAAACATAATTCCAACAAGATCGCAA
ACATTCGAGATTAAATACGATGAGCTATGAGACAACTATTCCATGCAAATCTAACAAAAAAGAATAAAGGGATCTGG
AGAATTATGGGTTAGAGGTGACCTTCAGAGTTTGGGTGAAACACAACTGGGGACGACCTACACCGAGGAGGAACTGC
CAAAATCTATCTGAAACCTAACAAATAAAAAGGGTCTATCTGTCAATAAACGAGCTCCCTATCGTCCATCTCCAATC
TTGTAAGGGGTGATCCTTACGCTTCCCTTTGTCCTCTCCCCCATCTACTAAGTAGACGCTAGCTGCGGATTGTTATT
ATGTTTTGGATAGAATACCTTTGCAAAATTGGAAGCTCCAGCTCCTCCTTGTTTTTCGGCAAGAGAAAGGCCAAAAT
ATCAGACCGTTCCGACGCCGGAGCTTCCTCAGAAAGCCTACCATCCGCAACATCGTTGCCAGGCCTTGCGAGGTTTG
CCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTATAGGAGATTGTTGAACAAAAAAGAAAAACATAACATATGATGAA
TGAATCATCAAATTAATTAATCAAGGTGACAAAGAAAGATTATATCCTTCCATTCCTAAGTCAAAAACCATTACAAT
GTACCGCCGGCAAAATGCTGCTAATAGAATTGACATTGTAAGTGGGGGATAGTGTCACGAGCTGCAATAGGAGGTAG
TGTCATCGTGACACTCTAGAGGTGATGCTTAAGAGGGTCACAAGGTCAATGGCAAAGCATGAGGGTTAAAGAGGATG
TTTACTGACATGTTGAAAGACAATGTCGTAATTAGTTAAAGTGAAGTACTGTGAAGTTAGTATTTCGAAAAACTGTA
ACATCGGAAGGGGTTCAATACATTTGACGACATATTTTTATGAAGTTATTAGGAATTAGTTACGAGAGATGGTTTTT
TCTTAGAATATTTTGATTTTGATGTTTCCTTGACACACATTATATTTCTCCATGTTCTTCTATGTATAAGTAATTTT
CTGTATCACTTAGAAACATTTCTTACTCTTTCCAGAAGCATCTCCAAACATCCCCCTAAACCAATAGCCCTAACATG
TCAATGTCACATGTCAATAGATGAAAGATCAACCTAAATGGTACCATATGTCCATACATAAAAGGACCCAAAAAAAT
AAATAAAGAAATAAATATGCACCTTCATTTTTAAGCGCCAGAAAAAGTAGAGAAGAATATAAGGTTTGAAGTGATCA
AGGGGATAAGCAGTTTAAGGTCGACTTGTTTGGAAACAATGCTAACCACCACCACTGCCACTCACAGTCTCAGCTCC
TCCTCCTCTGCTTCCCAACTCCCATCGCTCTTCCACTCTCTATCACAAAACCCAATCTCCCTCAGATTCTCCTCCAC
ATTAAAGCTAACCAAAACCAGAACCAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGCCAGCTCCCTGCTCTGA
GAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCCTCCACAGTGCCGTCGGAGATGAAGGCGTGGGTG
TA
```

```
>GPH23_ananassa_clone4
CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTACTAACGACCTTTACAGTGAGAGTGTGACCAGAG
GTGCCTGGGCGGAGCTGCCCAACCTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTCGAGTCTGCCATTTGATAAAA
GACCTGCCAGAATCCACACCACCAAACTCTTTAGCACTAATCCAATCCATAACAACTTCATAAAACACACATAGCAT
CAACATGCAATGATGTGGGTCCATAAGAACCATGAGTATGACATAGAGTCTTCAAGCTTCGATTTCCTTATTTGCTT
CGAAAGAAGCAAGTTCAGAGTCACACAAACCAGAATATAGATCTCAAATTTAATGAAACATATTCCTAAGAGCCTAA
AGAAATATAAAATCGTAACTGAACTTAATCTGAAATTGTCGTATAAATTGTAAATCGATCAAAAACAAACTTCAAGT
```

```
TCAGATTCAGAGAGCAGATCAGAGATAGCGTACAAGTGACCTAAGAAACAAAACAAAATTCCAACAAGATCGCAAAC
ATTCGAGATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAAGAGAATAAAGGGATCTGGAG
AATTAGGGGTTAGAGGTGACCTTAAGAGTTTGGGTGAAACACAACTGGGGAAGACAGAGACAGAGGAGGAACTGCGA
AGATCTATCTGAAACCAAACAAAGTAAAAAGGGTTTAGCTGTCAGTAACGAGCTCCTAACCGTCCATCTCCAATCTT
GTCAGGGGTGATCCTACGCGTCCACTTGTCCTTTCCCAGTTCTAACTATGTAGACGCTAGCTGCGGATTGTTATTAT
GTTTTGGATAGAATACCTTTGCAAAATAGGAAGCTCCTCCTTGTTTTTCTGCAAGAGAAAGGCCAAAATATCTGACC
ATTCCGACGCCGGAGCTTCCTCAGAAAGCCAGTACCATCCGCAACATCGATGCCAGGCCTTGCGAGGTTTGCCTCCG
CTTCTTTGGATTGTGTTTTTCGTGGTTTAGGAGATTGTTGAACAAAAAAGAAAAACATATATGATAAATGAATTATC
AAATTAATTAATCAAGGTGACGATACAAGATGAGAACACCAAAGGTTCAATAGTGTGTACTCTCAAGCCTAATACTA
ACACAACAAAGAAAGATTCTATCCTTCCATTCCCAGATCAAAAACCACTCTAATGTACCGCCGGCAAAGTGCTGCTA
ATAGAATTGACATTGTAAGTAGGGGATAGTGTCACGAGCAGCTCTAGGAGGTAGTGTCATCGTGACACTTTATTGGG
GTGGATGCTAAGGGGGTTCAGGTTATGAGCAAGCATGGGGGGTAAGGGGGATATCTACTGGCATTTTGAATGACAAT
GTTGTAAATGAAAACTTATATTTCAAGGTATTTTGATTTAATATTTAGAAAAACTGTAACATCAAAAGGGGTTCAAT
ACATTTGCCGACATATTTTTATGAGGTTTTTATGAATTAGTTATGAGAGATGGTTTTCCTTGGACTATTTAGATTTT
GATGTTTCCTTAACACACATTATATTTCTCCATTTTCTTGTAAGTAATTTTCTGTATCACTTAAAAACATTTCTTAC
TCTTCCCAGAAACATCTCCAAACATCCCTAAACCGATAGCTCTAACATGTCAATGTCAATAGATGAAAGATCAACCT
AAATGGTACCATATGTCCATACATAAAAAGACCCAAAAAGAAAATAAATAAGCACCTTCATTTTTAAGCGCCATAAA
AAGTAGAGAAGAATACAAGGTTTGAAGTGATCAAGGGGATAAGCAGTTTAAGGTCGACTTGTTCGGAAACAATGCTA
ACCACCACCACTGCCACTCTCAGCTGCTCCTCCTCCTCTGCTTCCCAACTCCCACCACTCTTCCACTCTCTATCACC
AAACCCAATCTCCCTCAGATTCTCCTCCACATTACAGCTAACCAAAACCAGAACCAGACCAACCCTTAAAACTCTCA
CTCGCCAAAAATGCCAGCTCCCTGCTCTGAGAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCCTCC
ACGGTGCCGTCGGAGATGAAGGCGTGGGTGTA

>GPH23_iinumae_clone2
CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTGCTAACGACCTTTACAGTGAGAGTGTGACCAGAG
GTGCCTGGGCGGAGCTGCCCAACCTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTTGAGTCTGCCATTTGATAAAA
GACCTGCCAGAATCCACGCCACCAAACTCTTTAGCACTAATCCAATCCATAACAACTTCATAAAACACACATAGCAT
CAACATGCAATAATGTGGGTCCATAAGAACCATGAGTATGACATAGAGTCTTCAAGCTTCGATTTCCTTATTTGCTT
CGAAAGAAGCAAGTTCAGAGTCACACAAACCAGAATATAGATCTCAAATTTAATGAAACATATTCCTAAGAACCTAA
AGCAATATAAAATCGTAACTGGACTTAATCTGAAATTGTCGTATAAATTGTAAATCGATCAAAAACAAACTTCAAGT
TCAGATTCACAGAGCAGATCAGAGATAGCATACAAGTGACCTAAGAAACAAAACAACATTCTAACAAGATCGCAAAC
ATTGGAGATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAAGAAACTAAAGGGATCTGGAG
AATTAGGGGTTAGAGGTCACCTTAAGAGTTTCGGTGAAACACAACACAACTGGGGAGACAGAGACAGAGGAGGAACT
GCGAATATCTATCTGAAACCAAACAAAGTAAAAAGGGTTTAGCTGTCAGTAACGAGCTCCTAACCGTCCATCTCCAA
TCTTGTCAGGGGTGATCCTACGCGTCCACTTGTCCTCTCCCAGTTCTAACTATGTACACGCTAGCTGCGGATTGTTA
TTATGTTTTGGATAGAATAGAATACCTTTGCAAAATAGGAAGCTCCTCCTTGTTTTTCGGCAAGAGAAAGGCCAAAA
TATCTGACCATTCCGACGCCGGAGCTTCCTCAGAAAGCCGGTTCCGTCCGCAACATCGATGCCAGGCCTTGCGAGGT
TTGCCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTTTAGGAGATTGTTGAACAAAAAAGAAAAACATAACATATGAT
GAATGAATTATCAAATTAATTAACCAAGGTGACAATACAAGATGAGAACACCAAGGGTTCAATAGTGTGTACTCTCA
AGCCTAATACTAACACAACAAAGAAAGATTTCTATCCTTCCATTCCCAAATCAAAAACCACTACAATGTACCGTCTA
ATTGAATTGACATTGTAAGTGAGAGATAGTGTCACGAGCTGCATTGGGAGATAGTGTCATCGTGACACTCTATGAAG
GGGATGCTTAAGAGGGTCGCATCAATGACAAACATGAGGGCAAATAGAAGGTCTACTGGCATGTCGAATGACAATGT
CGTAATTAGTTAAGTGAAACTTATATTTCAAGGTACTTTGACTTAGTATTTAGAAAAACTGTAACATCGAAAGGAGT
TCAATACATTTGACGACATATTTTTATGAGGTTTCTATAAATTAGTTATGAGAGATGGTTTTCCTTGGACTATTTTG
ATTTTGATGTTTCCTTAACACACATTATATTTAAAGTAATTTTCCGTATCACTTAAAAACATTTCTTACTCTTTCCA
GAAGCATCTCCAAACATCTCCCTAAATGTCAATGTCAATAGATGAAAGATCAACCTAAATGGTACCATATGTCCATA
CATAAAAAGACCCAAAAAGAAATAAATAAGCACCTTCATTTTTAAGCGCCATAAAAAGTAGAGAAGAATATAAGGTT
TGAAGTGATCAAGGGGATAAGCAGTTTAAGGTCGACTTGTTCGGAAACAATGCTAACCACCACCACTGCCAGTCTCA
GCTGCTCCTCCTCCTCTGCTTCCCAACTCCCACCACTCTTCCACTCTCTATCACCAAACCCAATCTCCCTCAGATTC
TCCTCCACATTACAGCTAACCAAAACCAGAACCAGAACCAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGCCA
GCTCCCTGCTCTGAGAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCCTCCACGGTGCCGTTGGAGA
TGAAGGCGTGGGTGTA

>GPH23_iinumae_clone5
CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTGCTAACGACCTTTACAGTGAGAGTGTGACCAGAG
GTGCCTGGGCGGAGCTGCCCAACCTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTTGAGTCTGCCATTTGATAAAA
```

```
GACCTGCCAGAATCCACGCCACCAAACTCTTTAGCACTAATCCAATCCATAACAACTTCATAAAACACACATAGCAT
CAACATGCAATAATGTGGGTCCATAAGAACCATGAGTATGACATAGAGTCTTCAAGCTTCGATTTCCTTATTTGCTT
CGAAAGAAGCAAGTTCAGAGTCACACAAACCAGAATATAGATCTCAAATTTAATGAAACATATTCCTAAGAACCTAA
AGCAATATAAAATCGTAACTGGACTTAATCTGAAATTGTCGTATAAATTGTAAATCGATCAAAAACAAACTTCAAGT
TCAGATTCACAGAGCAGATCAGAGATAGCATACAAGTGACCTAAGAAACAAAACAACATTCTAACAAGATCGCAAAC
ATTGGAGATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAAGAGACTAAAGGGATCTGGAG
AATTAGGGGTTAGAGGTCACCTTAAGAGTTTCGGTGAAACACAACACAACTGGGGAGACAGAGACAGAGGAGGAACT
GCGAATATCTATCTGAAACCAAAACAAGTAAAAAGGGTTTAGCTGTCAGTAACGAGCTCCTAACCGTCCATCTCCAA
TCTTGTCAGGGGGGATCCTACGCGTCCACTTGTCCTCTCCAGTTCTAACTATGTACACGCTAGCTGCGGATTGTTAT
TATGTTTTGGATAGAATAGAATACCTTTGCAAAATAGGAAGCTCCTCCTTGTTTTTCGGCAAGAGAAAGGCCAAAAT
ATCTGACCATTCCGACGCCGGAGCTTCCTCAGAAAGCCGGTTCCGTCCGCAACATCGATGCCAGGCCTTGCGAGGTT
TGCCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTTTAGGAGATTGTTGAACAAAAAAGAAAAACATAACATATGATG
AATGAATTATCAAATTAATTAACCAAGGTGACAATACAAGATGAGAACACCAAGGGTTCAATAGTGTGTACTCTCAA
GCCTAATACTAACACAACAAAGAAAGATTTCTATCCTTCCATTCCCAAATCAAAAACCACTACAATGTACCGTCTAA
TTGAATTGACATTGTAAGTGAGAGATAGTGTCACGAGCTGCATTGGGAGATAGTGTCATCGTGACACTCTATGAAGG
TGATGCTTAAGAGGGTCGCATCAATGACAAACATGAGGGCAAATAGAATGTCTACTGGCATGTCGAATGACAATGTC
GTAATTAGTTAAGTGAAACTTATATTTCAAGGTACTTTGACTTAGTATTTAGAAAAACTGTAACATCGAAAGGAGTT
CAATACATTTGACGACATATTTTTATGAGGTTTCTATGAATTAGTTATGAGAGATGGTTTTCCTTGGACTATTTTGA
TTTTGATGTTTCCTTAACACACATTATATTTAAAGTAATTTTCTGTATCACTTAAAAACATTTCTTACTCTTTTCCA
AACCATCTCCAAACATCTCCCCAAATGTCCATGGCCATAATAGATGAAAGATCAACCTAAATGGTACCATATGTCCA
TACATAAAAAGACCCAAAAAGAAATAAATAAGCACCTTCATTTTTAAGCGCCATAAAAAGTAGAGAAGAATATAAGG
TTTGAAGTGATCAAGGGGATAAGCAGTTTAAGGTCGACTTGTTCGGAAACAATGCTAACCACCACCACTGCCAGTCT
CAGCTGCTCCTCCTCCTCTGCTTCCCAACTCCCACCACTCTTCCACTCTCTATCACCAAACCCAATCTCCCTCAGAT
TCTCCTCCACATTACAGCTAACCAAAACCAGAACCAGAACCAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGC
CAGCTCCCTGCTCTGAGAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCCTCCACGGTGCCGTTGGA
GATGAAGGCGTGGGTGTA

>GPH23_mandshurica_clone3
CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTACTAACGACCTTTACAGTGAGAGTGTGACCAGAG
GTGCCTGGGCGGAGCTGCCCAACCTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTCGTGTCTGCCATTTGATAAAA
GACCTGCCAGAATCCACACCACCAAACTCTTTAGCACCAATCCAATCCATAACAACTTCATAAAACACACATAGCAT
CAACATGCAATAATGTGGGTCCACAAGAACCATGAGTATGACCATAGTCTTCAAGCTTCGATTTCCTTATTCGCTTC
GAAAGAAGCAAGTTCAGAGTCACACAAACCAGAACATAGATCTCAAAAATTTAATGAAACATATTCCTAAGAACCTA
AAGAAATATAAAATCGTAACTGAACTTATTCTGAAATTGTCGTATAAATTGTAAACCGATCAACACAAACTTCAAGT
TCAGATTCACAGAGCAGATCAGAGATAGCATACCAGTGACCTAAGAAACAAAACAACCTTCTAACAAGATCGCAAAC
ATTGGAGATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAAGAGAATAAAGGGATCTGGAG
AATTAGGGGTTAGAGGTCACCTTAAGAGTTTGGGTGAAACACAACCTGGGGAGACAGAGACAGAGGAGGAACTGCCA
AGATCTATCTGACCACCAAACCAAGTaAAAAGGgTtaaGCTATCAgTAACgAgCtCCtAACCGTCCaTCTCCAATCt
TGTCagGgGtGAtCCTACGCGTCCACTTGTCCTCTCCCAGTTCTAACTGTGTAGACGCTAGCTGCGGATTGTTATTA
TTTTTTGGATAGAATACCTTTGCAAAATAGGAAGCTCCTCCTTGTTTTTCGGCAAGAGAAAGGCCAAAATATCTGAC
CATTCCGACGCCGGAGCTTCCTCAGAAAGCCAGTACTATCCGCCACATCGATGCCAGGCCTTGCGAGGTTTGCCTCC
GCTTCTTTGGATTGTGTTTTTCGTGGTTTAGGAGATTGCTGAACAAAAAAGAAAAACATAACATATGATGAATGAAT
TATCAAATTAATTAATCAAGGTGACAATACAAGATGAGAACACCAAAGATTCAATAGTGTGTACTCTCAAGCCTAAT
ACTAACACAACAAAGAAAGATTCTATCGTTCCATTCCCAAATCAAAAACCATTACAATAGACCGCCGGCAAAGTGCT
GCTAGGAGGTAGTGTCATCGTGACACTCTATGAAGGTGATGCTTAAGAGGGACACGTCAATAGCAAGTATGAGGGAA
AAAAAGGATGTTTACTGGCATGTCGAATGATAATTTCGTAATTAGTTAAGTGAAACTTATATTTCAAGGTACTTTGA
CCTAGTATTTGGAAAAACCGTAACATCTAAATACACTTGACGACATATTTTTATGAGATTTCTATGAATTAGTTATG
AGAGATGGTTTTCTTTAGACTATTTTGATTTTGGATGTTTCCTTAACACACATTATATTTCTCCATTTTCTTTTTAT
GTATAAGTAATTTTCTGTATCACTTAAAAACATTTCTTACTCTTTCCAGAAGCATCTCTATCCCCCTAAACCAAATA
CACTAACATGTCAATGTCCACCCAAAAGAAATAAATAAGCACCTTCATTTTTAAGCGCCATAAAAAATAGAGAAGAA
TATAAGGTTTGAAGTGAACAAGGGGATAAGCAGTTTAAGGTCGACTTGTTCGGAAACAATGCTGACCACCACCACTG
CCACTCTCAGTCTCAGCTCCTCCTCCTCTACTTCCCAACTCCCACCACTCTTCCACTCTCTAACACCAAACCCAATC
TCCCTCAGATTCTCCTCCACATTACAGCTAACTAAAACCAGAACCAGACCAACCCTTAAAACTCTCACTCGCCAAAA
ATGCCAGCTCCCTGCTCTGAGAGTGTCTGCTAAATACGAAGCTTCCCCTGCCACAGCTGAGGCCTCCACGGTGCCGT
CGGAGATGAAGGCGTGGGTGTA
```

From fosmid 10B08

>FvescaParent_10B08Fb
GTATGTCCTATTGATTAGATCGTAAATGATTAATTAGATAGGTAATTACTTTCTTGAAAGCGGGCAAGCCGTGATAG
TCTTGAAAGAGAGCGAGCTTTCTGAAGATGGATTTCCCGTTTGTTTTGAGCCCCGCCGCGTCTGGGTTGCTTGCAAC
CCACGACTCAAGTAGGTCAAGCGAAAGCTGCAAACAAGTGTGTTCATGACCATCAGTACATAGTCTGGAAGTTCAAT
GCTTACCTAATCAACACTACTCAAGTCAAGTGTCCAGTACTGATCGACCTACTTAGACCACGTACGTCATATATTTT
TCTTCTATATATTTTCAAGGCTAGAATGAGAACTAAGAACCTAGCTAGCTAGCTAACCTGATTCTCTGCAAGTCCCA
TCTGGATGATTCCGGTGGGATTTTTAACTTGATCATACGGGTTCTTCTCGTATTCTTCCCACCCTAGGAAGTATGAC
GAGTTCTGGCCATGAGAGTCGCAGCTAGCTTTCTTCGACAACATTTTCGAGCAGTTAAACAGATCAAGAGCTTTTCA
ATTCTGAGAGAGAGAGGGACAGAGGAGCAAAGAGAGCTAGACGTAGATAGAGAAGGTTTTGTAAGCAGCTCAGTTTG
GTTTGTGGAAAATATTTAGGTAAGGCTGCATGGATATAAATAGGTGCTGTTGATTCGTTTTACTTATTAGCTTAAAC
AAACACATATGAGTTGGACCTCATCCGAATCTTTTATCTTAATTCTACTCGTACTTTTTTTTTTTTTT

>FnubicolaParent_10B08Fb
AGTTAGTTAATAATATGTCCTATTGATTAGATCGTAAATGATTAATTAGATAGGTAATTACTTTCTTGAAAGCGGGC
AAGCCGTGATAGTCTTGAAAGAGAGCGAGCTTTCTGAAGATGGATTTCCCGTTTGTTTTGAGCCCCGCCGCGTCTGG
GTTGCTTGCAACCCACGACTCAAGTAGGTCAAGCGAAAGCTGCAAACAAGTGTGTTCATGACCATCAGTACATAGTC
TGGAAGTTCAATGCTNACCTAATCAACACTACTCAAGTCAAGTGTCCAGTACTGATCGACCTACTTAGACCACGTAC
GTCATATATTTTTCTTCTATATATTTTCAAGGCTAGAATGAGAACTAAGAACCTAGCTAGCTAGCTAACCTGATTCT
CTGCAAGTCCCATCTGGATGATTCCGGTGGGATTTTTAACTTGATCATACGGGTTCTTCTCGTATTCTTCCCACCCT
AGGAAGTATGACGAGTTCTGGCCATGAGAGTCGCAGCTAGCTTTCTTCGACAACATTTTCGAGCAGTTAAACAGATC
AAGAGCTTTTCGATTCTGAGAGAGAGGGACAGAGGAGCAAAGAGAGCTAGACGTAGATAGAGAAGGTTTTGTAAGCA
GCTCAGTTTGGTTTGTGGAAAATATTTAGGTAAGGCTGCATGGATATAAATAGGTGCTGTTGATTCGTTTTACTTAT
TAGCTTAAACAAACACATATGAGTTGGACCTCATCCGAATCTTTTATCTTAATTCTACTCGTACTTTTTTTCTTTTT
TCTTTTTCATGTGCATAAGCAAATGCATTTCCGATTAAACATAAAAATGTACTGTCGAAACATCATTTCCAGCCAAA
TCCAAACATTCGCTCTAAAAAAGCTACATTTGCTATTAGATTCAATAACACAAAACCAAGCAAACATTAATCTTCAT
AAATACCAAAATTGGCCTCAATACCCAACTTAAAAACGACCTCAGTCCAGAGGAACCTCAACGTCTCCGTCGGTGAT
CTCCTGCTGCAAGTCCTTGGGGTCCTTCCCATCCAC

>11D02_vesca
GAGCTGCTGTGTGAACCAAATGGTACAGAGAAGCCGTTTGCCAAACCTACCCATGATCCAATCAAATGCATAAACTT
TAAGAACTCAAATACCAAACGATCAAACATAATGACTGAAATGAACAAAAATCAAATGGGCAAAGACTAAATGAGAA
AACAGACCTCTTTGTAAACTGGGTTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCAGCCATT
CCTGTTTATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCAGACAAACACAATTTCATACAGA
AGTCTTCCCACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAATGATAACTTTTCAATAGTATAAGAAGGA
GATGAGGACAGTACCAGAGACTGCAGCTACTTTGTGCCACAGCATAGGATTCATTGCTGTATTCTTCCCTTTGCTTG
GTTTCCTTTCAGTCTCTTCGACTTTCTTCTAAAACGACGTAGTCGGTGCAACTGTGCAAGTCTTCTTGTGATGCAAT
TTTCTTTTCTAGGTGATTTTTTTTTCTTTTATAATTAATTTGGTTTTATTTTTCCAAATAATACCTGAAAGACTTTTT
TTTTCGATAGGATTGCAGTAATTTTTTTTGGACAGTATTACGGGACACTGTGACAGCTTTAGAGTTTGAATCTTAGG
TTGGATGATTTAAGTATCTTAGTTGAATGGATGTTATGACATATTGGTGATTAGTATTAGAGTAATGAGAAAGAGAA
AATAAAATGAAAATACAGTACTGGCAATAAACACAATACGGTGGAGCAATCAACAATGCAATAGATTGACAAAGAAA
TGAAGACCTAAAAAAACCATTGCATTAATGCAATAGTGTTGATATTCCAATCTCTCCTGAATAGTATTACAACTCTC
CTGGACAAGTCATAACTGTGGGGGGTAATGGTGTAAACAAACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTT
GCTGGGCAGACATAGCACCCCATATATCATATCAGATGGGGTTAATGCTACCCAGGTGTGACATATTTGTACAGTTA
AACCTAATTTTGTCTAAAGAATGCTAAAATCGAACTCCCAAGCAACCGAATCTTCTGTTCCCCTGCTTTAGTATGTT
GTGGTTATGCCTCAGCTTCCCCAGCAGCATGAATCCGCTCGTCTGGAGTTACAGCATGAAGCAGCTCATCTCTTGTT
GCAGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGTAGCTCATCTGGCATTGCAGCATGAAGCTGCTCGTCTGG
AGTTGCAGCATTAAGTAGTCCTTCTGGAGTTGCAGCAGGATCCAGGTCCCAACACTTACCAGGTAGGTTAGTCTCTT
CTGCGTCGAGTAACCATGCGGGCACCTGGTGAGAAAAGCGTAACATCTCTCTTCTCGGAATCCATAGAATGGCGCTT
CTGTCCGTATCAGTCCGGTATACTGACCTAAATCCAGCCAACTTCACAAGGGGTGAGACACAAACACCAATCTCTTC
AGAGTAATCATCAAGAACCTCCACCATTTGATATTGGTGCCTCACTTCATCTGGAGTTGAAC

>11D02_viridis

```
GAGCTGCTGTGTGACCAAATGGGTACAGAAGAAGCCNGTTTGCCAAACCTACCCATGATCCAATCAATGCATAAACT
TTAAGAACTCAAATACCAAACGATCAAACATAATGACTGAAATGAACAAAAATCAAATGAGCAAAGACTAAATGAGA
AAACAGACCTCTTTGTAAACTGGGTTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCTGCCAT
TCCTGTTTATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCAAACACAATTTGATACAGAAGT
CTTCACACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAATGATAACTTTTCAATAGTATAAGGAGAAGAT
GAGGACAGTACCAGAGACTGCAGCTACTTTGTGCCATAGCATAGGATTCATTGCTGTATTCTTCCCCTAACTTGGTT
TCCTTTCAGTGTCTTCGACCTTCTTCTAAAACGACGGAGTCGGTGAAACTGTGCAAGTCTTCTTGTGAATTTTCTTT
TCTAGGTGATTTTTTTTTTCTTTTATAATTAATTTGGTTTTATTTTTCCAAATAATACCTGAAAGACTTTTTTTTTT
TTTTTTTGATAGAAATACCTAAAAGACTTCATAAAAGCTGTTAAGGCTTCATTTAGGATTGCAGTAATTTTTTTTGG
ACAGTATTACGGGACACTGTGACAGCTTGAGTTTGAATCTTAGGTGGGATGATTTAAGTATCTTAGTTGAATGGATG
TTATGACATATTGGTGATTAGTATTAGAGTTATGAGAAAATAAAATGAAAATACAGTACTGGCAATAAACACAATAC
GGTGGAGCAATCAACAAAGCAATAGATTGACAAGAAATGAAGACCTAAAAAAAACCATTGCATTAATGCAATAGTGT
TGATTTTCCAATCTCTCCTGAATAGTATTACAACTCTCCTGGACAAGTCATAACTGTGGGGGGTAATGGTGTAAGCA
AACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTTGCTGGGCAGACATAGCACCCCATAAATCATATCAGATGG
GGTTAATGCTACCCAGGTGTGACATATTTGTACAGTTAAACCTAATTTTGTCTAAAGAATGCTAAAATCGAACTCCC
AAGCAACCAAATCTTCTGTTCCCCTGCTTTAGTATGTTGTGATTATGCCTCTGCTTCCCCAGCAGCATGAATCCGCT
CGTCTGGAGTTACAGCATGAAGCAGTTCGTCTCTTGTTGCAGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGT
AGCTCGTCTGGCATTGCAGCATGAAGCTGCTCGTCTGGAGTTGCAGCATTAAGTAGTCCTTCTGGAGTTGCAGCAGG
ATCTAGGTCCCAACACTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNN

>11D02_iinumae
GAGCTGCTGTGTGAACCAAATGGTACAGAGAAGCCGTTTGCCAAACCTACCCATGATCCAATCAAATGCATAAACTT
TAAGAACTCAAATACCAAACGATCAAACATAATGACTGCAATGAACAAAAATCAAATGGGCAAAGACTAAATGAGAA
AACAGACCTCTTTGTAAACTGGGTTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCTGCCATT
CCTGTTTATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCAGACAAACACAATTTGATACAGA
AGTCTTCACACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAATGATAACTTTTCAATACTATAAGAAGGA
GATGAGGACAGTACCAGAGACTGCAGCTACTTTGTGCCATAGCATAGGATTCATTGCTGTATTCTTCCCTTAACTTG
GTTTCCGACGGAGTCGGTGCAACTGTGCAAGTCTTCTTGTGATGCAATTTTCTTTTCTAGGTGTTTTTTTTTCTTTT
ATAATTAATTTGGTTTTATTTTTCCAAATAATACCTGAAAGACTTTTTTTTTTTCGATAGAAATACCTGTTAAGACT
TAAGACTTCATTTAGTATTGCAGTAATTTTTTTGGACAGTATTACGGGACACTGACAGCTTTAGAGTTTGAATCTTA
GGTTGGATGATTTAAGTATCTTAGTTGAACGGATGTTATGACATATTGGTACTTAGTATTAGAGTTATGAGAAAAGA
AAAAATGAAAATACAGTACTGGCAATAAACACAATTCGGAGGAGCAATCAACAATGCAATAGATTGGCAAGAAATGA
AGACCTAAAAAAACCATTGCATTAATGCAATAGTGTCGATTTTCCAATCTCTCCTGAATAGTATTACAACTCTCCTG
GACAAGTCATACCTGTGGGGGGTAATGGTGTAAACAAACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTTGCT
GGGCAGACGTAGCACCCCCTAAATCATATCAGATGGGGTTAATACTACCCAGGTGTGACATATTTGTACAGTTAAAC
CTAATTTTGTCTAAAGAATGCTAAAATCGAACTCCCAAGCAACCGAATCTTCTGTTCCTCTGCTTTAGTATGTTGTG
GTTATGCCTCAGCTTCCCCAGCAGCATGAATCCGCTCGTCTGGAGTTACAGCATGAAGCAGCTCGTCTCTAGTTGCA
GCATGAGGTAGCTCGTCTCTTGTTGCAGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGTAGCTCGTCTGGCAT
TGCAAGATGAAGCTGCTCGTCTGGAGTTGCAGCATTAAGTAGTCCTTCTGGAGTTGCAGCAGGATCCAGGTCCCAAC
ACTTACCAGGTAGGTTAGTCTCTTCTGCGTCGAGTAACCATGCGGGTACCTGGTGGGAAAAGCGTAACATCTCTCTT
CTCGGAATCCATAGAATGGCGCTTCTGTCCGTATCAGTCCGGTATACTGACCTAAATCCAGCCAACTTCACAAGGGG
TGAGACACAACACCAATCTCTTCAGAGTAATCATCAAGAACCTCCACCATTTGATATGGTGCCTCACTTCATCTGGA
GTNNNNN

>11D02_nubicola
GAGCTGCTGTGTGACCAAATGGTACAGAGAAGCCNGTTTGCCAAACCTACCCATGATCCAATCAAATGCATAAACTT
TAAGAACTCAAATACCAAACGATCAAACATAATGACTGAAATGAACAAAAATCAAATGGGCAAAGACTAAATGAGAA
AACAGACCTCTTTGTAAACTGGGTTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCAGCCATT
CCTGTTTATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCAGACAAACACAATTTCATACAGA
AGTCTTCCCACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAATGATAACTTTTCAATAGTATAAGAAGGA
GATGAGGACAGTACCAGAGACTGCAGCTACTTTGTGCCACAGCATAGGATTCATTGCTGTATTCTTCCCTTTGCTTG
GTTTCCTTTCAGCCTCTTCGACTTTCTTCTAAAACGACGTAGTCGGTGCAACTGTGCAAGTCTTCTTGTGATGCAAT
TTTCTTTTCTAGGTGATTTTTTTTTCTTTTATAATTAATTTGGTTTTATTTTTCCAAATAATACCTGAAAGACTTTTT
```

TTTCGATAGGATTGCAGTAATTTTTTTTGGACAGTATTACGGGACACTGTGACAGCTTTAGAGTTTGAATCTTAGGT
TGGATGATTTAAGTATCTTAGTTGAATGGATGTTATGACATATTGGTCATTAGTATTAGAGTTATGAGAAAGAGAAA
ATAAAATGAAAATACAGTACTGGCAATAAACACAATACGGTGGAGCAATCAACAATGCAATAGATTGACAAAGAAAT
GAAGACCTAAAAAAACCATTGCATTAATGCAATAGTGTTGATATTCCAATCTCTCCTGAATAGTATTACAACTCTCC
TGGACAAGTCGTAACTGTGGGGGGTAATGGTGTAAACAAACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTTG
CTGGGCAGACATAGCACCCCATAAATCATATCAGGTGGGGTTAATGCTACCCAGGTGTGACATATTTGTACAGTTAA
ACCTAATTTTGTCTAAAGAATGCTAAAATCGAACACTCCCAAGCAACCGAATCTTCTGTTCCCCTGCTTTAGTATGT
TGTGGTTATGCCTCAGCTTCCCCAGCAGCATGAATCCGCTCGTCTGGAGTTACAGCATGAAGCAGCTCGTCTCTTGT
TGCAGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGTAGCTCATCTGGCATTGCAGCATGAAGCNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN

Clones from unspecific amplification
The size fragments amplified with primers for the vector were not the expected, i.e., they had lower molecular weight than the that had been amplicon cloned.

>11D02_nilgerrensis_unspecific
ATAAAACCTTAAAACCAGTTTTGGAATATCTAATAACAATCACAATAAACTTCTGATAATGAGATATTAATCCTCAC
ATATCTTATTCAACAAGNCCTTAAACAACATTTCAAACCTCACTTAAAAATATTCCTCCACTTTTCATAGGTGTGAC
AATAGGTGCAGTTTTTCCCGTTTGCGCAAAAACTTTTCCCCTCATTGAAATATTTACATTTTTTCTTCTTATAGAAA
GAGTCACGAGGCACCGGAGCTTCCACCGCTGGCTCAGGATGCTGCTGTGCGAACGCCGCCTAATGATGGGGAGGACA
AACCGCCTGGAGATGCTGAGGATAACTCTGATAACCCGCTTGAGGATAGAGAGAATAACTCACCTGAGGATGGGGAT
GATAACTCTGATAACCCGCCTGGGGATGGAGAGGATAATCAATATCCATAAATCAACACAAACCATAAATTGTTCAA
ATAAGTTATTGGTGTGAACCAAATCATAAAACCTTAAACCAGTCTAGAAATGTCTAATAACAAGCACAATAAACTTC
TGGTAATGAGGCATTAATCCTCACATCTTATTCAACAGGCCTTAAACAACATCTCAAACCTCACTTCATCTGGAGTT
GAAC

>11D02_F_ananassa_3_unspecific
AAAGACGTTTTGCAGATGTTCACTCCAGATGAAGTGAGGGAGTGTGATGCCATTAAATGGCAGCAACTATAAGACTT
GGAGAACTGAGTTAGACCTGAACCTAGCTCAGCAGAATGCAGATTGGTGTCTAAGTGTTCCAATGCCTACTGAACTT
GGCGCAGCTAGAGATAATTGGCTGAGAGCTAATAAGATCTGTAAGCTTACCATTATACGGACTATGACAGATGTTGT
GAAAGGTGGTATCCCTGAGAAAGAGGTAGCTAGTGAGTTTTTGGAGGCCATAGCTGAGAGGTTTGCTGTGAGTGACA
AGGCTGAAACCAGCATGTTGCTGGATCAGCTGCATAGCATGAAGTATGACATGAAGCTGAACATTAGGGAGTATATC
CTGAAGATGATAGACATAGCCTCTAAGCTTACTGCCCTTAAGATGACCATAGAGGAGGACTTGGTGGTTATGCTGGT
TCTGAAATCCCTGCCTGTAGAATTTGATCAGCTGAAAACTGCCTATAACACCCAGAAGGATAAGTGGTCGTTGAATG
AGTTAATTGCTGTGTGTGTCCAGGAGCATGAGAGGATTAAGAGGGTAACCATTCACCTAGTTACCACCAAACCTCAG
TGGAACAAAGCTGAGAAGAAAACTGGTCCTTCTAAGAACTTGGGAGTTGGCAAGAAGGCCATGAAGGTCAGTGGTAA
CAAGGGAGGCATTAAGTGTTTTTTCTGCAAGGTTAAGGGTCACATGAAACCTGACTGTGGCAAGTATAAGACCTGGA
AAACTATGATAGGAAATGAACCAGTCAAAAACTTTTATGTTTAGGTTAATTTTAGCTAATGTTCCAACTGAAACTTG
TAGTTTT

>11D02_R_ananassa_3_unspecific
ATCTAAAGGCAGCGTGTGACCAATCCAAAGGGTTGTACTTCTACTTGCTTTCTTTTTAGATTGTTTAGCAAATTTAC
CCTTAACACACTCGATGCAATCCACCAAATCAGAAAAGTCTAAATCCGGCAAAATTCCGCTGTTTTACCAACAATTT
TAATCTCTCCTTTGAGACATGTCCAAGTCTCCTGTGCCAAAGAAAAGCAGACTTCTCATTCATTTTTCTTTTATTTC
CAGTAATAGAATCATCAACACCATTCTCAATCAACAAAACTTCAGTGTGGCAAGTATTAGGTAAAGACCAGTAATCA
TCAATAAACTGAGCAGTAGCAAGATAATGAGAAGCACGACTTATTTTCATTCCATAAGAATCAATCAAAAACTGACA
ATTGTCTTTAACCAAAACAGCAACTGAAATTAAGTTCCTAGACATGGAAGGAACATAATACACTTGCTCTAAAACTA
GAAAAACTTCTGGCCTCAAAACTAACTTAACAAAACCAACTGCTTCTATTTCTACTCTCGTGCCTTCTCCAACATGA
ATCCTGACTTCATTTTCCCTTGAACTCAACAGGCTTTGAACCCTCTTGTAAAGAATAGATAATGTGCCCCTGTGAAC
CTGTATCACACTCCCAGTAACTGTATAACATTTCTTTAATTCACCACCAACTTAGAAATTCCCTTTTATACCATTAC
CTTCCAATCCTTCAATCTAATATATTTTATAACTAACCCAAGATTGTTCCCATAACCTCCCTAAAAAAAATATAATA
TGAACTCTCTTTTTATCTCATCATTAGCAATGACTATCTCC

```
>11D02_ananassa_7_unspecific
NNGGGACGTTTTGCAGATACTGCTGTGTGACCAAACCCTAAATAAACCCCACCATCCAGCCTTGCCACCTCTATTGC
ATGCCTTTTGATCGTTTCCTCCTCACACTTGTTTTCTTTCTCTTCCTCCCCCAAAACAAGAACCCAAAGTCCCCCAA
AACGCATCATATATAGATACAGACGCAGCGAGTGTTATTATTAACCCGAAGAAAAACCCAAAGAGCTAATCGACAGA
GAAGAAGAAGAAAGAGGCTTTATATAAAGAGAGAAAACGCTTGCTGTTGATTCAGAGCAACCAGCCCTTCTCTTTTT
CCCTTCTTCTCTCTGTGTGTAATATTTCAATGGCCGTTGAAGCTCGGCACCTCAATCTATTCCCCTCTCAACAACTC
TTCTGCAACAACAGGTACTGATCCTCACTTCATCTGGAGTTGAAC
```

```
>11D02_ananassa_9_unspecific
GAGCTGCTGTGTGAACCAAAAAAGAGAAAGACAGAAAAAAGAGCAGGAGGATGGAGTTGCCAAAAAGGCTGATCTGG
TGCGGCAAGGTTGATTTGTGTTTTGGTCCTTACTGATTTGTGTCCATTGGATTGCTTATAAATGACGTGTCAGCTTG
TTAGTGATGTCCAAGCAGACAAGCCAAATCCTATGAAAAGAGAGTCAATTACATATAAGTCTATTTTGAGACATTTT
CTCCCATATAGTTTCATCCAAACATTTTTACCCATATAAGTCCACCTAAAGCTAAATAGAGTAGTATATTAGATATT
AAAGTTTAACATAATTACAGGTTGTCATCTAGCAAAAAAAAAAAAATATTTCTCTTATTTACGAATATACCACTAGA
GTAAAAGGTCAACAACCACCCTCCTCTCAAAAAAAGTCTACGGTCGACTTCCAGCGATGTTTCCAATTAACTTTCGG
TGAGGATTTTGGTTCACACAGCAGCTC
```

```
>17O22_vesca
AAAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCTTAATTTGACAACAGACA
TAGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACTAGAGACTGCATTTCTTATAAAGGCAATGGAATCG
TAGAGACTGCATTTCTTACTCAGTACTGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTT
CCCGAAATTCAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACTCTACATTCATTCGATCCCCTTATC
ACTGTAGGGCTTCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATAGTAGCCTAGTTTAGTTTCTTATGCT
GAAGCAAAATATGTAATCACCTACGCTACAGAATAGTGTTACTTGTTACCGGACATGTTCACAATCTTTGAAGATGA
AGAACGGTACCAGTTACCCAACATAATCATAGTTATTTTGGCCTATTGATATTTTGATTAACGTGTAATTGATCGCT
ACTTGAATGATGTATATTATGAATGGCACTATTTAATATTTTGGGCTGCTACCTACTCTTCAACAAACTCTAATTAA
TTAACCAAACATCAGTGTCACAAGTCACACCAACCTAGTTAAACTTTCCATTATAAGTAGCTTTCCCAATAACCTAC
CTCCCAAAAATAGTTACTTTAAAGCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATTCTGGAT
TTGTTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTCTGGTCATCAGTGTCAAACTAG
AAGAACCGTGAATTGCGACCCCTCAGAATGTCAAAATGAGATCACTGTGATTCCTTTTAAAATTTTAACAGCGATTC
TTCTACAAAAGATGGACTAAATTCCACCTTGTACTGTACAAAAAACGAGTTTGAGTAGTGGGAATCGTTCCAATATA
TTTCTGCTCTGTTTACCAATTGCCAGGATGATACAAACATCTAAACTCTACAGGAACCCTTTTCTAGCAAAAGAATG
AGAAGAAAGAACTCTACAAGAATCCAAAGCGCGAAAACAAAATCAGAACTAAGACTAGACATGAACAAATTTGCTGC
AGCCTCCACTGATGAGCTTCTCCAGCAAGAACAAAAGAATCAAACCAGATAAAATGGAAAATCTCCTCTCACGTTGG
AACAATATCATTGATTTCAGATTTTGTCTCAGATTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATG
GATTCAGAAAATTTGCTACAAAAGCCCATAACTTGTAAACATCATCGAAGTTTGTGAGGAAACCC
```

```
>17O22_viridis
TAAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCTTAATCTGACAACGGACA
TAGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACTAGAGACTGCATTTCTTATAAAGGCAATGGAATCG
TAGAGACTGCATTTCTTACTCAGTACTGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTT
CCCGAAATTCAACAGCCATCAGAGTTCATCTGCCAATCAAGGCAAATATGACTCTACATTCATTCGATCCCCTTATC
ACTGTAGGGCTCCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATTGTAGCCTAGTCTAGTTTCTTATGCT
GAAGCAAAATATGTAATCACCTAGGCTACAGAATAGTGTTACTTGTTACCGGACATGTTCACAATCTTTGAAGATAA
AGAACGGTACCAGTTACCCAACATAATCATACTTGTTTTGGCCTATTGATATTTTGATTAATATGTAATTGATCGCT
ACTGGAATGATGTATATTATATTATGAATGGCACTATTTAATATTTTGGGCTGCTACCTACTCTTCAACAAACTCTA
ATTAATTAACCAAACATCAGTGTCACAGGTCACACCAACCTAGTTAAACTTTCCATTATAAGTAGCTTTCCCAATAA
CCTACCTCCCAAAAATAGTTACTTTAAAAGCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATT
CTGGATTTGTTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTTGGTCATCAGTGTCA
AACTAGAAGAACTGTGAATTGCGACTACACCAGGATGCCTTTGGTCACTTACCAACCTCAAGAAAAGGACCCCCTCA
GAATGTCAAAATGAGATCACTGTGATTCCTTTTAAAATTTTAACAGTGATTCTTCTACAAAAGACTAAATTCCACTT
TGTACTGTACAAAAAACGAGTTTGAGTAGTGGGAATCGTTCCAATATATTTCTGCTCTGTTTACCAATTGCCAGGAT
GATACAAACATCTAAACTCTACAGGAACCCTTTTCTAGCAAAAAAATGAGAAGAAAGAACTCTACAAGAATCCAAAG
```

TGCGAAAACAAAATCAGAACTAAGACTAGACATGAACAAATTTGCTGCAGCCTCCACTGATGAGCTTCTCCAGCAAG
AACAAAAGAATCAAACCAGATAAAATGGAAAATCTCCTCTCACGTTGGAACAATATCATTGATTTCAGATTTTGTCT
CAGATTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATGGATTCAGAAAATTTGCTACAAAAGCCCAT
AACTTGTAAACATCATCGAAGTTTGTGAGGAAACCC

>17O22_iinumae
AAAATGGGTTGACGAGTTCGTGAACAACACTTTACGACCCAAAGCGTCCAATACTTCTTAATTTGACAACGGACATA
GTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACTAGAGACTGCTTTTCTTATAAAAGCAATGGAATCGTA
GAGACTGCATATCTTACTCAGCACTGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTATAACTTCC
CGAAATTCAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACTCTACATTCATTCGATCCCCTTATCAC
TGTAGGGCTCCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATTGTAGCCTAGTTTAGTTTCTTATGCTGA
AGCAAAATATGTAATCACCTAGGCTACAGAAGAGTGTTACTTATTACCGGACATGTTCACAATCTTTGAAGATGAAG
AACGGTACCAGTTACCCAACATAATCATAGTTATTTCGGAATATTGATATTTTGATTAATATGTAATTGATCGCTAC
TTGACTGATGTATATTATGAATGTCACTATTTAATATTTTGGGCTGCTACCTACTCTTCAACAAACTTTAATTAATT
AACCAAACATCAGTGTCACAAGTCACACCAACCTAGTTAAACTTTCCATTATAAGTAGCTTTCCCAATAACATACCT
CCCAAAAATAGTTACTTTAAAGCTGGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATTCTGGATTT
GTTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTCGGTCATCAGTGTCAAACAAGAA
GAACCGTGAATTGCGACTATACCAGGATGCCTTTGGCCACTTGCCAACCTCAAGAAAAGGACCCCTCAGAATGTCAA
GATGAGATCACTGTGATTCCTTTTAAAATTTTGACAGTGATTCTTCTACAAAAGATGGACCAAATTCCACCTTGTAC
TGTACAAAAAACGAGTTTGAGCAGTGGGAATCGTTCCAATATATTTCTGCTCTGTTTACCAATTGCCAGGATGATAC
AAACATCTAAACTCTACAGGAACCCTTTTCTAGCAAAAAAAATGAGAAGAAAGAACTCTACAAGAATCCAAAGCGCG
AAAACAAAATCAGAACTAAGACTAGACATGAACAAATTTGCTGCAGCCTCCACTGATGAGCTTCTCCAGCAAGAACA
AAAGAATCAAACCAGATAAAATGGAAAATCTCCTCTCACGTCGGAACAATATCATTGATTTCAGATTTTGTCTCAGA
TTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATGGATTCAGAAAATTTGCTACAAAAGCCCATAACT
TGTAAACATCATCGAAGTTTGTGAGGAAACCC

>17O22_nubicola
AAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCTTAATTTGACAACGGACAT
AGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACTAGAGACTGCATTTCTTATAAAGGCAATGGAATCGT
AGAGACTGCATTTCTTACTCAGTACTGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTTC
CCGAAATTCAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACTCTACATTCATTCGATCCCCTTATCA
CTGTAGGGCTCCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATTGTAGCCTAGTTTAGTTTCTTATGCTG
AAGCAAAATATGTAATCACCTAGGCTACAGAATAGTGTTTACTTGTTACCGGACATGTTCACAATCTTTGAAGATGA
AGAACGGTACCAGTTACCCAACATAATCATAGTTATTTTGGCCTAATTTTGATTAATATGTAATTGATCGCTACTTG
AATGATGTATATTATGAATGGCACTATTTAATATTTTGGGCTGCTACCTACTCTACAACAAACTCTAATTAATTAAC
CAAACATCAGTGTCACAAGTCACACCAACCTAATTAAACTTTCCATTACAAGTAGCTTTCCCAATAACCTACCTCCC
AAAAATAGTTACTTTAAAGCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTGCCTATAATTCTGGATTTGTT
ACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTTGGTCATCAGTGTCAAACTAGAAGAA
CCGTGAATTGCGACTATACCAGGATGCCTTTGGTCACTTACCAACCTCAAGAATGTCAAAATGAGATCTCTGTAATT
CCTTTTAAAATTTTAACAGTGATTCTTCTACAAAAGATGGACTAAATTCCACCTTGTACTGTACAAAAAACGAGTTT
GAGTAGTGGGAATCGTTCCAATATTATTTCTGCTCTGTTTACCAATTGCCAGGATGATTCAAACATCTAAACTCTAC
AGGAACCCTTTTCTAGCAAAAAAATGAGAAGAAGGAACTCTACAAGAATCCAAAGCGCGAAAACAAAATCAGAACTA
AGACTAGACATGAACAAATTTGCTGCAGCCTCCACTGATGAGCTTCTCCAGCAAGTACAAAAGAATCAAACCAGATA
AAATGGAAAATCTCCTCTCACGTTGGAACAATATCATTGATTTCAGATTTTGTCTCAGATTCTTCGTCAACAGTAGA
TAGTCCGCCTTCTCTGATGAAGGAAGGATTCAGAAAATTTGCTACAAAAGCCCATAACTTGTAAACATCATCGAAGT
TTGTGAGGAAACCC

>17O22_mandshurica
AAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCTTAATTTGACAACAGACAT
AGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACTAGAGACTGCATTTCTTATAAAGGCAATGGAATCGT
AGAGACTGCATTTCTTACTCAGTACTGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTTC
CCGAAATTCAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACCACATTCATTCGATCCCCTTATCACT
GTAGGGCTTCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATTGTAGCCTAGTTTAGTTTCTTATGCTGAA
GCAAAATATGTAATCACCTAGGCTACAGAATAGTGTTACTTGTTACCGGACATGTTCACAATCTTTGAAGATGAAGA
ACGGTACCAGTTACCCAACATAATCATAGTTATTTTGGCCTATTGATATTTTGATTAATGTGTAATTGATCGCTACT
TGAATGATGTATATTATGAATGGCACTATTTAATATTTTGTGCTGCTACCTACTCTTCAACAAACTCTAATTAATTA

```
ACCAAACATCAGTGTCACAAGTCACACCAACCTAGTTAAACTTTCCATTATAAGTAGCTTTCCCAATAACCTACCTC
CCAAAAATAGTTACTTTAAAGCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATTCTGGATTTG
TTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTTGGTCGTCAGTGTCAAACTAGAAG
AACCGTGAATTGCGACTATACCAGGATGCCTTTGGTCACTTACCAACCTCAAGAAAAGGACCCCTCAGAATGTCACA
ATGAGATCACTGTGATTCCTTTTAAAATTTTAACAGTGATTCTTCTACAAAAGATAAGACTAAATTCCACCTTGTAC
TGTACAAAAAACGAGTTTGAGTAGTGGGAATCGTTCCAATATATTTCTGCTCTGTTTACCAAGTGCCAGGATGATAC
AAACATCTAAACTCTACAGGAACCATCTTCTAGCAAAAAAATGAGAAGAAAGAACTCTACAAGAATCAAAGCGCGAA
AACAAAATCAGAACTAAGACTAGACATGAACAAATTTGCTGCAGCCTCCACTGAGGAGCATCTCCAGCAAGAACAAA
AGAATCAAACCAGATAAAATGGAAAATCTCCTCTCACGTTGGAACAATATCATTGATTTCAGATTTTGTCTCAGATT
CTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATGGATTCAGAAAATTTGCTACAAAAGCCCATAACTTG
TAAGGCATCATCGAAGTATTGTGAGGAAACCC
```

>27F10_vesca

```
CCTGCAGGGTTTTTCATCATGTAAGGACCTCCATTGTCAGTAGCTTTATGCATATCATCTTCATCACAACAGCTGAA
GCAGCTCATGATTCCTTTAAACACACACAAAAAAACCCACAGTCAAAATGAGGAAATGAACAATACCCAAGTCATGA
ACACACAAAATTCAGTAAAAAAGTAAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCAGACCTTTGGAGAC
AAATTTCGTTGCTTAATGTAATAAGCAACAAAAAAATTCAGCTCAGCTGGATCAAAGCCCAGATGAAAAAGATTAAAA
CTTTAAACAAGAAAATAAAGATCAGAGAAAGAAAATATGATGGGTAGATCGGGAGAGATAAAATTACCTGAATCTGA
AGTGGGGGAAGTGAGTCAGTGAAGGACTGAGTTGGTGGAGTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGCG
AAGGATGCGCGGCGCAGGATAGGAGGGAAAAGGGTGCGTAGGATAACCCAATCAATGAACCAGATGAGAATACGCTA
GTGATTTTGATTATGAATTCTATAAATTCTATAAAAATTTATTTCATTTCTTAATTCTTACTCTGTTTCGGTGTTGG
CCAGATTTGACTCTTCTGTGCTTCAGTTTTGACCATTTACTTTTATAACCTCAGGAAGGGTTCAAGCGCGGCCTGCC
ACGTGGTGAATTCAAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCCTCCCGGTAACAGTAA
CTTTATCGACAAAACGCTTCTTATTTTATTTTATTTTTTTTGGCGAGCAAAACGCTTCTTATTTGTTTTGGGTCTGT
ACGCTTTTGGGTCTTATTTGTCAAGTTTCAATCACTAGCAGGAAGACTTGCGTATAATTGAAATAGCCACTATCTAT
ACTCTATATGCAAACACAAGAGTAGAGAAGGAGAACCAGAATACATTTCCA
```

>27F10_viridis

```
TGCGGGTTTTTCATCATGTAAGGACCTCCaTTGTTCGGTAGCTTTATGCATATCATCTTCATCACAACAGCTGAAGC
AGCTCATGATTCCTTTAAACACACAAAAAAAAAACACCCATAATCAAAATGAGGAAATGAACAATACCCTAGTCATG
AACACACAAAATTCAGTAAAAAAGAAAAAAGGGATCCGCTTCAAGCCAATCCCATCAAACTTGCAGACCTTTGGAGA
CAAATTTCGTTGCTTAATGTAATAAGCAACAAAaAATTCAGCTCAGCTGGATCAAaGCCCAGATGAAAAAGATTAAA
ACTTTAAACAAGAAAATAaAGATCAGAGGAAGAAAATaTGATGGGNAGATCGGGAGAGATAAAATTACCTGAATCtG
AaGTGGGGGAAGTGAGTCCGTGAAGAAGTGaGTTGGTGGANTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGC
GCAAGGATGCCCGGCGCAGGATAGGAGGGAAAAGGGTGCGTAGGAtAACCCACTCCANGAACCANATGACAATGCNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTATG
CAAAAaTACACTCaTATTTATGTAGAAAACGAGAATTGAACCTCTAACCTCTTACAAACAACCTATGAAATGTATAa
TATATGTAAAGACGaTTAAAATATATGTATAATATATAATATATGTATTGTTTTATATTTATAACATACTATAGATA
TAAATACAATTAAAATATAAAATGTTCAAATTTTAGCAAGAGGTATGTTTCGAAACCATGACTGCTCTGATGGAAAA
TATGACCACTTACGATCAAAACAAAGCTATCATTGCATTATATTTGTGAAAAAAATTATATTTATCACTTCATTTTT
TGGGCCACAATCTAAGTTTAGTAGAGGCCTATTACCAACCGTACCAACTAAGTCGGTATACCAACATCGATGGTTGG
TTTTGATAGAGGATTTTGCCTACCAATCATAAGTTGGTTGGTACATGATATTGGTAAATAAAGTCGGTATATCTACC
AATGCCAGCCCTACTTGAAACTTAGCCGGAAGACTTCATATAATTGAAATAGCTGAGATACACACTTGCTATATGCA
AACACAAGAGTAGAGAAGGAGAACCAGAATACATTTCCA
```

>27F10_iinumae

```
CCTGCAGGTTTTTCATCATGTAAGGACCTCCATTGTCAGTAGCTTTATGCATATCATCTTCATCACAACAGCTGAAG
CAGCTCATGATTCCTTTAAACACACAAAAAAAAACCCACAATCAAAATGAGGAAATGAACAATACCCAAGTCATGAAC
GCACAAAATTCAGTAAAAAAGAAAAAAGGGATCCGCTTCAATACAATCCCATCAAACTTGCAGACCTTTGGAGACAA
ATTTCGTTGCTTAATGTAATAAGCAACAAAAATCCAGCTCAGCTGGATCAAAGCCCAGATGAAAAAGATTAAAACTT
```

TACCCAAGAAAATAAAGGTCAGAGGAAGAAAATATGATGGGTAGATCGGGAGAGATAAAATTACCAGAATCTGAAGT
GGGGGAAGTGAATCAGTGAAGGACTGAGTTGCTGGAGTCTTGGGAGATCTGAGCTCTAAAGCCGGCGAAGGATGCGC
GGCGCAGGATAGGAGGGAAAAGGGTGCGTAGGATAACCCAATCAATGAACCAAATGAGAATACGCTAGTGATTTTGA
TTATGAATTCTATAAATTCTACAAAAATTTATTTCATTTCTTAATTCTTACTCTGTTTCGGTGTTGGCCAGATTTGA
CTCTTCTGTGCTTCAGTCATGACTTTGACCATTTACTTTTATAACCCCAGGAAGGGTTCAAGCGCGGCCTGCCACGT
GGTGAATTCTGGTTCGTCCGGAAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCCTCCGGGT
AACAGTAACTTTATCGTTTTACCGCTAGCATGTCTCTGTCTGTCGACATATATAACGCTTCTTATTTGTTTTGGGTC
TCTACGCTTTTGGGTCTTATTTGTCAAGTTTCAATCACTTGAAACTATTGAAATAGCTGAAATACACACTTACTATA
TGCAAACACAAGGGAGAGAGGAGAACCAGATCATTCTA

>27F10_nubicola
CCTGCGGGTTTTTATCATGTAAGGACCTCCATTGTCAGTAGCTTTATGCATATCATCTTCATCACAACAGCTGAAGC
AGCTCATGATTCCTTTAAACACACACAAAAAAAACCCACAATCAAAATGAGAAAATGAACAATACCCAAGTCATGAA
CACACAAAATTCAGTAAAAAAGAAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCAGACCTTTGGAGACAA
ATTTCGTTGCTTAATGTAATAAGCAACAAAAAATTCAGCTCAGCTGGATCAAAGCCCAGATGAAAAAGATTAAAACT
TCAAACAAGAAAATAAAGATCAGAGGAAGAAAATATGATGGGTAGATCGGGAGAGATAAAATTACCTGAATCTGAAG
TGGGGGAAGTGAGTCAGTGAAGGACTGAGTTGGTGGAGTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGCGAA
GGATGCGCGGCGCAGGATAGGAGGGAAAAGGGTGCGTAGGATAACCCAATCAATGAACCAAATGAGAATACGCTAGT
GATTTTGATTATGAATTCTATAAATTCTACAAAAAATTTATTTCATTTCTTAATTCTTACTCTGTTTCGGTGTTGGC
CAGATTTGACTCTTCTGTGCTTCAGTTTTGACCATTTACATTTATAACCCCGGGAAGGGTTCAAGCGCGGCCTGCCA
CGTGGTGAATTCAAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCCTCCCGGTAACAGTAAC
TTTATCGTTTTACCGCTAGTATGTCTCTGTCTGTCGACATGACGCTTCTTATTTCTTTTGGGTCTCTACGCTTTTGG
GTCTTATTTGTCAAGTTTCAATCACTAGCAGGAAGACTTGCGTATAATTGAAATAGCCATTATCTATACTCTATATG
CAAACACAAGAGAGAGAAGGAGAACCAGAATCATCCA

>27F10_mandshurica
CCTGCAGGGCTTTTTATCATGTAAGGACCTCCATTGTCAGTAGCTTTATGCATATCATCTTCATCACAACAGCTGAA
GCAGCTCATGATTCCTTTAAACACACACAAAAAAACCCACAATCAAAATGAGGAAATGAACAATACCCAAGTCATGA
ACACACAAAATTCAGTAAAAAAGTAAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCAGACCTTTGGAGAC
AAATTTCGTTGCTTAATGTAATAAGCAACAAAAAAATCAGCTCGGCTGGATCAAAGCCCAGATGAAAAAGATTAAAA
CTTTAAACAAGAAAATAAAGATCAGAGGAAGAAAATATGATGGGTAGATCGGGAGAGATAAAATTACCTGAATCTGA
AGTGGGGGAAGTGAGTCAGTGAAGGACTGAGTTGGTGGAGTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGCG
AAGGATGCGCGGCGCAGGATAGGAGGGAACAGGGTGCGTAGGATAACCCAATCAATGAACCAAATGAGAATACGCTA
GTGATTTTGATTATGAATTCTATAAATTCTATAAAAATTTATTTCATTTCTTAATTCTTACTCTGTTTCGGTGTTGG
CCAGATTTGACTCTTCTGTGCTTCAGTTTTGACCATTTATTTTTATATCCTCAGGAAGGGTTCAAGCGCGGCCTGCC
ACGTGGTGAATTCAAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCCTCCCGGTACAGTAAC
TTTATCGTTTTACCGCTAGTATGTCTCTGTCAGTACTCTGTCGACATAACGCTTCTTATTTGTTTTGGGTCTCTACG
CTTTTGGGTCTTATTTGTCAAGTTTCAATCACTAGCAGGTAGACTTGCGTATAATTGAAATAGCCACTATCTATACT
CTATATGCAAACACAAGAGAGAGAAGGAGAACCAGAATACATTTCCA

>27F10_ananassa_2
CCTGCAGGTTTTTATCATGTAAGGACCTCCATTGTCAGTAGCTTTATGCATATCATCTTCATCACAACAGCGGAAGC
AGCTCATGGACTCCTTTAAACACACAAAAAAAACCCACGATCAAAATGAGGAAATGAACAATACCTAAGTCATGAAC
ACACAAAATTCAGTAAAAAAGAAAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCACCCCTTTGGAGACAA
ATTTCGTTGCTTAATGTAATAAGCAACAAAAATTCAGCTCAGCTGGATCAAAGCCCAGATGAAAAAGATTAAAACTT
TACCCAAGAAAATAAAGGTCAGAGGAAGAAAATATGGTGGGTAGATCGGGAGAGATAAAATTACCAGAATCTGAAGT
GGGGGAAGTGAATCAGTGAAGGACTGAGTTGCTGGAGTCGTGGAAGATCTGAGCCCGGCGAAGGATGCGCGGCGCAG
GATCGGAGGGAAAAGGGTGCGTAGGATAACCCAACCAATGAACCAAATGAGAACACGCTAGTGATTTTGATTATGAA
TTCTATAAATTCTACAAAAATTTATTTCATTTCTTAATTCTTACTCTGTTTCGGTGTTGGCCAGATTTGACACTTCT
GTGCCTCAGTCATGACTTTGGCCATTTACTTTTATAACCCCAGGAAGGGTTCAAGCGCGGCCTGCCACGTGGTGAAT
TCTGGTTCGTCCGGAAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCCTCCCGGCAACAGTA
ACTTTATCGTTTTACCGCTAGTATGTCTCTGTCTTATTTGTTTTGGGTCCCTACGCTTTTGGGTCTTATTTGTCAAG
TTTCAATCACTTGAAACATAGCAGGAAGACTTGCATATAATTGAAATAGCTGCAATACACACTTGCTATATGCAAAC
ACAACAAGAGAGAGGAGGAGAACCAGAATCATCCA

```
>29G10_vesca
TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGAGCTTTGGAGAGGTAGGAAGAGTTTATTTCTAGAGGGAGGCTACC
CATTTGAAGTAGAGATTGGACTAAAAACAACTTGAAAGGAAGATGGGGAGGATAAATAAAAAGGATAGAAACTGCTC
AAGTGCTTAACAATGGTTGTAGACGAGTTGTGTCTTGCTGCATATATTGAAGAGATTATATAGAGGTGCATGTAGGA
TGAAGACGCCGTATCTTAAATTTTGATTTGGTTCTTCTCACACACCAGAGATTGAGTTCGGATCATCGGATCCGAAA
AATCAAGTCCTTGTGTATAAAAGCACGTTACGGAGTGATCCCACTCATCAATAAGTTATCGGACTTAATTATTGTCA
CGGTGGACCACGTCAGTCTGGCATATCGATCATCACTCCCAATCTTGTCGATCATCAATTTGGCATGCATATCAGAC
CCAAGCCATTACTTGCTTCTATGAACGTATTTATATCATTTCTAATCACCCAGAATTATGGATAATATTTCTTATTC
ACAACGACGATTGGCTTCTTGGTGTGTTGCGCTTTGTTAGGACAGTTCATTGAATTTCAGGAATCCACAATTGGGTG
CTGCCTTCTTCT
```

```
>29G10_nubicola
TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGGGCTTTGGAGAGGTAGGAAGAGTTGATTTCTAGAGGGAGGCTACC
CATTTGAAGTAGAGATTGACTAAAAACAACTTGAAAGAGGAAGATGGGGAGGATAGAAACTGCTCAAGTGCTTAACA
ATGGTTGTAGACGAGTTGTGTCTTGCTGCATATATTGAAGAGATTATATAGAGGTGCATGTAGGATGAAGACACCGT
ATCTTAAATTTTGATTTGGTTCTTCTCACACACCAGAGATTGAGTTCGGATCATCGGATTCGAAAAATCAAGTCCTT
GTGTATAAAAGCACGTTACGGAGTGATCCCACTCATCAATAAGTTATCGGACTTAATTATTGTCACGGTGGACCATG
TCAGTCTGGCATATCGATCATCACTCCCAATCTTGTCGATCATCAATTTGGCATGCATATCAGACCGAAGCCATTAC
TTGCTTCTATGAATGTATTTATATCATTTCTAATCACCCAGAATTATGGATAATATTTCTTATTCACAACGACGATT
GGCTTCTTGGTGTGTTGCGCTTTGTTAGGACAGTTCATTGAATTTCAGGAATCCACAATTGGGTGCTGCCTCT
```

```
>29G10_nilgerrensis
TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGGGCTTTGGAGAGTTAGGAAGAGTTGATTTCTAGTGGGAGGCTACC
CATTTGAAGTAGAGATTGGACTAAAAACAACTTGAAAGGAAGATGGGGAGGATAAATAAAAAGGATAGAAACTGCTC
AAGTGCTTAACAATGGTTGTGGACGAGTTGTGTCTTGCTGCATATATTGAAGAGATTATATAGAGGTGCATGTAGGA
TGAAGACACCGTATCTTAAATTTTGATTTGGTTCTTCTCACACACCAGAGATTGAGTTCGGATAATCGGATTCGAAA
AATCAAGTCCTTGTGTGTAAAAGCACGTTGCGGAGTGATCCCACTCATCAATAAGTTATCGGACTTAATTATTGTCA
CGGTGGACCACGTCAGTCTGGCATATCTATCATCACTCTCAATCTTGTCGATCATCAATTTGGCTACATATCAGACC
GAAGCCATTACTTGCTTCTATGAATGTATTTATATCATTACTAATCACCCAGAATTATGGATAATATTTCTTATTCA
CAACGACGATTGGCTTCTTGGTGTGTTGCGCTTTGTTAGGACCAGTTCATTGAATTTCAGGAATCCACAATTGGGTG
CTGCCTTCTTCT
```

```
>29G10_mandshurica
TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGAGCTTTGGAGAGGTAGGAAGAGTTTATTTCTAGAGGGAGGCTACC
CATTTGAAGTAGAGATTGGACTAAAAACAACTTGAAAGGAAGATGGGGAGGATAAATAAAAAGGATAGAAACTGCTC
AAGTGCTTAACAATGGTTGTAGACGAGTTGTGTCTTGCTGCATATATTGAAGAGATCATATAGAGGTGCATGAAGGA
TGAAGACACCGTATCTTAAATTTTGATTTGGTTCTTCTCAGTTCTCACACACCAGAGATTGAGTTCGGATCATCGGA
TCCGAAAAATCAAGTCCTTGTGTATAAAAGCACGTTACGGAGTGATCCCACTCATCAATAAGTTATCGGACTTAATT
ATTGTCACGGTGGACCACGTCAGTCTGGCATATCGATCATCACTCCCAATCTTGTCGATCATCAATTTGGCATGCAT
ATCAGACCGAAGCCATTACTTGCTTCTATGAACGTATTTATATCATTTCTAATCACCCAGAATTATGGATAATATTT
CTTATTCACAACGACGATTGGCTTCTTGGTGTGTTGCGCTTTGTTAGGACAGTTCATTGAATTTCAGGAATCCACAA
TTGGGTGCTGCCTTCTTCT
```

```
>32L07_vesca
GAGTTGAAAAACGGGTCGAATCCCGGCACCACCGTCCGCGTCGCGTAGGACTTGAATCCTTCCAAGGTCACCTCCTT
GATGTACATAGCTGCCCTCGCCGGAGAGGTGCGGACGCTAATCGGAAGCCGATTTTGGAGAGATTTAGTGTCGGTGA
TAGATCGGAACCCTAGAAATCTGAGCTTCTGGTTTTTGCTTTCGGAAGTTGAGAGTCTGAAATGACATGGTTCGAAT
TTCTTTTTGTTGTTTTCCGCTTTTTTGGTGGGTTCGAATTTTTAGACCAAGGCGGGAGATATTTGGGCCAGTGATTT
ATATCTTGGGCTCACTCTGGGACTCATGTCTTTGGGCCTCGTCGACCTCGAGGTGCTCATGAAGTCCGGCCGTCCTC
AGGGTCGGAAACACCGCGGTACTACTGACTACTGTGTCATCGCTTTAGAATTTCATTAATTGGCTTTGCGAGCTATA
AATAATTGTGATTTGGTTTGAATTTAGGTAAGTTTTAGTATTAGTATTTATCAACGGGAATTGCGGAGATGAGAAAA
GTTGAGGTTGATTTGGGGGAGTGTGGTGTTGTTAGTTAGTTGAATTATTAGAAACGAAAAAATAACAGAAGAATATA
```

```
AATGTGGATGGATTATTGGATTAAGATTTGATTCAACGGAAGAAGGAGGCGTGGTGTGTGTTTTGATAGTCTAATTT
GAACTGTTTTGCTTCTGACAGCTAAAATCTATCCGGTGGTGAAAAATCAGCATCGGCTACTATGTACACTTTTAATC
GGCAACGCATTAGCGATGGAGGTGACTTGTCTAATTTACTAAGTTTATTTAGGTTGTTACTGGTACATTTTATGTGT
TTATTGCCGTGGATGTAGTTTGTATGGGCCAGTTGACCAGCAGTTTCAAATGGCAGGCCAATAGGGCCAACCTAGAT
TGTAGTTGAATTTTGGGAAGGAAAAAAAAAAGCAAACCAAAAGACATCACCACGAGCCACTTTGGCCTATCTATATA
TATTACTTCCTTGCTTAATGTGTTGCTCAATTGCTAAACAATATCATCAATGTCTAAAATAACGCGCCTCAAGGCTA
AGGCAAGGGAAGGCGTGCCTTAGGACGACCTCTGAAAGACATTTGATATCAAAGGTGTGATTGAGGCGCGCGATCAA
GACGACGAGGTCAAGGTGCCTAATACAACATCAGGTTATAGGTTTGAATCTCACTTTGAGAAATGTGATGGTTTGAA
CGGTTAAATCTATTGTCTTTTTATATTGTATGGGCGGTAAAATTAAATGTTAAACTTCGGTAAATTGTCAAATGTTT
AATAGTATAAGAATCTACATATAGTAGGTGTAAAATAGATACCGAAATGATAATATTTTGTGAATAACGTACGTCAT
ATGATTTAATATTAAGACTTTGTACGATTTAACGTTACACATTAAAATTGTAGATAAAAAGTTTATATCATCATCAA
CATCGATGTTCGAATAAATTTTATAACGTTCAATGCGGTACAAATCTCCCAATGACTATTATCGAGTACAACGTCCA
TATCCGACACATGATATAGGCTATCAAATTATCAAAACCCTTTGATCCGATTCTGTAGCTTTGACGACTATAAGCTT
AGTTAAGTTTAGTAGGACTCACCGCAATTTCGCACTAGTAGGACAAAAAGATGGTAAGATTCCTTTCATTTTTCTTC
TTTACTATCCTTCTTTTCCTCAATTTTTCCCTAGAATCCTACAACAAGAAAGGACTTTGGCCCCTTGTGCTCCTTTA
TCATCTTAAAAGCATCACCACCATCCCCTATATAGATGCATATTCACTATCAAGCTACCCAAGTATGCAAATTTATA
GCATCTCATTATCTTGTTTCCTCTAGCTATTCTACTCAATGCATATCAACAACCTGACCCAGTTCTCCTATAATTGC
TGGCAGATAGTAATACCAATTACTCCAGAATCTTCACACCCAGAACTTGAAATTACACGACCTCAATACTCCAAACA
GTACAAAACAACCCAGATGATCAAAACACATAACATTCTTTATTTCATCTTATTGGGAAAATCTCTATATCTATTAT
CTTCATTATTCAATTTTTCTACACTGCATGCTATACATGTTACAAAAGAGAAAGAAAAGACACTAGTCCATATCACA
TAGGCCATGTCCTTCCCAATTCTAACCCAACAATTCAAGGACCACACCCATGAGTAGTGGCACTGAATCACTGAATC
GTCGCCTTCACAACTACACTACCTATCCAACCCAGACTCAACACAGATGAAAATTCACAGCAGCTAAGAATATAGTA
CTAGTTTTGCTCTATCTTTTTTCTTTACCAAAACAAAAAAAACCCTGTAGTAACCAATATAACCGCTAACAGCTTTT
CCCATCCTGCCCATAACAGCTTTTCCCCTGCAGTATGGGAAACCCTTATCTAAAACCCCCCGATTTATAGTAACAAA
AAAATAAATAAAATAATTTACTTTCCTCATTTACCATTTTACCCTCATCTTCTCCTTCATTGCCACTTGAACCCCCA
CTCTCCATGCTCCTTGAACCTTCTCAACACCCTTTCTAGGGCAATGTCAAAAGCGTCTTTTACCGTCTCCAACCCCT
CCTGCGGTTTCGCGTACAGAAAATTCGGTATGTAATCGATAACTTTCTCCCGCATTTTCCT
```

>32L07_viridis

```
GAGTTGAAAAACGGGTCGAATCCCGGCACCACCGTCCGCGTCGCGTAGGACTTGAATCCTTCCAAGGTCACCTCCTT
GATGTACATAGCTGCGCTCGCCGGAGAGGTGTGGACGCTAATCGGTAGCCGATTTTGAAGAGATTTAGGGTCGGTGA
TAGATCGGAACCCTAGAAAATACGTCACCACGAGCCACTTTGGCCTATCTATGTTACTTCCTTGCTTAATGTGTTGC
TCAATTGCTCAACAATATCATCAATGTCTAAAATAACGCGCCTCGAGGCTAAGGCAAGGGAAGGCGTCACCTTAGGA
CGACCTCTGAAAGACATTTGATATCAAAGGTGTGATTGAGGCGCGCGATCAAGACGACGAGGTCAAGGTGCCTAATA
CAACATCAGGTTATAGATTTGAATCTCACTTTGAGAAATGTGATGGTTTGAACGGTTAAATCTATTGTCTCTTTATA
TTGTATGGGTGGTAAAATTAAATGTTAAATTTCGGTAAATTGTCAAATGTTTAATAGTATAAGAATCTACATATAGT
AGGTGTAAAATAGATACCGAAATGATAATATGTTGTGAATAATATACGTCATATGGTTTAATATTAAGACTTTGTGC
GATTTAATGCTACACATTAAAATTGTAGATAAAAAATTTATATCATTATCATCATCGATGTTCGAATAAATTTTATA
ACGTTCAATGCGGTACAAATCTCCCAATGACTATAATCGAGTACAACGTCCATATCGGACACATGATATAGGCTATC
AAATTATCAAAACCCTTTGATCCGATTCTGTAGCTTTGACGACTATAAGCTTAGTTAAGTTTAGTAGGACTCACCGC
AATTTCGCACTAGTAGGACAGAAAGATGGTAAGATTCCTTTTATTTTTCTTCTTTACTATCCTTCTTTTCCTCATTT
TTTCCCTAGAATCCTACAACAAGAAAGGACTTTGGCCCCTTGTGCTCCTTTATCATCTTAAAAGCATCACCACCATC
CCCTATTTAGATGCATATTCACTATCAAGCTACCCAAGTATGCAAATTAATAGCATCTCATTATCTTGTTTCCTCTA
GCTATTCTACTCAATGCATATCAACAACCTGACCCAGTTCTCCTATAATTGCTGGCAGATAGTAATACCAATTACTC
CAGAATCTTCACACCCAGAACTTGAAATTACACGACCTCAATACTCCAAACAGTACTGTCAGTACAAAACAACCCAG
ATGATCAAAACACTTAAAATTCTTTATTTCATCTTATTGCTATCTCTATCATCTTCATTATTCAATTTTTCTACACT
GCATGCTATACATGTTACAAAAGAGAAACAAAAGACACTAGCCCATATCACATAGGCCATGTCCTTCCCAATTCTAA
CCCAACAATTCAAGGACCACACCCATGAGTGGCACTGAATCACTGAATCGTCACCTTCACAACCACACTACCTATCC
AACCCAGACACAGATGAAAATTCACAGCAGCTAAGAATATAGTACCAATTTTGCTCTATCTTTCTTCTTTACCAAAA
CAAAAAAGATCCTGTAGTAACTAATATAACAGCTAACAGCTTTTCCCATCCTGCCCATAACAGCTTTTCCCCTGCAG
TATGGGAAACCCTGATCTAAATCCCCCGATTTATAGTAACAAAAAAATAAATAAAATAATTTGCTTTCCTCATTTAC
CATTTTACCCTCATCTTCTCCTTCATTGCCACTTGAACCCCCACTCTCCATGCTCCTTGAACCTTCTCAACACCCTT
TCTAGGGCAATGTCAAAAGCGTCTTTTACCGTCTCCAACCCCTCCTGCGGTTTCGCGTACAGAAAATTCGGTATGTA
ATCGATAACTTTCTCCCGCATTTTCCT
```

```
>34D20_vesca
GCAGAAAGAAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTCAGGGGATAAAGGAAGTA
TTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCTCTCCAAGAAATGGAGCAAGTCCTACTTCCTACGC
GAATTTGATTTCTACAAGGTGAGCAACATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGT
TATGTGAATCATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGATTATCGGC
GAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCTTGCAAACACTTGCTATGGTGTAATG
ACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTTTTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAA
ATGGATTTAGTACTTCCATTAAGTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTCCATGCACAGGAGAACTGC
AGGTGAACCGTATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGTGTGGATGAACTATACG
TAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGTTGGGTATGAATCTCCCTCCCGGCCACACTAGAC
CACACTTTTGAACTGGCGGATTCCATCCGTCCTAGATTTTGTGCCGACTATCACAATAGTGTAATTAAGTTGGTCCT
CCTAGCCATAGTTTCTAGTACTATTCTACTGATATCATGTATTGCCTCAGCTTTTGACAATGGAATATGATGAATTT
GGAATGAATACAAAAACTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAAAAGATGGGTACATGTTTGCTTATT
CTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTCATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCAGT
ACTTGTATAGTTGTATTGATCTGATATAACATAAATTTAATGAATCTAATAGACATTTTTCCTAGTTAACAGAGGAT
AGGTCTCCGGCTGACCTTATCCTACAAGGAAATAGAAACGTACAATTAACGCATTATACACAAGACTGGTCTATATA
AGGCATCAAATTCTCTTTATCTGTTTCATTGATCATATTGTCCTCTTTATCTGTTTCATACTTTCATTGATCATATT
GTCTAGTACTGGAAGAGCTATATTTATCAGATAACAGAAAGTGCTTACTTGCTGGTTCATACTCAATATGGATCCGA
AGGTCCTTAGTTACAATGGTGTTGACCTGAGCATGAGCGACTTGGATCTTCTTAGAGGCCCTTGTTACTTAACCGAT
AGCATCATTTGATTCTATTTCACTCATCTTACTTCCCATTATGATGATGATATCCTTCTGGTTTCCCCTAATATCTC
TGATCTTCTGGTAAATTCTCCGGATCCCGAGGATGAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGA
AAGTTGTGATCTTCGCAGTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG
TATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTGGAAGCTAGGAAAATGTA
TACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCACCAATAACTCCAGCTGGGACTAGTAGTTTGGCTA
CCAACAACAGTTCTACAATGAGACACGAGTGCCACTCTACGCAGTCGCGGCGATTTATAGACTATACCAAGACAATG
CTTGGGGTTTGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCATTATCATCC
TTCGGAAGTGTGCTATCCCATGCAAAAAATCACCAATAGTAATCATGGAGATGATGATGATGTTAATGAACCTTGGT
ATAGAGAAGAGACTCTTATGCCTCAGCAGACGAATTTTTACGACTGCG

>34D20_iinumae
NNNNNNNNNAACTGATGTGCTTTCCGGAGGGACTGACAGTAGAAAAGGACAGTGCAGTTCAGGGGATAAAGGAAGTA
TTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCTCTCCAAGAAATGGAGCAAGTCCTACTTCCTACGC
GAATTTGATTTCTACAAGGTGAGCAACATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGT
TATGTGAATCATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGATTATCGGC
GAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCTTGCAAACACTTGCTATGGTGTAATG
ACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTTTTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAA
ATGGATTTAGTACTTCCATTAACTAGTAACTGAGGAATCCAATTGCACTCTGTGTTTACATGCACAGGAGAACTGCA
GGTGAACTATATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGTGTGGATGAACTATAAGT
AGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGTTGGGCATGAGTCTCCCTCCCGGCCACACTAGACC
ACCTAGATTTTGTGCCGACTATCACAATAGTGAAAGTTGGTCCTCCTAGCTATAGTTTCTAGTACTATTCTACTGAT
ATCATGTTTCGTCTCAGCTTTTGACAATGGAATATGATGAATATGGAATGAACAAAACCTGCTTTGTCCATCTATTA
GCATTTTCTGAAACCCAAAAGATGGGTACATGTTTGCTTATTCTCTTTATCTAGTGCATCATGTGAGTTATCAAGTT
CATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCACTACTTGTGTAGTTGTATTGATCTAAATTTTTCCTAG
TTAACAGAGGATAGGTCTCCGGCTGACGTTATCCTACAAGGAAACAGAAACGTACAATTAACGGATTCACAAGACTG
GTCTATATAAGGCATCAAATTCTCTTTATCTGTTTCATTGATCATATTGTCTAGTACTGGAAGAGCTATATTTATCT
GATAACAGAAAGTGCTTACTTGCTGGTTCCTACTCATTATGGATCCGAAGGTCCTTAGTTACAAAGGTGTTGACCTG
AGCATGAGCGACTAGGATATTCTTAGAGGACCTTATTACTTAACCGATAGCATCATTCGATTCTATTTCACTTATCT
TACTTCCCATTATGATGATGATATCCTTCTGGTTTCCCCTAATATCTCTGATCTCCTGGTAAATTCTCCGGATCCCG
AGGATGAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATCTTCGCAGTGAATGATAAC
AAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTGTATTTCAGAAAATCAAACGCATTCGTACA
TTACGACAGCTTGGGGGGTAACAATAGTTTGGAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTC
CAGCAACACAAGCACCAATAACTCCAGCTGGGACTAGTAGTTTGGCTACCAACAACAGTTCTACAATGAGACACGAG
TGCCACTCTACGCAGTCGCGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTTTGGGGTTTTGTTGTCAACTA
CATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCATTATCATCCTTCGGAAGTGTGCTATCCCATGCAAAAAA
TCACCAATAGTAATCATGGAGATGATGATGATGTTAATGAACCTTGGTATAGAGAAGAGACTCTTTGCCTCAGCAGA
CGAATTTACGACTNNN
```

>34D20_viridis

```
NNNGAGAAGAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTCAGGAGATAAAGGAAGTA
TTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCTCTCCAAGAAATGGAGCAAGTCCTACTTCCTACGC
GAATTTGATTTCTACAAGGTGAGCAACATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGT
TATGTGAATCATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGATTATCGGC
GAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCTTGCAAACACTTGCTATGGTGTAATG
ACTGCAAGTTACGGGCAATAGGGGCTTCGGGTCTTTTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAA
ATGGATTTAGTACTTCCATTAACTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTTCATGCACAGGAGAACTGC
AGGTGAACCATATGTCTATATAGATATGTCGAATGTTAGATAGGATACATAGTATGAGGGTGTGGATGAACTATACG
TAGAGCACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGTTGGGCATGAATCTCCCTCCCGGCCACAGTAGAC
CACACTTTTGAACTGGCGGATTCCATCCGGCCTAGATTTTGTGCCGACTATCACAATAGTGTAAGTTGGTCCTCCTA
GCTATAGTTTCTAGTACTATTCTACTGATATCATGTTTTGTCTCAGCTTTTGACAATGGAATATGATGAATATGGAA
TGAACAAAAGCTGCTTTGTCCATCTTTTAGCATTTTCTGAAACCCAAAAGATGGGTACATGTTTGCTTATTCTCTTT
ATCTAGTGCATCATGTGAGTTATCAAGTTCATGTTTATGTATTCTGCTGATTTAGGAATTAGGATTGCAGTACTTGT
ATAGTTGTATTGATCTGATATAACTCCCATTAATGAATCTAATATAAATTTTTTCCTAGTTAACAGAGGATAGGTCT
CCGGCTGACCTTATCCAACAAGGAAACAGAAACATACAATTAACGCATTATACACAAGACTGGTCTATATAAGGCAT
CAAATTCTCTTTATCTGTTTCATTGATCATATTGTCTAGTACTGGAAGAGCTATATTCATCTGATAACAGAAAGTGC
TTACTTGCTGGTTCATACTCAATATGGATCCGAAGGTCCTTAGTTACAATGGTGTTGACCTGAGCATGAGCGACTTG
GATCTTCTTAGAGGCCCTTATTACTTAACCGATAGCATCATTCGATTCTATTTCACTAATCTTAGTTCCCATTATGA
TGATGATATCCTTCTGGTTTCCCCTAATATCTCTGATCTTCTGGTAAATTCTCCGGATCCCGAGGATGAGCTTAGAG
CCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATATTTGCAGTGAATGATAACGAAGATCCGAGTCGA
AGCGACGGCGGAAACCATTGGAGCTTGCTGGTGTATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGG
GGGTAACAATAGTTTGGAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCAC
CAATAACTCCAGCTGGGACTAGTAGTTTGGTTACCAACAACAGTTCTACAATGAGACACGAGTGCCACTCTACGCAG
TCGCGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTTTGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTA
CTTGCGTCTGTTTGGAAATTATCATTATCATCCTTCGGAAGTGTGTTATCCCATGCAAAAAATCACGAATGGTAATC
ATGGAGGTGATGATGATGTTAATGAACCTTGGTATAGAGAAGAGACTCTTATGCCTCAGCAGACGAATTTACGACGC
G
```

>34D20_mandshurica

```
NNNNNGAAGAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGAGCAGTTCAGGGGATAAAGGAAGTA
TTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCTCTCCAAGAAATGGAGCAAGTCCTACTTCCTACGC
GAATTTGATTTCTACAAGGTGAGCAACATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGT
TATGTGAATCATCTGCAAATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGATTATCGGC
GAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCTTGCAAACACTTGCTATGGTGTAATG
ACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTTTTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAA
ATGGATTTAGTACTTCCATTAAGTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTCCATGCACAGGAGAACTGC
AGGTGAACCGTATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGTGTGGATGAACTATACG
TAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGTTGGGTATGAATCTCCCTCCCGGCCACACTAGAC
CACACTTTTGAACTGGCGGATTCCATCCGTCCTAGATTTTGTGCCGACTATCACAATAGTGTAATTAAGTTGGTCCT
CCTAGCCATAGTTTCTAGTACTATTCTACTGATATCATGTATTGCCTCAGCTTTTGACAATGGAATATGATGAATTT
GGAATGAATACAAAGACTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAAAAGATGGGTACATGTTTGCTTATT
CTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTCATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCAGT
ACTTGTATAGTTGTATTGATCTGATATAACATAAATTTAATGAATCTAATAGACATTTTTCCTAGTTAACAGAGGAT
AGGCCTCCGGCTGACCTTATCCTACAAGGAAATAGAAACGTACAATTAACGCATTATACACAAGACTGGTCTATATA
AGGCATCAAATTCTCTTTATCTGTTTCATTGATCATATTGTCCTCTTTATCTGTTTCATACTTTCATTGATCATATT
GTCTAGTACTGGAAGAGCTATATTTATCAGATAACAGAAAGTGCTTACTTGCTGGTTCATACTCAATATGGATCCGA
AGGTCTTTAGTTACAATGGTGTTGACTTGAGCATGAGCGACTTGGATCTTCTTAGAGTCCCTTGTTACTTAACCGAT
AGCATCATTCGATTCTATTTCACTCATCTTACTTCCCATTATGATGATGATATCCTTCTGGTTTCCCCTAATATCTC
TGATCTTCTGGTAAATTCTCCGGATCCCGAGGATGAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGA
AAGTTGTGATCTTCGCAGTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG
TATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTGGAAGCTAGGAAAATGTA
TACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCACCAATAACTCCAGCTGGGACTAGTAGTTTGGCTA
CCAACAACAGTTCTACAATGAGACACGAGTGCCACTCTACGCAGTCGGGGCGATTTATAGACTATACCAAGACAATG
CTTGGGGTTTGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCATTATCATCC
```

TTCGGAAGTGTGCTATCCCATGCAAAAAATCACCAATAGTAATCATGGAGATGATGATGATGTTAATGAACCTTGGT
ATAGAGAAGAGACTCTTATGCCTCAGCAGACGAATTTTACGAGTGCG

>34D20_nubicola
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNATGTTGGCATCCAGACGGCATCTGGTTTTGGAGTCCCTCTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGA
ATTTGATTTCTACAAGGTTAGCAACAGCCTGCAAACTAGATATATTTTGTTTTTCTAACTATTACAGTGTGTATTAT
GTGAATCATCTGCATATTATCTATATCTAACTCTATGGTATAATGATTCAGAACTACAAGGAAAGATTATCGGCGAG
AAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCTTGCAAACACTTGCTATGGTGTAATGACT
GCAAGTTGCGGGCAATAGGGGCTTCAGGCCTTTTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATG
GATTTAGTACTTCCATTAACTAGTAACTGAGGAATCCAATTGCAATCTGTGTTTTCCATGCACAGGAGAACTGCAGG
TGAACTTTATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGTGTGGATGAACTATACGTAG
AACACCCAGAAAACCAGAGAAAGTAAAGAGGAACTGCAGGTTGGGCATGAATCTCCCTCCCGGCCACACTAGACCAC
ACTTTTGAACTGGCGGATTCCATCCGGCGTAGATTTTGTGCCGACTATCACAATAGTGTAAGTTGGTCCTCCTAGCT
ATAGTTTCTAGTACTATTCTACTGATATCATGTTTTGTCTCAGCTTTTGACAATGGAATATGATGAATATGGAATGA
ACAAAAGCTGCTTTGTCCATCTGTTAGCATTTTCTGAAACCCAAAAGATGGGTACATGTTTGCTTATTCTCTTTATC
TAGTGCATCATGTGAGTTATCAAGTTCATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCAGTACTTGTATA
GTTATTGATCTAATATAACATAAATTTAATGAATCTAATATAAATTTTTCCTAGTTAAGTCAAAAGACCTTATCCTA
CAAGGAAACAGAAACGTACAATTAACGCATTATACACAAGACCGGTCTATATAAGACATCAAAATCTCTTTATCTGT
TTCATATTGATCATATTGTCTAGTACTGGAAGAGCTATATTTATCTGATAACAGAAAGTGCTTACTTGCTGGTTCAT
ACTCAATATGGATCCGAAGGTCCTTAGTTACAAAGGTGTAGACCTGAGCATGAGCGACTTGGATCTTCTTAGAGGCC
CTTGTTACTTAACCGATAGCATCATTCGATTCTATTTCACTCATCTTACTTCCCATTATGATGATGATATCCTTCTG
GTTTCCCCTAATATCTCTGATCTTCTGGTAAATTCTCCGGATCCCGAGGATGAGCTTAGAGCCTTTGCGGAGTCTGA
CCAACTTGGTAAAAGAAAAGTTGTGATCTTTGCAGTGAATGATAACGAAGATCCGAGTCGAAGCGACGGTGGAAACC
ATTGGAGCTTGCTGGTGTATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTG
GAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCACCAATAACTCCAGCTGG
GACTAGTAGTTTGGTTACCGACAACAGTTCTACAATGAGACACGAGTGCCACTCTACGCAGTCGCGGCGATTTATAG
ACTATACCAAGACAATGCTTGGGGTTTGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGA
AATTATCATTATCATCCTTCGGAAGTGTGTTATCCCATGCAAAAAATCACCAATAGTAATCATGGAGATGATGATGA
TGTTAATGAACCTTGGTATAGAGAAGGAGACTCTTATGCCTCAGCAGACGAATTACGACGCA

>34D20_nilgerrensis
NGCAGAAAGAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTCAGGGGATAAAGGAAGTA
TTATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCTCTCTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCG
AATTTGATTTCTACAAGGTGAGCAACATGCCGGCAAACTAGTTATATTTTGTTTTTCTTACTATTACAGTGTGTGTT
ATGTGAATCATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGATTATCGGCG
AGAAGGTGTTTTGCATGCAAGCAGCAGCAAATGCTATGGGCCAATTTCCCTTGCAAACACTTGCTATGGTGTAATGA
CTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTTTTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAA
TGGATTTAGTACTTCCATTAACTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTCCATGCATAGGAGAACTGCA
GGTGAACCATATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGTGTGGATGAACTATACGT
AGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGTTGGGTATGAGTCTCCCTCCCGGCCACACTAGACC
ACACTTTTGAACTGGCGGATTCCATCCGGCCTAGATTTTGTGCCGACTATCACAATAGTGTAAGTTGGTCCTCCTAG
CTATAGTTTCTAGTACTATTCTACTGATATCATGTTTTGTCTCAGCTTTTGACAATGGAATATGATGAATATGGAAC
AAAGCTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAAAAGATGGGTACATGTTTGCTTATTCTCTTTATCTAG
TGCATCATGTGAGTTATCAAGTTCATGTCTATGCATTCTGCTGATTTAGGAATAAGGATTGCAGTACTTGTATAGTT
GTATTGATCTGATATAACATAAATTTAATGAATCTAATATAAATTTTTCCTAGTTAAGCCAAAAGACCTTATCCTAC
AAGGAAACAGAAACGTACAATTAACGCATTATACACAAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGTT
TCATTGATCATATTGTCTAGTACTGGAAGAGTTATATTTATCAGATAACAGAAAGTGCTTACTTGCTGGTTCATACT
CAATATGGATCCGAAGGTCCTTAGTTACAATGGTGTTGACCTGAGCATGAGCGACTTGGATCTTCTTAGAGGCCCTT
GTTACTTAACCGATAGCATCATTCACTTATCTTACTTCCCATTACGATGATGATATCCTTCTGGTTTCCCCTAATAT
CTCTGATCTTCTGGTAAATTCTCTGGATCCTGAGGATGAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAA
GGAAAGTTGTGATCTTCGCAGTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTG
GTGTATTTCAGAAAATCAAACGCATTCGTACACTACGACAGCTTGGGGGGTAACAATAGTTTGGATGCTAGGAAAAT
GTATACAGTATTCAAGAAACTTGTGGCAGCTCCAGCAACACAAGCACCAACTCCATCTGGGACTAGTAGTTTGGTTA
CCAACAACAGTTCTACAATGCGACACGAGTGCTACTCTACGCGGTCGCGGCGATTTATAGACTATACCAAGACAATG
CTTGGGGTTTGGGGTTTTGTTGTCAATTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCATTATCCTTC

GAAAGTGTGTTATCCCATGCAAAAGATCACCAATAGTAGTAATCATGGAGATGATGAAGAGCGTAATGATGATGATG
TTAATGAACCTTGGTATAGAGAAGAGACCGTTATGCCTCAGCAGACGAATTTACGACTNNN

>34D20_ananassa
NNNNNNNNNAACTGATGCGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTCATGGGATAAAGGAAGTA
TTAATGTTAGGCATCCTAGACGGCATCTGGTTTTGGAGTCCCTCTCCAAGAAATGGAGCAAGTCCTACTTCCTACGC
GAATTTGATTTCTACAAGGTGATCAACATGCCTGCACACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGT
TATGTGAATCATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACCAGGAAAGATTATCGGC
AAGAAGGTGTTTTGCATGCATGCAGCACAAAATGCTATGGGCCAATTTCCCTTGCAAACACTTGCTATGGTGTAATG
ACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTTTTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAA
ATGGATTTAGTACTTCCATTAACTAGTAACTGAGGAATACAATTGCACTCTGTGTTTTCCATGCACAGGAGAACTGC
AGGTGAACTATATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGTGTGTATGAACTATAAG
TAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGTTGGGCATGAGTCTCCCTCCCGGCCACACTAGAC
CAC*ACTTTTGAACTGG*CGGATTCCATCCGGCCTAGATCTTGTGCCGACTATCACAATAGTGTAAGTTGGTCCTCCTA
GCTATAGTTTCTAGTACTATTCTACTGATATCATGTTTCGTCTCAGCTTTTGACAATGGAATATGATGGATATGGAA
TGAACAAAACCTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAAAAGATGGGTACATGTTTGCTTATTTTCTTT
ATCTAGTGCATCATGTGAGTTATCAAGTTCATGTTTATGCATTCTGCTGATTTAGGATTTAGGATTGCACTACTTGT
ATAGTTGTATTGATCTAAATTTTTCCTAGTTAACAGAGGATAGGACTCCGGCTGACCTTATCCTACAAGGAAACAGA
AACGTACAATTAACGGATTCACAAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGTTTCATTGATCATATT
GTCCAGTACTGGAAGAGCTATATTTATCTGATAACAGAAAGTGCGTACTTGCTGGTTCATACTCAATATGGATCCGA
AGGTCCTTAGTTCCCAAGGTGTTGACCTGAGAATGAGCGACTTGGATCTTCTTAGAGGCCCTTATTACTTAACCGAT
AGCATCATTCAATTCTATTTCACTTATCTTACTTCCCATTATGATGATGATATCCTTCTGGTTTCCCCTAATATCTC
TGATTTTCTGGTAAATTCTCCGGATCCCGAGGATGAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGA
AAGTTGTGATCTTCGCAGTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG
TATTTCAGAAAATCAAACGCATTGGTACATTACGACAGCTTGGGGGGTAACAATAGTTTGGATGCTAGGAAAATGTA
TACAGCATTCAAGAAACTTGTGGCAGTTCCAGCAACACAAGCACCAACTCCAGCTGGGAGTAGTAGTTTGGTTACCA
ACAACAGTTCTACAATGGGACACGAGTGCTACTCTACGCAGTCGCGGCGGTTTATAGACCATACCAAGATAATGCTT
CGGGTTGGGGGTTTTGTTGTCAAATACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCATCATCATCCTTC
GGAAGTGTGTTATCCCATGCAAAAGATCACCAATAGTAATCATGGAGATGATGAAGAGCTTAATGATGATGATGTTA
ATGAACCTTGGTATAGAGAAGAGACTCTTATGCCTCGGCAGACGAATTTTACGACTGCG

>40M11_vesca
CAACATTTTGGTGGCCTTCTTGACATTCCAGTTTCTGGCCCTCAGATGCCTTGCAATGGATGCATCAGAACAGTATG
TGGACAGCTTCTCGGGTACTGCCTTTAACAATTTTCTCACCTCATTAATCTGCAAACAATAAGATTTTTTAGGCAAA
GCAGAACTATGAGTTCCCCAAACTAATAGCTTTCAAACAAGTAGAGGAGCACATTTACTAAAGATACCTTTGCCTGC
TGCTCTTCACTTGTTAAAATACTCTCAGAGCCATTTGAGGAAGATTTTTTTATTCCCGCACTCATAGTTTTGAGGGG
AAACTCTGCAAATCAACAATGGAGATTTCAAAACTTATGTCCTAGTTTCACAGTTCCCTTCGGTCTCCCATCACCAT
CAAATACAATAAATTTCAATATATTTAACAAAAAAATTGCTCTTCATCCCACAAAACACAGAGTCCTCATCTTCATT
GTTCAATATATCATTTGAAATTAACAACTTTTATTCTTCTAGTCAACCACATTTCGCAGCTACTTGTTTAACTCATA
AACCCTTTCTTCCGATCCATAGCTATCAAATATCCAATCTAAACGAGACTACTACTTTGTTCACAACGAATCCAACA
CAAAAGGATCAAAAAAACCATCCAAAACTCATGCACAACATAATCAACCAAATATTTTAACCACAAAAACAAGCACA
ATTCTCCAAAGTACAAAAAGAAATGGGCTTTAGACACCAGGAAGGCATATCAAACCGGCCCACACACGTTAAAGGGA
TACAAAGATCTCACCTGGACCAAAGACAGAACTGGGTGGTTGCTGACTGAGCAAAGCCAATATCTCGGAGCTCCTCA
GATGTCGGAGAGACCCATCTGAACCCAAGTCAACTGCACTGTTACAGCAACTACAAAACGCAAAGATAGAGAGAGAG
AGAGAGAGAGAGAGAGAGAGAGAGAGAGAAGATGAGAATATTACACAGGGATCTGGGATCAGAGTGATGATGCTGTG
TTTGTGTTTGAGTTTGAAACTGAAGCCCAACCACGACCAACACGTGCTATTTGTAAGCCGAACCCATCTGCCTTCCT
TTCCTTCCATGTCTGTTCTGTGGTGTAGACTTTTCGGACAAAGCTTCCGTGGTGTCGTTTTTTTCCCAGGGTGGAAA
GGTCTGAACTGCCCCGGGGGACAACGGCGTGGTGTGGTGTGGCCGCTGCCCTTTTGAGGAAATTCACGTGGATTATG
GTGTGTCCTGCTTGTACTGTTGTCGGAGTTTACTAGGAACAATGAAATCATATCTATTTCTCATAAAAGGAAGCGTA
TTCTTTTTTATTAACCTTTTATTAATTACCTAGATTAGCTATTTCAAGTCAAAATTCATATATCGAAATATATGCTT
CTTGTCGCTAGACGCTATATTAGGTAACGAACATTTTAGGTAGTCGAACTGGTAGGCAGCGGCGGAGCTAGGATTTG
TTATTAGAGAGGTTAAGATGTAAAAGGTAAGTTTATGTATATGACACTTATTTATATATTTTTCCATTTTTAATGTA
AAATTCTATTATGATTTGGAAATCTGAGAAATATCTTATTTTCATTTGAGAAATGTCTTATTTTCATATGAGAAATG
GTGTGTGTGGGAGCTTTTGGGGGGGGGGGGGGGAGCTAAAGGCATTTTTCTATTTATATTAGTTGTTTAGAAGAAAA
AGAATGAACTATTTGGACATGACCCCTCGACCCAGCATGACCTGATTTTGAGATTAGAAGGGCACAAGTAACCAAAA

```
ATGTTGTTAGTTGGGACTAAAATCTTTCAATTGATCTTATGGCAGGTGGGCAAAAATTGTAAGCATTGCCCCCCAA
TTTCGAAGAGATGCCTTGAACTAAAAACCAGGAGTAGATGCAAAACAAAGACAAGAGCAACAAATTGGTAATAAACT
AGAGTGCATCTGAAACGTGCCCTATGCCAGCCAGCTTTGTCCATCGTAAGGACGAGAGCGACAATAACTATGTGACC
ACATGTGATGATGTGAACGGTACTCTTGTCATTTCGACGATCTTTCTTCTGAGATTTTTTCTATTCCTGGGTCACAT
ATGCGAAAGAAAACAGAGACAAAAGCATAGAAAAGAATTTGATATGATTGATTCCATAAGTGAAAGATGTATGTACA
TGTTAGCCGACACTACAAATGCTGCTAGAAAGAAATATACTACCTCCTAATACAATTTCTGTACCCCCATTTCTACA
CAGAATGTTGTAGCAAACTAGCAATGCCACCTTTCTGGTTTAAATTTCAAACAACAGAAACATTCTCCTTCATGTGT
TTGCATGTTTCCGAAAAATGTACAGAGCATCATTCAAATGGGATCTTTCTGCAATTGAAGCACCAGTATAGAACCAA
AGAGTCAGCAGTGTGTTTGTCGGAAGAGAACTATCGAGCCATGACTGTTTTCGCATCACAGGTTGCAGCTCAGATAT
TACAGCGTACAAGCTTTGGCGACACCACCATTTGGATTCCTGGGTTTCATTTGGTAGCCATGCATGAGAAGAAATGA
CAACAGAGGAGAAAAGAAAACTACGCGCATTCATAACTAAGTTCTATGCTTTTGTCTTTTATTTCTTGTAAGAGAAG
TTTTACCTCTGACGAATGTAAGCTGGTGGATTAAGGGTATGATTTTTCCACCATCTACCCAAAAGCATCCAGCTGCT
TGCAAGAAATCAAAGGAACAACATCAGTCAAGATGATTGTCTTCTTGCACAAGCCGAAAGCCCCAGCCAGCCCTCCT
CCCCCTTATTGCCAAACTTCACAACTAACCAAAAGCTCATCTAGCACCAAGTTTGGCTCTTCTAACTCAAGGAGTCC
GGCATCTGTAGTCACCACAGGAAACCAGCACTCTAAGCTCTGTGGCAGAAACATTACACGGTAGGAAAACAAAGAAT
GCCACCGAGCTCCTATGATGCAGAAAGAGGATTCATAACACCCATCAATCACTAGTTAGATCTCTCTCTTTCCAACA
ACCCCAGCAGGACTATGCTTCCTTTGTATACCATATTGCCAAGCATCTCTAGACTTATCACCCGACTGATCAAAGTC
GTCCAAAGAATTCTCAAACTGTGGTTTCATAGGCCG
```

>40M11_mandshurica

```
CAACATTTTGGTGGCCTTCTTGACATTCCAGTTTCTGGCCCTCAGATGCCTTGCAATGGATGCATCAGAACAGTATG
TGGACAGCTTCTCGGGTACTGCCTTTAACAATTTTCTCACCTCATTAATCTGCAAACAATAAGATTTTTTAGGCAAA
GCGGAACTATGAGTTCCCCAAACTAATAGCTTTCAAACAAGTAGAGGAGCACATTTACTAAAGATACCTTTGCCTGC
TGCTCTTCACTTGTTAAAATACTCTCAGAGCCATTTGAGGAAGATTTTTTTATTCCCGCACTCATAGTTTTGAGGGG
AAACTCTGCAAATCAACAATGGAGATTTCAAAACTTATGTCCTAGTTTCACAGTTCCCTTCGGTCTCCCATCACCAT
CAAATACAATAAATTTCAATATATTTAACAAAAAAATTGCTCTTCATCCCACAAAACACAGAGTCCTCATCTTCATT
GTTCAATATATCATTTGAAATTAACAACTTTTATTCTTCTAGTCAACCACATTTTGCAGCTACTTGTTTAACTCATA
AACCCTTTCTTCCGATCCATAGCTATCAAATATCCAATCTAAACGAGACTACTACTTTGTTCACAACGAATCCAACA
CAAAAGGATCAAAAAAACCATCCAAAACTCATGCACAACATAATCAACCAAATATTTTAACCACAAAAACAAGCACA
ATTCTCCAAAGTACAAAAAGAAATGGGCTTTAGACACCAGGAAGGCATATCAAACCGGCCCACACACGTTAAAGGGA
TACAAAGATCTCACCTGGACCAAAGACAGAACTGGGTGGTTGCTGACTGAGCAAAGCCAATATCTCGGAGCTCCTCA
GATGTCGGAGAGACCCATCTGAACCCAAGTCAACTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
```

>63F17_vesca

```
CGCTCTATGGAAGGGACAAGAGACACTGAAATAGCAATGGGGTCCTACCAACCTCATCATACATGGGCAAGAAATCA
TTCTAGTCCTCTCGGACAGGTAATCACAGAATCCAGATTAGATGCAGGTTTTGAATTATTAGAGTCTATAAAGGGAC
ATAGTTACAACTGTTTGTATGCTTTTCCATTTTTTTTATTTTTTTATTTTTTGAGAATGTATGCTTTTTCACTTGTA
TGGCCTGAAGTTGCGAATGTTTTGGTTGATAGATATTTGGATATAGAATGTCACTATGGGCAGAGCACACAGGAACC
GTTGAGGACTGTTTTAGAGAACCAGAGAGTCTTGAATGTGTTAGGAGAGTTAGAGCAATGGGTGAGATGAACTGGAA
ACAATTTGCTGCTGAGGAGGTTACAGAGATGAGGGGTCATCTATTGAAGTATCCAGTTGAAATTGATCGAAAAGGCA
AAGTCACATCCCTTCCTGGATGTGAGAGTTTCCCCGATGCAGGAGGAAATATAACCGGTTCTTTCCTTGGCATTCAA
GAAAATTTGACAATTTGATCACCAGTTCAGTTTTATAGAAGAACTCAGTTAGTACAGTTTTGAAACGTTTTTTTGTT
GTATTTAGCAAACCCATAGGAGGATAGGGTTTTCTTTTATTCAACAGGGATATAGGCGCTTTTAGGGTTTCTTTTCC
TATTCAATTTCGTTCTTTGGTAGACCAAGTCGCTTCTTTGGCATTCAAGGAAACCTGAGCATTTGATCTGCCTGTCA
TCACATCCAGAGTTGCAGATTGTTTAGAGAAGAATTCCAATAAATTCCTTTTGTACAGTTTGGTTAACTTTTGGTAT
TCAACAACGCATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGATATGCAGGTAACATCTCTGACTATGCAT
CTTTGCTTTTTCTTCTTTTTTTGAGAACAAGGCATCTTGTTTATGTGTAGCCAACTTGAAGCACTGTATTTAAATAA
TGCTAAAACAGTGTTAATTTTGTTATAAAAGTGTAGGCAACAATGAACTTGAACTTGAACTTGAACTTGATAAGAAA
ATAGATCCCAGAGATGGTCTATCTACTACCCTTGACTACACAAGTTACTCATTCTTTACATGTGAAATGGCTATCCA
GAGCAGTCGTATTTCATGAGATATTAACAAGCTTTGGACGTCCAGTTGCCAGGTTCATCTCTACTCGGAGGCCAAGT
CGAGCAAGGGCAGGCACATCAACAGACCCCTTAA
```

>63F17_viridis

```
CGCTCTATGGAAGGGACAAGAGACACTGAAATAGCAATGGGGGTCCTACCAACCTCATCATACATGGGCAAGAAATC
ATTCTAGTCCTCTCGGACAGGTAATCACAGAATCCAGATTATATGCCGGTTTTGAATTATTAGAGTCTATAAAGGGA
CTAGTTACAACTGTTTTTTTCCACTTTTTTTTTTTTTTTTTTTTTTTTGAGAATGTATGCTTTTCACTTATATGGCC
TGAAGTTGCGAATGTTTTGGTTGATAGATATTTGGATATAGAATGTCACTATGGGCAGAGCACACAGGAACCGTTGA
GGACTGTTTTAGAGAACCAGAGAGTCTTGAATGTGTTAGGAGAGTTAGAGCAATGGGTGAGATGAACTGGAGACAAT
TTGCTGCTGAGGAGGTTACAGAGATGAGGGGTCATCTATTGAAGTATCCAGTTGAAATTGATCGAAAAGGCAAAGTC
ACATCCCTTCCTGGATGTGAGAGTTTCCCCGATGCAGGAGGAAATATAACCGGTTCTTTCCTTGGCATTCAAGAAAA
TTTGACAATTTGATCACCAGTTCAGTTTTATTGAAGAACTCGGTACAGTTTTGAAACGTTTTTTTGTTGTATTTAGC
AAACGCATAGGAGGATAGGGTTTTCTTTTATTCAACAGGGATATAGGCGCTTTTAGGGTTTCTTTTCCTATTGAATT
TCGTTCTTTGGTATACCAAGTCCCTTCTTTGGCATTCAAAGAAACCTAAGCATTTGATCTGCCTGTCATCACGTCTA
GAGTTGCAGATTGTTTAGAGAAGAATTCCAATAAATTCCTTTTGTACAGTTTGGTTAACTTCTGGTGTTCAACAACG
CATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGATATGCAGGTAACATCTCTGACTATGCATCTTTGCTTT
TTCTTCTGTTTTTGAGAACAAGGCATCTTGTTTATTTGTGGCCAACTTGAAGCACTGTATTTAAATAATGCTAAGAC
CGTGTCAATTTTGTTACAAAAGTCTAGGCAACAATGAACTTGAACTTGATAAGAAAATAGATCCTAGAGATGGTCTA
TCTACTACCCTTGACTACACAAGTTGCTCATTCTTTACATGTGAAATGGCTATCCAGAGCAGTCGTAAATTTCATGA
GATATTAACAAGCTTTGGACGTCCAGTTGCCAGGTTCATGTCTACTCGGAGACCAAGTGAGCAAGGGCAGGCACATC
AACAGACCCCTAAA

>63F17_mandshurica
CGCTCTATGGAAGGGACAAGAGACACTGAAATAGCAATGGGGGTCCTACCAACCTCATCATACATGGGCAAGAAATC
ATTCTAGTCCTCTCGGACAGGTAATCACAGAATCCAGATTAGATGCAGGTTTTGAATTATTAGAGTCTATAAAGGGA
CATAGTTACAACTGTTTGTATGCTTTTCCATTTTTTTATTTTTTTATTTTTTGAGAATGTATGCTTTTTCACTTATA
TGGCCTGAAGTTGCGAATGTTTTGGTTGATAGATATTTGGATATAGAATGTCACTATGGGCAGAGCACACAGGAACC
GTTGAGGACTGTTTTAGAGAACCAGAGAGTCTTGAATGTGTTAGGAGAGTTAGAGCAATGGGTGAGATGAACTGGAA
ACAATTTGCTGCTGAGGAGGTTACAGAGATGAGGGGTCATCTATTGAAGTATCCAGTTGAAATTGATCGAAAAGGCA
AAGTCACATCCCTTCCTGGATGTGAGAGTTTCCCCGATGCAGGAGGAAGTATAACCGGTTCTTTCCTTGGCATTCAA
GAAAATTTGACAATTTGATCACCAGTTCAATTTTATAGAAGAACTCAGTTAGTACAGTATTGAAACGTTTTTTCGTT
GTATTTAGCAAACCCATAGGAGGATAGGGTTTTCTTTTATTCAACAGGGATATAGGCGCTTTTAGGGTTTCTTCTCC
TATTCAATTTCATTCTTTGGTAGACCAAGTCGCTTCTTTGGCATTCAAGGAAACCTGAGCATTTGATCTGCCTGTCA
TCACATCCAGAGTTGCAGATTGTTTAGAGAAGAATTCCAATTCCTTTTGTACAGTTTGGTTAACTTTTGGTGTTCAA
CAACGCATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGATATGCAGGTAACATCTCTGACTATGCATCTTT
GCTTTTTCTTCTATTTTTGAGAACAAGGCATCTTGTTTATGTGTGGCCAACTTGAAGCACTGTATTTAAATAATGCT
AAAACAGTGTTAATTTTGTTATAAAAGTGTAGGCAACAATGAACTTGAACTTGAACTTGATAAGAAAATAGATCCTA
GAGATGGTCTATCTACTACTCTTGACTACACAAGTTACTCATTCTTTACATGTGAAATGGCTATCCAGAGCAGTCGT
AAATTTCATGAGATATTAACAAGCTTTGGACGTCCAGTTGCCACGTTCATCTCTACTCGGAGGCCAAGTCGAGCAAG
GGCAGGCACATCAACAGACCC

>63F17Rrc_ananassa_2
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNTTCAATTTCATTCTGGGATAGACCAAGTCACTTTTTTGGCGTTCAAGGAAACCTGAGCATTTGATCTGCCGG
TCATCACATCCAGAGTTGCAGATTGTTTATAGAAGAATTCCAATAAATTCCTTTTGTACAGTTTGGTTAACTTTTGG
TGTTCAACAACGCATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGATATGCAGGTAACATCTCTGACTATG
CATCTTTGCTTTTTCTTCTTTTTTTGAGAATAAGGCATCTTGTTTATGTGTAGCCAACTTGAAGCACTGTATTTAAA
TAATGCTAAAACAGTGTTAATTTTGTTATAAAAGTGTAGGCAACAATGAACTTGAACTTGAACTTGAACTTGATAAG
AAAATAGATCCCAGAGATGGTCAATCTACTACCCTTGACTACACAAGTTACTCATTCTTTACATGTGAAATGGCTAT
CCAGAGCAGTCGTATTTCATGAGATATTAACAAGCTTTGGACGTCCAGTTGCCAGGTTCATCTCTACTCGGAGGCCA
AGTCGAGCAAGGGCAGGCACATCAACAGACCCTCAA
```

```
>72E18_vesca
GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAAATCCGGCCTAAGCATTAATATCAAATCAGTCCTTGAGAT
TCGACATGCATAAAAAAGACAATAAAGGGTACAAAAACAACCACTCAAACAATCACAACATAATATCATTCAATACC
TTGACCATTCCGGTTCCATTATCACACACAAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAACA
TACACCACAATCAATGACAACAATGCCTCATTCACACAACAAAGAAATAGACATTCAAAAACAAAACACAATACACA
CTACTAATGTGGCACAGAAACCAAAGCATGATTCAAAACAAAACTAGAACATCTACATAGTTCTCTCACAATAGTAA
AGAAACGATCTTTGACAATCAAAAGGCATCGAAAGCTAGTAAAGAAACGATCTTTCAGATGGGAAATACCCAAATTT
GATTGCTACATGCATAAAACCCTCAAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATAAAAAAA
ATTCAAGAAAAAAAGAGAGAGAAAATTACAGATTTAAAGCGACGAACAATGAAAAGGAATGAGAGGCAAAGAGAAGA
GATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGGGAGAGAGAGAGAGAGATCGACGACGAAGCAGAGCGAAA
GAGACGAGTGTGGTGTTTGTGAGTTGAGGCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCA
GCCGTTTGATTTATGGACCGCGTCTATTGAGCCCTTGTGGGGCCATTACAGCTCCTTCCGCTGTTCCAGTCATTTTT
TTCTCCACCTTCTTTACCTTTTTGCCCCTCAGTCCCTTCCCTTTTCTCCCAATTCTTTCTCAACTCTTCTTAAACCT
AATTGCATTTCCCTAATTGCATTTTCATTTTAGTGCTTAGATCAATATGATTAAGAAGCTTCATTTTGTCAACACAA
GGCAACAAGGACACAGGGGAGCATTTCGATCATCGTTCCAGTCATTTTCGTATATAATTTGGGCTTGAAATGGTTGA
TCGGTCGTAAAATTTAAAATGACGTTTTGATATGATATCTATAAAGAGGACATAATTTACTTTGTATGTCAGGTTTT
AATATGGAACGAAGAGTATGGGTGAAAGTGTCATCCCACACATTTTAAAAGAGCTGTAATGTAGGGTAATGAGCACA
ACTGCAAGCTGCATCCTATAATGGATCAATCAGAACATTAAACAACGTAAAGAGGAAGGTATTTGCTTTACACAACC
TTATAAAATGATGAGGATCTACTCAAAATCCAGACTACCATGGTTGGCAAAATTAGATCCTCACTGTAACCAGCTAG
GCATTGGTAATGCATAATAGCTATAGCTAACTATAGGTGGGAGACTCATCATTGAGATCATAGAAAAACAAAGATGA
AAGAAAGAAATGAAGAAACAAGCAACAGCTATTCGAAAGCAAGTACAGAAGGGATTGTTCATGAAGTGTTCACCAAG
TCACAGCTTAGGGCATTCTTAGAAGTAACAAGCTTACCAACTTCCATTTACTTGTTTCAAGTTCATGATGATATTAA
CCATCCAACGAAAACATCCAAAGGTACTGTGACAGAAAGCTCAAGGGGATATCTGTGTTTAAAGCCACAACATGACT
AAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTGCAAATGACAAAACAGTTCAAATTGACTGCATAAA
TAGATTACTCTTGTATAGATCAACAAGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAATACTAACATCAAT
ATTGAAATTGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCGTAGGAAATGGAATTGAACACTTGA
CACCAAAGAGAACAATGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGGGGGTAGAATCACCCATCGACGTGGAT
ACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTGGACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGA
GCTTCTACATACAAACTCAAATGGATATGTTCTGGCGCTTGTAGAATATAATTATGTATACAAATATGCATGTACAG
AGCTTCCACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTCCAATAAAGGTGAGTTTAAATAGAC
CAAGATGTTAGCTAAAAAAAACAGACAAAACATTTAAGCAAAAGAAGAGCAGTAGAAGGTATTAAGATACCAAACAA
CATATTTGGGTTGGAGGACAAAGTAGTATAGAGGAGTGTACCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATT
GGTCATGATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAAACCCAACCACT
GATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAAACCACGTGGTATTTGATAGCCTCCACTT
CCATCAGGACTTCGCAGATTCCTCAGTAGAATATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAA
CCCA
```

```
>72E18_viridis
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnTTAATATCAAAATCAGTCCTTGAGA
TTCAACATGCATAACAAAGACAATAAACGGTACAAAAACAACCACTCAAACAATCACAACATAATATCATTCAATAC
CTTGACCATTCCGGTTCCATTATCACACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAAC
ATACACCACAATCAATGACAACAATGCCTCATTCACACAACAAAGAAATAGACATTCAAAAACAAAACACAATACAC
ACTACTAATGTGGCACAGAAACCAAAGCATGATTCAAAACAAAACTAGAACATCTACATAGTTCTCTCACAATAGTA
AAGAAACGATCTTTGACAATCAAAAGGCATCGAAAGCTAGTAAAGAAACGATCTTTCAGATGGGAAATACCCAAATT
TGATTGCTACATGCATAAAACCCTCAAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATAAAAAA
AAAATTCAAGAAATAAAGAGAGAGAAAATTACAGATCTAAAGTGACGAACAATGAGAAGGAATGAGAGGCAGAGAGA
AGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGGGAGAGAGAGAGAGATCGAAGACGAAGCAGAGCGAA
AGAGACGAGTGTGGTGTTTGTGAGTTGAGGCGAAGAATTGNACCNNNATANAGGAGTGNGATTGACNAGTTATCTCN
GCNGNTTGATTTATGGACCGCGTCCATTGTGCCNTTGTGGGGCCATNACNGCTCCTNCCNCTGTNCCNGCCATNTTT
ATCTCCACCTTCTNCACNTTTNTGCCNCNCNTGNCCNTCCCTTTTCTCCCNNTTCTNNCTCANATCTTCTNAANGCT
NATNGCNTTTNCCNNANNGCATTTGNATTTNNGNGCTTCNATCAATNCGNNNNNAANGCTNCGTNTTGTCNTNNCAN
GGNTNCAAGGANCCTNNGGGANGCNNGTTGATCGTCAANTNCGGCCTGGTNATANANAATATNNGACNTNAAATGGT
TGATTGNNTCGTTAAATTNGAAATAAnnnnnnnnnnnnnnnnnnnnnAAAAGGNCNTAATTTANTNNGTANNTCAA
```

GGTTTTAANTAGGGAAATGGAGGAGTGNGGGTGAAAAGTGTCCATCCCACACGTTTTAAANGAGCNTTAANGNAGGG
NAANGNGCAGNAACNGCAAGCTGCATNCCTNTAATNGGATCAATCAGAACATTAAGNCAACGTAAAGAGGAAGGTAT
TTGCTTTACACAACCTTATAAAATGATGNGGATNTACTCCAAAGTGAGGACTACCATGGTCGGCAAAATTAGTTCCT
GACTGTAACCAGCTAGGCATNGGCAATGCATAATAGCTATAGCTGACTATAGGTGGGAGACTCATCATTGAGATCAG
AGAAAAACNAAGATGAAAGAAAGAAATGATGAAACAAGAATAGTTATTCGAAAGCAAGTACAGAAGGGATTGTTCAT
GAAGTGTTCACCAAGTCACAGCTTAGGGCATTCTTAGAAGCAACAAGCTTGCCAACTTCCATTTACTTGTTTCAAGT
TCATGNTGATATTANCCATCCAACGAAAACATCCAAAGGTNCTGTGACTGAAAGCTCAAGGGGATATCNGTGTTTAA
AGCCACTACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTGCAAATGACAAAACAGTTCT
AATTGACTGCATAAGATAGATTACTCTTGTATAGATCAACAAGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCA
GAAAACTAACATCAATATTGAAATTGGATAAATATGCGATCTGAACTTCTTCACGTTGATGAGCTATCGTAGGAAAT
GGAATTGAACACTTGACACCAAAGAGAACAACGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGGAGGTAGAATC
ACCCATCGACGTGGATACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTGGACATGGCAGGGTGTCACAAAAGATCA
ATATTGCATGCAAAGAGCTTCTACATACAAACTCAAATGGAGATGTTCTGGCGCTTGTAGAATATAATTATGTATAC
AAATATGCATGTACAGAGCTTCCACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTCCAATAAAG
GTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAAAAAAAAAAGACAAAACATTTAAGCAAAAGAAGAGCAGTAGAAG
TTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAAAGTCGTATAGAGGAGTGTACCTTCTTTAAACGGCGGT
GTTTTCCTAGGGCCCAATTGGTCATGATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACTGTCTGTGTTGCA
ATATTAAAACCCAACCACTGATAAATCTCAGTTGTGTAATTTGCACATGTCACAATATTGAATAGAAAACCACGTGG
TATTTGATAGCCTCCACTTCCATCAGGACTTCGCAGATTCCTCAGTAGAATATGGCAATAGAAGTTCGCAATTTGAT
TTATTATCCCAAAACCAAACCCA

>72E18_iinumae
GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAAACCCGGCCTAAACATTAACATCAAAATCAGTCCTTGAGA
TTCGACATGCATAAAAAAGACAATAAAGGGTACAAAAACAACCACTCAAACAATCACAACATAATAGCATCCAATAC
CTTGACCATTCCGGTTCCATTATCACACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCTAC
ATACACCACAATCAATGACACCAATGCCTCATCGACACAACAAACAAATAGACATTCAAAAACAAAACACAATACAC
GATGCTAACATTTCCCTAATCTCTCCTCCATCAACTAAAATCTCCATTCCAAATCACACACTACTACACAGAAACCA
AAGCATGATTCAAAACAAACCAAGAACATCTACATAGTCCTCTCACAATAGTAAAGAAACGATCTTTGACAATCAAA
AGGCATCGAAAGCTAGTAAAGAAACGATCTTTCAGATGGGAAATACCCAAATTTGATTGCTATATACATAAAACCCT
CAAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATTAAAAAAAAATCAAGAAAAAAAGAAGAGAG
AAAATTACAGATCTAAAGCGACGAACAAATGAGAAGGAATGAGAGACAGAGAGAAGAGATGAGGAAGTTGACCTTTG
TGAATGAGAGTGAGAGAGAGAGAGAGATCGAAGACGAGGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG
GCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATTTATGGACCGCGTCCATT
GCGCCCTTGTGGGGGCCATTACAGCTCCTTCCGCTGTTCCAGTCATTTTTCTCCACCTTCTTCACCTTTCTGCCCCT
CGTTCCCTTTCCCTCTTCTCCCAACTCTTCTTAGCCTAATTGCATTTTCATTTTAGTGCTTAGATCAATATGATTNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNCCGAAAGCTCAATGGGATATCTGTGTTTAAAGCCACAACATGACTAAATATA
ATTGCTCCCAATTTCTAAAGTTACATTCGTTTTGTACAAATGACAAAACAGTTCAAATTGACTGCATAAATAGATTA
CTCTTGTATAGATCAACAAGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAATACTAACATCAATATTGAAA
CTGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCGTAGGAAATGGAATTGAACACTTGACACCAAA
GAGAACAATGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGGGGGTAGAATCACCCATCGACGTGGATACTTGGG
TCTTCCGTCCTTCCCATCAAATAGCTGGACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTA
CATACAAACTCAAATGGATATGTTCTGGCGCTTGTAGAATATAATTATGTATACAAATATGCATGTACAGAGCTTCC
ACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTCCAATAAAGGTGAGTTTAAATAGACCAAGATG
TTAGCTAAAAAAAACAGACAAAACATTTAAGCAAAAGAAGAGCAGTAGAAGGTATTAAGATACCAAACAACATATTT
GGGTTGGAGGACAAAGTAGTATAGAGGAGTGTACCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATTGGTCATG
ATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAAACCCAACCACTGATAAAT

CTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAAACCACGTGGTATTTGATAGCCTCCACTTCCATCAG
GACTTCGCAGATTCCTCAGTAGAATATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAACCCA

>72E18_nilgerrensis
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNAAAAAAGAGAGAGAGATTACAGATCTANGCGACGAACAATGAGAAGGAATGAGAGGCAGAGAGAAGAGATGAGGA
AGTTGACCTTTGTGAATGAGAGTGAGTGAGAGAGAGAGATCGAAGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGT
TTGTGAGTTGAGGCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATTTATGG
ACCGCGTCCATTGCGCCCTTGTGGGGCCATTACAGCTCCTTCCGCTGTTCCAGTCATTTTTTTCTCCACCTTCTGCC
CCATTCCCTTCCCTTTTCTCCCAATTCTTTCTCAACTCTTCTTAAACCTAATTGCATTTTCATTTTANTGCTTANAT
CAATATGATTAAGAAGCTCCATTTTGTCAACACAAGGCAACAAGGACNTANGGGAGCATGTCGATCATCGTTCCGGT
CACTTTCGTATATAATTTGGACTTAAAATGGTTGATCGATCGTAAAATTTGAAATGACGTTTTGGTATGATATTTGT
AAAGAGGACATAATTTACTATGGAACGAAGAGTATAGGTGAAAGTGTCATCCCACACATTTTAAAAGAGCTTTAATG
TAGGGGTAATGAGCACAACTACAAGCTGCATCCTATAAGGGATCAATCAGAACATTAAACAACGTAAAGAGGAAGGT
ATTTGCTTTACACAACCTTATAAAATGATGAGGATCTACTCAAAATCCAGACTACCATGGTTGGCAAAATTAGATCC
TGCCTGTAACCAGCTAGGCATTGGCAATGCATAATAGCTAGAGCTAACCATAGGTGGGAGACTCATCATTGAGATCA
TAGGAAAAAAAATGATGAAAACAAGCAACAGTTATTCGAAAGCAAGTACAGAAGGGATTGTTCATTAAGTGTTCACC
AAGTCACAGCTTAGGGCATTCTTAGAAGCAACAAGCTTACCAACTTCCATTTACTTGTTTCAAGTTCATGATGATAT
TTAGTCATGTTGTTTAAAGCCACAACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTGCA
GATGACAAAACAGTTCAAATTGACTGCATAAGATAGATTACTCTTGTATAGATCAACAAGCAAATCTCCAAGTTCTT
ATTACAAAGTCTAAGCAGAACACTAACATAAATATTGAAATTGGATAAATATGCGATCTGAACTTCTTCACGTTGAT
GACCTATCATAGGAAATGGAATTGAACACTTGACACCAAAGAGAACAACGAAGGTAGCCTCGCCAATCACTTCTACA
AGAATGGGGGTAGAATCACCCATCTACGTGGATACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTGGACATGGCAG
GGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTACATACAAACTCATATGGATATGTTCTGGCGATTGCAG
AATATAATTATGTATACAAATATGCATGTACAGAGCTTCTACATACAAACTCATACGAATACTTGTAAATTTAGGCA
ATTTAATTCCAATAAAGGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAAAAGACAAAACATTTAAGCAAAAGAA
GAGCAGTAGAAGTTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAATGTAGCATAGAGGAGTGTACCTTCT
TTAAACGGCGGTGCTTTCCTAGGGCCCAGTTGGTCATTATAGAAGCAGCAACTGCAGCAAAAAGATAACCAGCAACC
GTCTGTGTTGCAATATTAAAACCCAACCACTGATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAG
AAAACCACGAGGTATTTGATAGCCTCCACTTCCATCAGGACTTCGCAGATTCCTCAGTAGAATATGGCAATACAAGT
TCGCGATTTGATTTATTATCCCAAAACCAAACCCA

>72E18_mandshurica
GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAGACCCGGCCTAAACATTAACATCAAAATCAGTCCTTGAGA
TTCAACATGCATAACAAAGACAATAAAGGGTACAAAAACAACCACTCAAACAATCACAACATAATATCATTCAATAC
CTTGACCATTCCGGTTCCATTATCACACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAAC
ATACACCACAATCAAATGCTACATTCACACAACAAAGAAATAGACATTCAAAGACAAAACACAAACACACTACTAAC
GTGGCACGGAAACCAAAGCATGATTCAAAACAAAACTAGAACATCTACATAGTTCTCTCACAATAGTAAAGAAACGA
TCGTTGACAATCAAAAGGCATCGAAAGCTAGTAAAGAAACGATCTTTCAGATGGGAAATACCCAAATTTGATTGCTA
TATACATAAAACCCTCAAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATAAAAAAAAATTCAAG
AAAAAAAGAGAGAGAAAATTACAGATCTAAAGCGACGAACAGTGAGAAGGAATGAGAGGCAGAGAGAAGAGATGAGG
AAGTTGACCTTTGTGAATGAGAGTGAGTGAGGGAGAGAGAGAGAGAGATCGACGACGAAGCAGAGCGAAAGAGACGA
GTGTGGTGTTTGTGAGTTGAGGCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTT
GATTTATGGACCGCGTCTATTGCGCCCTTGTGGGGCCATTACAGCTCCTTCCGCTGTTCCAGTCATTTTTTTCTCCA
CCTTCTTCACCTTTTTGCCCCTCAGTCCCTTCCCTTTTCTCCCAATTCTTTCTCAACTCTTCTTAAACCTAATTGCA
TTTCCCTAATTGCATTTTCATTTTAGTGCTGAGATCAATATGATTAAGAAGCTTCATTTTGTCAACACAAGGCAACA
AGGACACAAGGGAGCATGTCGATCATCGTTCCAGTCATTTTCGTATATAATTTGGGCTTGAAATGGTTAATCAATCG
TAAAATTTAAAATGACGTTTTGATATGATATCTATAAGGAGGACATAATTTACTTTATATGTCAGGTTTTAATACGG
AAGGAAGAGTATGGGTGAAAGTGTCATCCCACACATTTTAAAAGAGCTGTAATGTAGGGTAATGAGCACAACTGCAA
GCTGCATCCTATAATGGATCAATCAGAACAATAAACAACGTAAAGAGGAAGGTATTTGCTTTACACAACCTTATAAA

```
ATGATGAGGATCTACTCAAAATCCAGACTACCATGGTTGGCAAAATTAGATCCTCACTGTAACCAGCTAGGCATTGG
TAATGCATAATAGCTATAGCTAACTATAGGTGGGAGACTCATCATTGAGATCATAGAAAAACAGAGATGAAAGAAAG
AAATGATGAAACAAGCAACAGCTATTCGAAAGCAAGTACAGAAGGGATTGTTCATGAAGTGTTCACCAAGTCACAGC
TTAGGGCATTCTTAGAAGCAACAAGCTTACCAACTTCCATTTACTTGTTTCAAGTTCATGATGATATTAACCATCCA
ACGAAAACATCCAAAGGTACTGTGACAGAAAGCTCAAGGGGATATCTGTGTTTAAAGCCACAACATGACTAAATATA
ATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTGCAAATGACAAAACAGTTCAAATTGACTGCATAAATAGATAC
TCTTGTATAGATCAACAAGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAATACTAACATCAATATTGAAAT
TGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCGTAGGAAATGGAATTGAACACTTGACACCAAAG
AGAACAATGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGGGGGTAGAATCACCCATCGACGTGGATACTTGGGT
CTTCCGTCCTTCCCATCAAATAGCTGGACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTAC
ATACAAACTCAAATGGATATGTTCTGGCGCTTGTAGAATATAATTATGTATACAAATATGCATGTACAGAGCTTCCA
CATACAAACTCATATGTATACTTGTAAATTTATGCAATTTAATTCCAATAAAGGTGAGTTTAAATAGACCAAGATGT
TAGCTAAAAAAAAACAGACAAAACATTTAAGCAAAAGAAGAGCAGTAGAAGGTATTAAGATACCAAACAACATATTT
GGGTTGGAGGACAAAGTAGTATAGAGGAGTGTTCCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATTGGTCATG
ATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAAACCCAACCACTGATAAAT
CTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAAACCACGTGGTATTTGATAGCCTCCACTTCCATCAA
GACTTCTCAGATTCCTCAGTAGAATATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAACCCA
```

>72E18_ananassa

```
GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAAACCCGGCCTAAACATTAACATCAAAATCAGTCCTTGAGA
TTCAACATGCATAAAAAAGACAATAACGGGTACAAAAACAACCACTCAAACAATCACAACATAATATCATTCAATAC
CTTGACCATTCCGGTTCCATTATCACACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAAC
ATACATCACAATCAATGACAACAATGCCTCAAACAAATAGACATTCAAAAACAAAACACAATACACAATGCTAACAT
TTCCCTAATATCTCCTCCATCAACTAAAATCTCCATGCCAGATCACCACAGAAACCAAAGCATGATTCAAAACAAAC
CAAGAACATCTACATAGTTCTCTCACAATAGTAAAGAAACGATCTTTGACAATCAAAAGGCATCGAAAGCTAGTAAA
GAAACGATCTTTCAGATGGGAAATGCCCAAATTTGATTACTATATACATAAAACTCCCAAATTGATACGAAATCAAA
CAATGCAGCAATCAAATCATTCCACAGAAAAAAAATTCAAGAAAAAAAAAAGAGAGAGAAAATTACAGATCTAAAGC
GACGAACAATGAGAAGGAATGAGAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAGTGAGGGAGAGAGAGAG
AGAGATCGAAGACGAAGCTGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAGGCGAAAGAATTGGAGCAAAATA
AAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATTTATGGACCGCGTCCGTTGCGCCCTTGTGGGGCCATTGCA
GCTCCTTCCGCTGTTCCAGTCATTTTTCTCCACCTTCACCTTTTTGCCCCTCATTCCCTTTCCCTCTTCTCCGATCC
CAACTTTCTCAACTCTTCTTAACCCACCCAGTTGCATTTTCATTGTAGTGCTTAAATCAATATGGTTAAGAAGCTTC
ATTTTGTCaacacaaggcaacaaggacataggggagcatgtggatgatcgtttttaaactacgttttggtatgatat
cgtaaggaggacataatttactttgtatgtcnggtttttaatacggaatgaagagtgtgggagaaagtgtcatccca
cacattcatttaannnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnCAAAATCCAGACTACCGTGGTGGGCAAAATTAGATCCTCACCTATAGTAGGGTAACTATAGGTG
GGAGACTNGTCATTGTGATCATAGAAAAAGAAATGATGAAACAAGCAACAGTTACTCGAAAGCAGGTACAGAAGGGA
TTGTTCATGAAGTGTTCACCAAGTCNCAGCTTAGGGCATTCTTAGAAGCAACAAGCTTACCAACTTCCATTTACTTG
TTTCAAGTTCAGGGTGACATTACCCNTCCAACGAAAACATCCAAAGGTACTGTGACTGAAAGCCCAAGGGGATATCC
GTGTTTAAAGCCACAACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTGCAAATGACAAA
ACAGTTCAAATTGACTGCATAAGATAGATTACTCTTGTATAGATCAACAAGCAAATCTCCAAATTCTTATTACAAAG
TCTAAGCAGAATACTAACATCAATATTGAAATTGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCG
TAGGAAATGGAATTGAACACTCGACACCAAAGAGAACAACAAAGGTAGCCTCACCAATCACTTCTACAAGAATGGGG
GTAGAATCACCCATCGACGTGGATACTTGGGTCGTCCGTCCTTCCCATCAAATAGCTGGACATGGCAGGGTGTCACA
AAAGATCAATATTGCATGTAAAGAGCTTCTACATACAAACTCAAATGGATATGTTCTGGCGCTCGTAGAATATAATT
ATGTATACAAATATGCATGTACAGAGCTTCCACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTC
CAATAAAGGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAACAGACAAAACATTTAAGCAAAAGAAGAGCAGTAG
AAGTTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAAAGTAGTATAGAGGAGTGTACCTTCTTTAAACGGC
GGTGTTTTCCTAGGGCCCAATTGGTCATGATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTT
GCAATATTAAAACCCAACCACTGATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAAACCACG
TGGTATTTGATAGCCTCCGCTTCCATCAGGACTTCGCAGATTCCTCAGTAGAATATGGCAATAGAAGTTCGCAATTT
GATTTATTATCCCAAAACCAAACCCA
```

APPENDIX E
GENE-PAIR HAPLOTYPE INDIVIDUAL LOCI ALIGNMENTS

To avoid identification of untrue polymorphisms and therefore poor alignments, "N"s were inserted where the chromatogram did not display distinct peaks. EcoRI sites are in green because EcoRI was used to extract inserts from TOPO vectors and detect polymorphic inserts in this vector. Simple Sequence Repeats (SSRs) are magenta-colored.

Gene Pairs Detected by Microcolinearity

GPH5: Single Nucleotide Polymophisms (SNPs) are in bold. SNPs that occur in more than one clone are likely to reflect real differences (as oppose to amplification or sequencing errors) and are colored red.

```
GPH5_ananassa_clone2      CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
GPH5_ananassa_clone7      CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
GPH5_viridis              CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
GPH5_iinumae              CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
GPH5_nilgerrensis         CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
GPH5_mandshurica          CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
GPH5_nubicola             CAATGCCATGGTCTCCGGTCTATTTCAACCGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
GPH5_vesca                CAATGCCATGGTCTCCGGTCTATTTCAACTGGGAAGTTCTTATGAGTGGGTGGTGACAAA 60
                          ***************************** ******************************


GPH5_ananassa_clone2      GAAGACCGGAAGATCATCGGGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAG 120
GPH5_ananassa_clone7      GAAGACCGGAAGATCATCGGGAATCAGATTTGTTCGCGCTTGCAGAAAGAGAATCCTCGAG 120
GPH5_viridis              GAAGACCGGAAGATCATCAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAG 120
GPH5_iinumae              GAAGACCGGAAGATCCTCAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAG 120
GPH5_nilgerrensis         GAAGACTGGAAGATCATCAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAG 120
GPH5_mandshurica          GAAGACCGGAAGATCATCAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAG 120
GPH5_nubicola             GAAGACCGGAAGATCATCAGAATCAGATTTGTTTGCTCTTGCAGAAAGAGAATCCTCGAG 120
GPH5_vesca                GAAGACCGGAAGATCATCAGAATCAGATTTGTTTGCGCTTGCAGAAAGAGAATCCTCGAG 120
                          ****** ******** ** ************** ** ***********************


GPH5_ananassa_clone2      TGAAGAGAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATCCTTGAGCAAACTCAA 180
GPH5_ananassa_clone7      TGAAGAGAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATCCTTGAGCAAACTCAA 180
GPH5_viridis              TGAAGACAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATTGTTGAGCAAACTCAA 180
GPH5_iinumae              TGAAGACAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATTGTTGAGCAAACTCAA 180
GPH5_nilgerrensis         TGAAGACAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATTGTTGAGCAAACTCAA 180
GPH5_mandshurica          TGAAGACAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATTGTTGAGCAAACTCAA 180
GPH5_nubicola             TGAAGACAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATTGTTGAGCAAACTCAA 180
GPH5_vesca                TGAAGACAAGATCCTAAGGAGGAACTCCGAGTCTGGTTTAGAATTGTTGAGCAAACTCAA 180
                          ****** *************************************  **************


GPH5_ananassa_clone2      GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_ananassa_clone7      GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_viridis              GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_iinumae              GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_nilgerrensis         GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAACT 240
GPH5_mandshurica          GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
GPH5_nubicola             GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAACT 240
GPH5_vesca                GGAACAAGAAGTAGCACCTCCCAAGAAGAAGAAGAAAAATGGGATCTACAGAAAAGAGCT 240
                          ********************************************************* **
```

```
GPH5_ananassa_clone2      TGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
GPH5_ananassa_clone7      TGCTCTTGCTTTCCACCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
GPH5_viridis              TGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
GPH5_iinumae              TGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
GPH5_nilgerrensis         TGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
GPH5_mandshurica          TGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
GPH5_nubicola             TGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
GPH5_vesca                TGCTCTTGCTTTCCTCCTACTCACAGCATCAGCAAGAAGTTTCCTATCAGCTCATGGAGT 300
                          ************** *********************************************


GPH5_ananassa_clone2      TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
GPH5_ananassa_clone7      TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
GPH5_viridis              TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
GPH5_iinumae              TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
GPH5_nilgerrensis         TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
GPH5_mandshurica          TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
GPH5_nubicola             TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
GPH5_vesca                TCACTTCTATTTCTTGCTTTTCCAAGGCTTGTCCTTTCTTGTTGTAGGCTTGGACTTAAT 360
                          ************************************************************


GPH5_ananassa_clone2      AGGTGAGCAGGTTAGCTAAAAGCTTCAAACAAAGCGTCAATTGCCCACAGTTATTCTTTG 420
GPH5_ananassa_clone7      AGGTGAGCAGGTTAGCTAAAAGCTTCAAACAAAGCGTCAATTGCCCACAGTTATTCTTTG 420
GPH5_viridis              AGGTGAGCAGGTTAGCTAAAAGCTTCAAACAAAGCGTCAATTGCCCACAGTTATTCTTTG 420
GPH5_iinumae              AGGTGAGCAGGTTAGCTAAAAGCTTCAAACAAAGCATCAATTGCCCACAGTTATTCTTTG 420
GPH5_nilgerrensis         AGGTGAGCAGGTTAGCTAAAAGCTTCAAACAAAGCGTCAATTGCCCACAGTTATTCTTTG 420
GPH5_mandshurica          AGG-GAGCAGGTTAGCTAAAAGCTTCAAACAAAGCGTCAATTGCCCACAGTTATTCTTTG 419
GPH5_nubicola             AGGTGAGCAGGTTAGCTAAAAGCTTCAAACAAAGCGTCAATTGCCCACAGTTATTCTTTG 420
GPH5_vesca                AGGTGAGCAGGTTAGCTAGAAGCTTCAAACAAAGCGTCAATTGCCCACAGTTATTCTTTG 420
                          *** ************** **************** ************************


GPH5_ananassa_clone2      ATAGATATATGCTGAA----CTGTAAGAGACATATTTCAAGCTCTTTGGTGTTCAAAGTT 476
GPH5_ananassa_clone7      ATAGATATATGTTGAA----CTGTAAGAGACATATTTCAAGCTCTTTGGTGTTCAAAGTT 476
GPH5_viridis              ATAGATATATGTTGAA----CTGTAAGAGACATATTTCAAGCTC--TGGTGTTCAAAGTT 474
GPH5_iinumae              ATAGATATATGTTGAA----CTGTAAGAGACATATTTCAAGCTCTTTGGTGTTCAAAGTT 476
GPH5_nilgerrensis         ATAGATATATGTTGAA----CTGTAAGAGACATATTTCAAGCTCTTTGGTGTTCAAAGTC 476
GPH5_mandshurica          ATAGATATATGTTGAA----CTGTAAGAGACATATTTCAAGCTCTTTGGTGTTCAAAGTT 475
GPH5_nubicola             ATAGATATATGTTGAATGAACTGTAAGAGACATATTTCAAGCTCTTTGGTGTTCAAAGTT 480
GPH5_vesca                ATAGATATATGTTGAA----CTGTAAGAGACATATTTCAAGCTCTTTGGTGTTCAAAGTT 476
                          *********** ****    ************************  *************


GPH5_ananassa_clone2      GGATTCAATTACATGTAGACACAGTTACCATTTTTCCATTTGAAACAGAAGGTAATATGC 536
GPH5_ananassa_clone7      GGATTCAATTACATGTAGACACAGTTACCATTTTTCCATTTGAAACAGAAGGTAATATGC 536
GPH5_viridis              GGATTCATTTACATGT-GACACAGTTACCATTTTCCCATATGAAATAGAAGGTAATATGC 533
GPH5_iinumae              GGATTCAATTACATGTAGACATAGTTACCATTTTCCCATTTGAAATAGAAGGTAATAT-- 534
GPH5_nilgerrensis         GGATTCAATTACATGTAGACACAGTTACCATTTTCCCATTTGAAACAGAAGGTAATATGC 536
GPH5_mandshurica          GGATTCAATTACATGTAGACACAGTTACCATTTTCCCATTTGAAATAGAAGGTAATACGC 535
GPH5_nubicola             GGATTCAATTACATGTAGACACAGTTACCATTTTCCCATTTGAAATAGAAGGTAATACGC 540
GPH5_vesca                GGATTCAATTACATGTAGACACAGTTACCATTTTCCCATATGAAATAGAAGGTAATATGC 536
                          ******* ******** **** ************ **** ***** ***********


GPH5_ananassa_clone2      ATGATATAAATATCAAGTTAATTGTACAATGATATTATTTGTAATAAGTGAGAATAATGA 596
GPH5_ananassa_clone7      ATGATATAAATATCAAGTTAATTGTACAATGATATTATTTGTAATAAGTGAAAATAATGA 596
GPH5_viridis              ATGATATAAATATCAAGTTAATTGTACAGTGATAT---TTGTAACCAGTGAAAATAATGA 590
GPH5_iinumae              -----------ATCAAGTTAATTGTACAATAATAT---TTGTAATCAGTGAAAATAATGA 580
GPH5_nilgerrensis         ATGATATAAATACCAAGTTAATTGTACAATGATAT---TTGTAATCAGTGAAAATAATGA 593
GPH5_mandshurica          ATGATATAAATATCAAGTTAATTGTACAATGATAT---TTGTAATCAGTGAAAATAATGA 592
GPH5_nubicola             ATGATATAAATATCAAGTTAATTGTACAATGATAT---TTATAATCAGTGAAAATAATGA 597
GPH5_vesca                ATGATATAAATATCTAGTTAATTGTACAATGATAT---TTGTAACCAGTGAAAATAATGA 593
                                     * * ************* * ****   ** ***  ***** ********
```

```
GPH5_ananassa_clone2      CAATCTTCATAACAAAATTTCAGTTATCTTTCCATTGCTGTATGAACTGTTACCATTAGC 656
GPH5_ananassa_clone7      CAATCTTTATAACAAAATTTCAGTTATCTTTCCATTGCTGTATGAACTGTTACCATTAGC 656
GPH5_viridis              CAATCTTTATAACAAAATTTCAGTTATCTTTCCATTGCTGTATGAACTGTTACCATTAGC 650
GPH5_iinumae              CAATCTTTATAACAAAATTTCAGTTATCTTTTCACTGCTGTATGAACTGTCACCATTAGC 640
GPH5_nilgerrensis         AAATCTTTATAACAAAATTTCAGTTATCCTTCCATTGCTGTGTGAACTGTTACCATTAGC 653
GPH5_mandshurica          CAATCTTTATAACAAAATTTCAGTGATCTTTCCATTGCTGTATGAACTGTTACCATTAGC 652
GPH5_nubicola             CAATCTTTATAACAAAATTTCAGTGATCTTTCCATTGCTGTATGAACTGTTACCATTAGC 657
GPH5_vesca                CAATCTTTATAACAAAATTTCAGTTATCTTTCCATTGCTGTATGAACTGTTACCATTAGC 653
                           ****** **************** *** ** ** ****** ******** *********


GPH5_ananassa_clone2      CTCTCACACAAGAACAACTACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 716
GPH5_ananassa_clone7      CTCACACACAAGAGCAACAACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 716
GPH5_viridis              CTCTCACACAAGAACAGCAACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 710
GPH5_iinumae              CTCTCACACAAGAACAACAACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 700
GPH5_nilgerrensis         CTCTCACACAAGAACAACAACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 713
GPH5_mandshurica          CTCTCACACAAGAACAACAACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 712
GPH5_nubicola             CTCTCACACAAGACCAACAACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 717
GPH5_vesca                CTCTCACACAAGAACAACAACACCAAACAAACAGAACCAGACCAAATCACACCAATATAA 713
                          *** ********* ** * *****************************************


GPH5_ananassa_clone2      AACAGAATTGGATTTTCATGAAAGGCGGCAAGGCACAATCAATGAAGGAGA--AGACAAA 774
GPH5_ananassa_clone7      AACAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGA--AGACAAA 774
GPH5_viridis              AACATAATTGGATTTTCATGAAAGGCAGCAAGGCATGATCAATGAAGGAGA--AGACAAA 768
GPH5_iinumae              AACAGAATTGGATTTCCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGA--AGACAAA 758
GPH5_nilgerrensis         AACAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGAGAAGACAAA 773
GPH5_mandshurica          A-CAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGA--AGACAAA 769
GPH5_nubicola             A-CAGAACTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGA--AGACAAA 774
GPH5_vesca                AACAGAATTGGATTTTCATGAAAGGCAGCAAGGCACAATCAATGAAGGAGA--AGACAAA 771
                          * ** ** ******* ********** ********  **************  *******


GPH5_ananassa_clone2      GAATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATC 834
GPH5_ananassa_clone7      GAATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATC 834
GPH5_viridis              GAATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATC 828
GPH5_iinumae              GAAACCTTTTGTCATAGGGATTGAACCTGAATTATCTGGAGTGTTTCTGGCTGTCATATC 818
GPH5_nilgerrensis         GAATCCTTTTGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATC 833
GPH5_mandshurica          GAATCCTTTCGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATC 829
GPH5_nubicola             GAATCCTTTCGTCATATGGATTGAATCTGAATTATTTGGAGTGTTTCTGGCTGTCATATC 834
GPH5_vesca                GAATCCTTTTGTCATATGGATTGAATCTGAATTAGTTGGAGTGTTTCTGGCTGTCATATC 831
                          *** ***** ****** ******** ********  ************************


GPH5_ananassa_clone2      TCATATGCAGGCATGTTACATGTCTG----------CATTTGGTGACAAAAGCTAAATCT 884
GPH5_ananassa_clone7      TCATATGCAGGCATGTTACATGTCTG----------CATTTGGTGACAAAAGCTAAATCT 884
GPH5_viridis              TCATATGCAGGCATGTTACATGTCTG----------CATTTGGTGACAAAAGCTAAATCT 878
GPH5_iinumae              TCATATGCAGGCATGTTACATGTCTG----------CATTTGGTGACAAAAGCTAAATCT 868
GPH5_nilgerrensis         TCATATGCAGGCATGTTACATGTCTG----------CATTTGGTGACAAAAGCTAAATCT 883
GPH5_mandshurica          TCATATGCAGGCATGTTACATGTCTG----------CATTTGGTGACAAAAGCTAAATCT 879
GPH5_nubicola             TCGTATGCAGGCATGTTACATGTCTG----------CATTTGGTGACAAAAGCTAAATCT 884
GPH5_vesca                TCATATGCAGGCATGTTACATGTCTCATGATGTCTTCATTTGGTGACAAAAGCTAAATCT 891
                          ** **********************           ************************


GPH5_ananassa_clone2      TAACATGACCTAAGAATTAAGACATATTGGACCATTGGGCTTAATCATAGTCTAAGCCCA 944
GPH5_ananassa_clone7      TAACATGACCTAAGAATTAAGACATATTGGACCATTGGGCTTAATCATAGTCTAAGCCCA 944
GPH5_viridis              TAA----------GAATTAAGACATATTGGACCATTGGGCTTAATCATAGTCTGAGCCCA 928
GPH5_iinumae              TAA----------GAATTTAGACGTATTAGACCATTGGGCTTAATCATCGTCCGAGCCCA 918
GPH5_nilgerrensis         TAACATGACCTAAGAATTAACACATATTGGACCATTGGGCTTAATCATAGTCTAAGCCCA 943
GPH5_mandshurica          TAACATGACCTAAGAATTAAGACATATTGGACCATTGGGCTTAATCATAGTCTAAGCCCA 939
GPH5_nubicola             TAACATGACCTAAGAATTAAGACATATTGGACCATTGGGCTTAATCATAGTCTAAGCCCA 944
GPH5_vesca                TAACCTGACCTAAGTATCAAGACATATTGGACAATTGGGCTTAATCATAGTCTAAGCCCA 951
                          ***          * **  * ** **** *** *************** ***  ******
```

```
GPH5_ananassa_clone2      AATCTGTACTAGCCCATAATATGCTTTTTATAGAAA--ATA--CT----GTGATCTTCAC 996
GPH5_ananassa_clone7      AATCTGTACTAGCCCATAATATGCTTTTTGTAGAAA--ATA--CT----GTGATCTTCAC 996
GPH5_viridis              AATCTGTACTAGCCCATAATATGCTTTTTATAGAAA--ACA--CTCTCTGTGATCTTCGC 984
GPH5_iinumae              AATCTGCACTAGCCCGTAATATGCTTTTTATAGAAA--ACAGACTCTCTGTGATCTTCGC 976
GPH5_nilgerrensis         AATCTGTACTAGCCCATAATATGCTTTTTATAGAAAC-AGAGATTCTCTGTGATCTTCAC 1002
GPH5_mandshurica          AATCTGTACTAGCCCATAATATTCTTTTTATAGAAAACAGAGGTTCTCTGTGATCTTCAC 999
GPH5_nubicola             AATCTGTACTAGCCCATAATATTCTTTTTATAGAAAACAGAGATTCTCTGTGATCTTCAC 1004
GPH5_vesca                AATCTGTACTAGCCCATAATATGCTTTTTATAGAAA--ACA----CTCTGTGATCTTCAC 1005
                          ****** ******** ****** ****** ******  * *        ********* *

GPH5_ananassa_clone2      CATTGAGGAGTCAAGTTACTCAGCCATGAAGTCAAGGTCAAGCCAAGTAGTGCAGTTGAG 1056
GPH5_ananassa_clone7      CATTGAGGAGTCAAGTTACTCAGCCATGAAGTCAAGGTCAAGCCAAGTAGTGCAGTTGAG 1056
GPH5_viridis              CATTGAGGAGTCAAGTTACTCAGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAG 1044
GPH5_iinumae              CATTGAGGAGTCAGGTTACTCAGCTCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAG 1036
GPH5_nilgerrensis         CATTGAGGAGTCAAGTTACACAGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAG 1062
GPH5_mandshurica          CATTGAGGAGTCAAGTTACTCGGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAG 1059
GPH5_nubicola             CATTGAGGAGTCAAGTTACTCGGCCCTGAAGTAAAAGTCCAGTCAAGTAGTGCAGTTGAG 1064
GPH5_vesca                CATTGAGGAGTCAAGTTACTCAGCCCTGAAGTAAAAGTCCAGTCA-GTAGTGCAGTTGAG 1064
                          ************* ***** * **  ****** ** *** ** ** **************

GPH5_ananassa_clone2      TTCAACTTGTTCTGGGTTCTTCAAAGTTCGAAACTTTAAGCTTCAATGGAGGAAGAGAAG 1116
GPH5_ananassa_clone7      TTCAACTTGTTCTGGGTTCTTCAAAGTTCGAAGCTTTAAGCTTCAATGGAGGAAGAGAAG 1116
GPH5_viridis              TTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAACTTTAAGCTTCAATGGAGGAAGAGAAG 1104
GPH5_iinumae              TTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAACTTTAAGCTTCAATGGAGGAAGATAAG 1096
GPH5_nilgerrensis         TTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAACTTTAAGCTTCAATGGAGGAAGAGAAG 1122
GPH5_mandshurica          TTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAACTTTAAGCGTCAATGGAGGAAGAGAAG 1119
GPH5_nubicola             TTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAGCTTTAAGCGTCAATGGAGGAAGAGAAG 1124
GPH5_vesca                TTCAACTTGTTCTGGGTTCTTCAAAGTTTGAAACTTTAAGCTTCGATGGAGGAAGAGAAG 1124
                          **************************** *** ******** ** *********** ***

GPH5_ananassa_clone2      GATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATGTATAAAAGCTTGAAG-GA 1175
GPH5_ananassa_clone7      GATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGCTTGAAG-GA 1175
GPH5_viridis              GATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGCTTGAAG-GA 1163
GPH5_iinumae              GATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGCTTGAAG-GA 1155
GPH5_nilgerrensis         GATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGCTTGAAG-GA 1181
GPH5_mandshurica          GATGCCTTTTATGTTGCTCCAACGGGAGATGTGGTTGGCATATATACAAGCTTGAAG-GA 1178
GPH5_nubicola             GATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGCTTGAAG-GA 1183
GPH5_vesca                GATGCCTTTTATGTTGTTCGAAAGGGAGATGTGGTTGGCATATATAAAAGCCTGAAAAGA 1184
                          **************** ** ** ****************** **** **** ****  **

GPH5_ananassa_clone2      TTGCC--AAAACCAAGCTGG-TTCATCGGTAAA-GTTTT--GATCTTTT--AAGCCTTTT 1227
GPH5_ananassa_clone7      TTGCC--AAAACCAAGCTGG-TTCATCGGTAAA-GTTTT--GATCTTTT--AAGCCTTTT 1227
GPH5_viridis              TTGCC--AAAACCAAGCTGG-TTCATTGGTAAAAGTTTT--GATCTTTT--AAGCCTTTT 1216
GPH5_iinumae              TTGCC--AAAACCAAGCTGG-TTCCTCGGTAAA-GTTTT--GATCTTTT--AAGCCCTTT 1207
GPH5_nilgerrensis         TTGCC--AAA-CCAAGCTGG-TTCCTCGGTAAA-GCTTT--GATCTTTT--AAGCCTTTT 1232
GPH5_mandshurica          TTGCC--AAAACCAAGCTGG-TTCATCGGTAAA-GTTTT--GATCTTTT--AAGCCTTTT 1230
GPH5_nubicola             TTGCC--AAAACCAAGCTGG-TTCATCGGTAAA-GTTTC--GATCTTTT--AAGCCTTTT 1235
GPH5_vesca                TTGCCCAAAACCCAAGCTGGGTTCATCCGAAAAAGTTTTTGAATCTTTTTAAAGCCCTTT 1244
                          *****  *** ********* *** *  * *** * **    *******  ***** ***

GPH5_ananassa_clone2      ----GTAATTTG----ATCACTCCCATTGTTTTATCAATTTTTGATTTCCCATTTGATTA 1279
GPH5_ananassa_clone7      ----GTAATTTG----ATCACTCCCATTGTTTTATCAATTTTTGATTTCCCATTTGATTA 1279
GPH5_viridis              ----ATAATTTG----ATCACTCTCATTGTTTTATCAATTTNNNNNNNNNNNNNNNNNNN 1268
GPH5_iinumae              ----ATAATTTG----ATTACTCTCATTGTTTTATCAATTTTTGATTTCCCATTTGATNN 1259
GPH5_nilgerrensis         ----ATAATTTG----ATTACCCTTATTGTTTTATCAANNNNNNNNNNNNNNNNNNNNNN 1284
GPH5_mandshurica          ----GTAATTTG----ATCACTCCCATTGTTTTATCAATTTTAGATTTCCCATTTGATTA 1282
GPH5_nubicola             ----ATAATTTG----ATCACTCTCATTGTTTTANNNNNNNNNNNNNNNNNNNNNNNNNN 1287
GPH5_vesca                TTTAATAATTTGGAATNCCACCTTCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1304
                               *******       **
```

```
GPH5_ananassa_clone2      CATTACTGGGTCTTGTTTATTTTGTTGAAATAACTATGCCCTTTCGTTCTAGCATGCAAC 1339
GPH5_ananassa_clone7      CATTACTGGGTCTTGTTTATTTTGTTGAAATAACTATGCCCTTTCGTTCTAGCATGCAAC 1339
GPH5_viridis              NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1328
GPH5_iinumae              NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1319
GPH5_nilgerrensis         NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1344
GPH5_mandshurica          CATTACTGGCTCTTGTTTATTTTGTTGAACTAACTATGCCCTTTCGTTCTAACATGCAAC 1342
GPH5_nubicola             NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1347
GPH5_vesca                NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1364


GPH5_ananassa_clone2      TGAAATTTACTGCTAGATTGTATTGTTGTGCCGTTATGGTGTTCATTATGTAAAAGAGAA 1399
GPH5_ananassa_clone7      TGAAATTTACTGCTAGATTGTATTGTTGTGCCGTTATGGTGTTCATTATGTAAAAGAGAA 1399
GPH5_viridis              NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1388
GPH5_iinumae              NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1379
GPH5_nilgerrensis         NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1404
GPH5_mandshurica          TGAAAATAACTGCTAGATTGTATAGCTGAGCCTTTATGGTGTTCATTATGTAAAAGAGAA 1402
GPH5_nubicola             NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1407
GPH5_vesca                NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1424


GPH5_ananassa_clone2      TGAATTCTGGTGGTGGGTATAGAGTACCTCCCTGATTTTTTATGAGATACTATGCTTCTG 1459
GPH5_ananassa_clone7      TGAATTCTGGTGGTGGGTATAGAGTACCTCCCTGATTTTTTATGAGATACTATGCTTCTG 1459
GPH5_viridis              NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1448
GPH5_iinumae              NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1439
GPH5_nilgerrensis         NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1464
GPH5_mandshurica          TGAATTCTGGTGGTGGGTATAAAGCACCTCCCTGAGATTATATGAGATACTATGCTTCTG 1462
GPH5_nubicola             NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1467
GPH5_vesca                NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1484


GPH5_ananassa_clone2      GAAAATGTTATAAA-------GATGAAAACTAGTTTTTCTAAGAACTTGATG-----TTA 1507
GPH5_ananassa_clone7      GAAAATGTTATAA--------GATGAAAACTAGTTTTTCTAAGAACTTGATG-----TTA 1506
GPH5_viridis              NNNNNNNNNNNNN---------NNNNNNNNNNNNNNNNNNNNGAACTTGATG-----TTA 1494
GPH5_iinumae              NNNNNNNNNNNNN--------NNNNNNNACTAGTTTTTCTAAGAACTTGATG-----TCA 1486
GPH5_nilgerrensis         NNNNNNNNNNNNNN-------NNNNNNNNNNNGTTTTTCTAAGAACTTGATG-----TTA 1512
GPH5_mandshurica          GAAAATGTTATAA--------GATGAAAACAACTTTTTCTAACAACTTGATG-----TTA 1509
GPH5_nubicola             NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTAAGAACTTGATG-----TTA 1522
GPH5_vesca                NNNNNNNNNNNNN--------NNNNNNNNNNNNNNNNNNNNNGTATTAGATGAAAACTAG 1536
                                                                      * * ****     *


GPH5_ananassa_clone2      TTT--------ACTTGATGAGTTGATGGAGGATTACGTATGTGGTTTGGTTTTGTTTTTA 1559
GPH5_ananassa_clone7      TTT--------ACTTGATGAGTTGATGGAGGATTACGTATGTGGTTTGGTTTTGTTTTTA 1558
GPH5_viridis              TTT--------ACTTGATGAGTTGACGGAGGATTACATATAAGGCTTGGTTATGTTTTTA 1546
GPH5_iinumae              TTT--------ACTTGATGAGTTGATGGAGGATTACACATGTGGTTTGGTTTTGTTTTTA 1538
GPH5_nilgerrensis         TTT--------ACTTGATGAGTTGATGGAGGATTACATATGTGGTTTGGTTTTGTTTTTA 1564
GPH5_mandshurica          TTT--------ACTTGATGAGTTGATGGAAGATTACGTATGTGGTTTGGTTTTGTTTTTA 1561
GPH5_nubicola             TTT--------GCTTGATGAGTTGATGGAGGATTACATATGAGGTTTGGTTATGTTTTTA 1574
GPH5_vesca                TTTTTCTAAGAACTTGATGAGTTGATGGAGGATTACATATGAGGTTTGGTTATGTTTTTA 1596
                          ***         ************* *** ******  **  ** ****** ********


GPH5_ananassa_clone2      GGTATTCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1619
GPH5_ananassa_clone7      GGTATTCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1618
GPH5_viridis              GGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1606
GPH5_iinumae              GGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1598
GPH5_nilgerrensis         GGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1624
GPH5_mandshurica          GGTATTCAATCCTTCTGAAATTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1621
GPH5_nubicola             GGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1634
GPH5_vesca                GGTATGCAATCCTTCTGTAAGTGTGTTTAAAGGGTATGGTTTGCCTAAGGAGGCCGAGGA 1656
                          ***** *********** ** ***************************************
```

```
GPH5_ananassa_clone2      GTACCTTGTCTCACATGGGCTTAAGAATGCTTCATGTACTATCAGTGCCAGTGATGTGAA 1679
GPH5_ananassa_clone7      GTACCTTGTCTCACATGGGCTTAAGAATGCTTCATGTACTATCAGTGCCAGTGATGTGAA 1678
GPH5_viridis              GTACCTTGTCTCGCATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAA 1666
GPH5_iinumae              GTACCTTGTCTCACATGGGCTTAAGAATGCTTCATATACTATCAGCGCCAGTGATGTGAA 1658
GPH5_nilgerrensis         GTACCTTGTCTCACATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAA 1684
GPH5_mandshurica          GTACCTTGTCTCGCATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAA 1681
GPH5_nubicola             GTACCTTGTCTCGCATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAA 1694
GPH5_vesca                GTACCTTGTCTCGCATGGGCTTAAGAATGCTTCATATACTATCAGTGCCAGTGATGTGAA 1716
                          ************ ********************** ********* **************


GPH5_ananassa_clone2      AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGAT---TTCAT 1736
GPH5_ananassa_clone7      AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGAT---TTCAT 1735
GPH5_viridis              AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATGATTTCAT 1726
GPH5_iinumae              AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGAT---TTCAT 1715
GPH5_nilgerrensis         AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGAT---TTCAT 1741
GPH5_mandshurica          AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATGATTTCAT 1741
GPH5_nubicola             AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATGATTTCAT 1754
GPH5_vesca                AGATGGTCTGTTTGGAAGCCTTGTTGCTTGTCCTTACCAGGTTTGAATTGATGATTTCAT 1776
                          ****************************************************   *****


GPH5_ananassa_clone2      GTGTTCTAGTTTCTGTTTGGGCATCTGTTATTTTCATGGCATGTGGCGTGGAGCTAGTTG 1796
GPH5_ananassa_clone7      GTGTTCTAGTTTCTGTTTGGGCATCTGTTATTTTCATGGCATGTGGCGTGAAGCTAGTTG 1795
GPH5_viridis              GTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAGCTAGTTG 1786
GPH5_iinumae              GTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAGCTAGTTG 1775
GPH5_nilgerrensis         GTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAGCTAGTTG 1801
GPH5_mandshurica          GTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTCGCGTGGAACTAGTTG 1801
GPH5_nubicola             GTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAACTAGTTG 1814
GPH5_vesca                GTGTTCTAGTTTCTGTTTGGGTATCTGTTATTTTCATGGCATGTGGCGTGGAACTAGTTG 1836
                          ********************* ********************** ***** * *******


GPH5_ananassa_clone2      CATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCAGTTTTTTACTATAAT 1856
GPH5_ananassa_clone7      CATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCAGTTTTTTATTATAAT 1855
GPH5_viridis              CATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAAGTCGGTTTTTTATTATAAT 1846
GPH5_iinumae              CATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCGGTTTTTTATTATAAT 1835
GPH5_nilgerrensis         CATATGGTATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCGGTTTTTTATTATAAT 1861
GPH5_mandshurica          CATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCGGTTTTTTATTATAAT 1861
GPH5_nubicola             CATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTCGGTTTTTTATTATAAT 1874
GPH5_vesca                CATATGATATTAAATCTTTTGTTTGGTTCCTCATTGTATAATTTGGTTCTTTATTATAAT 1896
                          ****** ********************************** *  *** **** ******


GPH5_ananassa_clone2      CTTTCAGCAGCCAGCATCATCTATGGTTAATTCAGGGTTCAAAATTGCGCCTAATCAGCT 1916
GPH5_ananassa_clone7      CTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCGCCTAATCAGCT 1915
GPH5_viridis              CTTTCAGCAGCCAGCATCTTCCATGGTTAATTCAGGGTTCAAAATTGCACCTAATCAGCT 1906
GPH5_iinumae              CTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCACCTAATCAGCT 1895
GPH5_nilgerrensis         CTTTCAGCAGCCAGCATCTTCTATGGTTGATTCAGGGTTCAAAATTGCACCTAATCAGCT 1921
GPH5_mandshurica          CTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCACCTGATCAGCT 1921
GPH5_nubicola             CTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCACCTGATCAGCT 1934
GPH5_vesca                CTTTCAGCAGCCAGCATCTTCTATGGTTAATTCAGGGTTCAAAATTGCACCTGATCAGCT 1956
                          ****************** ** ****** ******************* *** *******


GPH5_ananassa_clone2      TACACCAAAGAGAATGCCGGATGTGGTAAATTCAGGCGTCAATGACCCTCCACAAAAGAG 1976
GPH5_ananassa_clone7      TACACCAAAGAGAATGCCGGATGTGGTAAATTCAGGCGTCAATGACCCTCCACAAAAGAG 1975
GPH5_viridis              TACACGAAAGAGAATGCCGGATGTGGTCAATTCAGGCGTCAATGACCCTCCACAAAAGAG 1966
GPH5_iinumae              TACACGAAAGAGAATGCCGGATGTGGTAAAGTCAGGCGTCAATGACCCTCCACAAAAGAG 1955
GPH5_nilgerrensis         TACACGAAAGAGAATGCCGGATGTGGTAAATTCAGGCGTCAATGACCCTCCACAAAAGAG 1981
GPH5_mandshurica          TACACGAAAGAGAATGCCGGATGTGATAAATCCAGGTGTCAATGACCCTCCACAAAAGAG 1981
GPH5_nubicola             TACACGAAAGAGAATGCCGGATGTGATAAATTCAGGTGTCAATGACCCTCCACAAAAGAG 1994
GPH5_vesca                TACACGAAAGAGAATGCCGGATGTGATAAATTCAGGTGTCAATGACCCTCCACAAAAGAG 2016
                          ***** ******************* * **  **** ***********************
```

```
GPH5_ananassa_clone2      ATCGCTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTTTGATAGAAGACCTTC 2036
GPH5_ananassa_clone7      ATCGCTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTTTGATAGAAGATCTTC 2035
GPH5_viridis              ATCGCTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTTTGATAGAAGATCTTC 2026
GPH5_iinumae              ATCATTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTTTGATAGAAGATCTTC 2015
GPH5_nilgerrensis         ATCGCTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTTTGATAGAAGATCTTC 2041
GPH5_mandshurica          ATCACTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTATGATAGAAGATCTTC 2041
GPH5_nubicola             ATCATTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTATGATAGAAGATCTTC 2054
GPH5_vesca                ATCACTGGATGTAAGTATCATATGCTACATGGAACTTTTGTAGTATGATAGAAGATCTTC 2076
                          ***  *************************************** ********** ****


GPH5_ananassa_clone2      TATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCTTTCAGCAGCCAGCGTCT 2096
GPH5_ananassa_clone7      TATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCTTTCAGCAGCCAGCGTCT 2095
GPH5_viridis              TATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCTTTCGGCAGTCAGCATCT 2086
GPH5_iinumae              TATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATCTTTCAGCAGCCAGCGTCT 2075
GPH5_nilgerrensis         TATTTGGTTACTCATTGTATGATTCAGGTTTTTATTATAATCTCTCAGCAGCCAGCGTCT 2101
GPH5_mandshurica          TATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATTTTTCAGCAGCCAGGGTCT 2101
GPH5_nubicola             TATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATTTTTCAGCAGCCAATGTCT 2114
GPH5_vesca                TATTTGGTTACTCATTGTATGATTCGGGTTTTTATTATAATTTTTCAGCAGCCAGCGTCT 2136
                          ************************* *************** * ** **** **   ***


GPH5_ananassa_clone2      TCTATGGTTAATTCAGGCTTCAAAGTTGCACCTGATCAGTTTACACGAATGCGATTGCCG 2156
GPH5_ananassa_clone7      TCTATGGTTAATTCAGGCTTCAAAATTGCACCTGATCAGTTTACACGAATGCGATTGCCG 2155
GPH5_viridis              TCTATGGTTAATTCAGGCTTCAAAATTGCACCTTACCAGCTTGCACGAATGAGATTGCCG 2146
GPH5_iinumae              TCTATGGTTAATTCAGGCTTCAAAATTGCACCTTATCAGCTTACACGAATGAGATTGCCG 2135
GPH5_nilgerrensis         TCTATGGTTAATTCAGGCTTCAAAATTGCACCTTATCAGCTTACACGAATGAGATTGCCG 2161
GPH5_mandshurica          TCTATGGTTAATTCAGGCTTCAAAATTGCACCTCATCAGCTTACGCGAATGAGATTGCCG 2161
GPH5_nubicola             TCTTTGGTTAATTCAGGCTTCAAAATTGCACCTCATCAGCTTACGCAAATGAGATTGCCG 2174
GPH5_vesca                TCTATGGTTAATTCAGGCTTCAAAATTGCACCTCATCAGCTTACGCGAATGAGATTGCCG 2196
                          *** ******************** ******** * *** ** * * **** ********


GPH5_ananassa_clone2      GATGTGGTAAATTCAGG-TGTCAATTACCCTCCACAGAGGACATTGCCGGATGTAAGTAT 2215
GPH5_ananassa_clone7      GATGTGGTAAATTCAGG-TGTCAATTACCCTCCACAGAGGACATTGCCGGATGTAAGTAT 2214
GPH5_viridis              GATGTGATAAATTCAGG-TGTCAATGACCCTCCACAGAGGACATTGCCAGATGTAAGTAT 2205
GPH5_iinumae              GATGTGGTAAATTCAGG-TGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGTAT 2194
GPH5_nilgerrensis         GATGTGGTAAATTCAGNGTGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGCAT 2221
GPH5_mandshurica          GATGTGGTAAATCCAGG-TGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGTAT 2220
GPH5_nubicola             GATGTGGTAAATTCAGG-TGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGTAT 2233
GPH5_vesca                GATGTGGTAAATTCAGG-TGTCAATGACCCTCCACAGAGGACATTGCCGGATGTAAGTAT 2255
                          ****** ***** ***  ******* ********************** ******** **


GPH5_ananassa_clone2      CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAGATCCAGACACCTTCAG 2275
GPH5_ananassa_clone7      CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGACACCTTCAG 2274
GPH5_viridis              CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGATACCTTCAG 2265
GPH5_iinumae              CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCGGATACCTTCAG 2254
GPH5_nilgerrensis         CTTATGCTACATGGAAATTTTATTCTGGCCAGATTGGTATGAAAATCCAGATACCTTCAG 2281
GPH5_mandshurica          CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGATACCTTCAG 2280
GPH5_nubicola             CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGATACCTTCAG 2293
GPH5_vesca                CTTATGCTACATGGAAATTTTGTTCTGGCCAGATTGGCATGAAAATCCAGATACCTTCAG 2315
                          ********************* *************** ***** **** ** ********


GPH5_ananassa_clone2      TCTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGAT 2335
GPH5_ananassa_clone7      TTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGAT 2334
GPH5_viridis              TTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGAT 2325
GPH5_iinumae              TTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGAT 2314
GPH5_nilgerrensis         TTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCCGCAAATGAT 2341
GPH5_mandshurica          TTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGAT 2340
GPH5_nubicola             TTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGA- 2352
GPH5_vesca                TTTGGCTGGATTATGGAGTTGCGTTGATCACTTGTTTATTGTATTTATTCTGCAAATGAT 2375
                          * ************************************************ ********
```

```
GPH5_ananassa_clone2      GTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGTAATCGAA 2395
GPH5_ananassa_clone7      GTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGTAATCGAA 2394
GPH5_viridis              GTTTT-CAGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGTAATCGAA 2384
GPH5_iinumae              GTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGTAATCGAA 2374
GPH5_nilgerrensis         GTTTTTCGGCTTCCAGTTTTTCTGCACATAAGCATTTTAAAGCTGATAATTGTAATCGAA 2401
GPH5_mandshurica          GTTTTTTGGCTTCCAGTTTTTCTTCACATAAGCATTTTAAAGCTGATCATTGTAATCAAA 2400
GPH5_nubicola             --TTTTTGGCTTCCAGTTTTTCTTCACATAAGCATTTTAAAGCTGGTCATTGTAATTGAA 2410
GPH5_vesca                GTTTTTTGGCTTCCAGTTTTTCTTCACATAAGCATTTTAAAGCTGATCATTGTAATCGAA 2435
                            ***   *************** ********************* * ********  **


GPH5_ananassa_clone2      CTCAAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATT 2455
GPH5_ananassa_clone7      CTCAAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCAC-ACTAAGATCACAATT 2453
GPH5_viridis              CTCGAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGGTCACAATT 2444
GPH5_iinumae              CTCGAGTAATTCTGCTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATT 2434
GPH5_nilgerrensis         CTCGAGTAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATT 2461
GPH5_mandshurica          CTCGAATAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATT 2460
GPH5_nubicola             CTCGAATAATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATT 2470
GPH5_vesca                CTCGAATTATTCTACTACTGGTGTAAGTTGCCTTGTGTCACCACCACTAAGATCACAATT 2495
                          *** * * ***** ****************************** ****** ********


GPH5_ananassa_clone2      TCGTATTTTATGATCAACACTGAATACCTATGTCTAGTGTCATGATTATAGTCATGTGAA 2515
GPH5_ananassa_clone7      TCGTATTTTATGATCAACACTGAATACCTATGTCTAGTGTCATGATTATAGTCATGTGAA 2513
GPH5_viridis              TCGTATTTTATGATCAACACTGAATACCTATGTCTAGTGTCGTGATTATGGTCATGTGAA 2504
GPH5_iinumae              TCGTATTTTATGATCAACACTGAATACCTATGTCTAGTGTTGTGATTATGGTCATGTGAA 2494
GPH5_nilgerrensis         TCGTATTTTATGATCAACACTGAATACCTATGTCTAGTGTCGTGATTATGGTCATGTGAA 2521
GPH5_mandshurica          TTGTATTTTATGATCAACACCGAATACCTATGTCTAGTGTCGTGATTGTGGTCATGTGAA 2520
GPH5_nubicola             TCGTATTTTATGATCAACACCGAATACCTATGTCTAGTGTCGTGATTATGGTCATGTGAA 2530
GPH5_vesca                TCGTATTTTATGATCAACACCGAAGACCTATGTCTAGTGTCGTGATTATGGTCATGTGAA 2555
                          * ****************** *** ***************  ***** * **********


GPH5_ananassa_clone2      GTGGATTTCTTAATATATGCCTCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2575
GPH5_ananassa_clone7      GTGGATTTCTTAATATATGCCTCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2573
GPH5_viridis              GTGGATTTCTTAATATATGCCTCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2564
GPH5_iinumae              GTGGATTTCTTAATATATGCCTCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2554
GPH5_nilgerrensis         GTGGATTTCTTAATATATGACTCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2581
GPH5_mandshurica          GTGGATTTCTTAATATATGCCTCATCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2580
GPH5_nubicola             GTGGATTTCTTAATATATGCCTTGTCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2590
GPH5_vesca                GTGGATTTCTTAATATATGCCTCGTCTATGTCTTCATCCAGCAATCCTGCATACTTGAGT 2615
                          ******************* **  ************************************


GPH5_ananassa_clone2      TTGATGGTGCTTCAAAAGGAAATCCTGGACCATCTGGTGCAGGAGCTGTACTCCGTGCTG 2635
GPH5_ananassa_clone7      TTGATGGTGCTTCAAAAGGAAATCCTGGACCATCTGGTGCAGGAGCTGTACTCCGTGCTG 2633
GPH5_viridis              TTGATGGAGCTTCAAAAGGAAGTCCTGGATTATCTGGTGCAGGAGCTGTACTCCGTGTTG 2624
GPH5_iinumae              TTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGTGCAGGAGCTGTACTCCGTGCTG 2614
GPH5_nilgerrensis         TTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGTGCAGGAGCTGTACTCCGTGCTG 2641
GPH5_mandshurica          TTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGTGCAGGAGCTGTACTCCGTGCTG 2640
GPH5_nubicola             TTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGTGCAGGAGCTGTACTCCGTGCTG 2650
GPH5_vesca                TTGATGGAGCTTCAAAAGGAAATCCTGGATTATCTGGTGCAGGAGCTGTACTTCGTGCTG 2675
                          ******* ************* *******  ********************** **** **


GPH5_ananassa_clone2      AAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATAT 2695
GPH5_ananassa_clone7      AAGATGGAAGTGTTGTATGCGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATAT 2693
GPH5_viridis              AAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATAT 2684
GPH5_iinumae              AAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATAT 2674
GPH5_nilgerrensis         AAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATCTTTTAGGATATATAT 2701
GPH5_mandshurica          AAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATTTTTTG-ATATATAT 2699
GPH5_nubicola             AAGATGGGAGTGTTGTATGTGGAGTTCATGAAAGCATTGTGAATTTTTTTTATATATAT 2710
GPH5_vesca                AAGATGGGAGTGTTGTATGTGGAGTTCATGAAAACATTGTGAATTTTTTG-ATATATAT 2734
                          ******* *********** ************* ********** ****   ********
```

```
GPH5_ananassa_clone2      TTTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGCTACTATAGTTGCACCGGCT 2755
GPH5_ananassa_clone7      TTTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGTTACTATAGTTGCACCGGCT 2753
GPH5_viridis              ATTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGTTACTGTAGTTGCACCGGCT 2744
GPH5_iinumae              TTTTGTTTTTGTAAAAGTGGATCTCTTTATAACATTGGGTTTACTATAGTTGCACCGGCT 2734
GPH5_nilgerrensis         TTTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGTTACTATAGTTGCACCGGCT 2761
GPH5_mandshurica          TTTTGTTTTTGTAAAAATGGATCTCTTTATAACATTGGGGTTGCTATAGTTGCACCGGCT 2759
GPH5_nubicola             TTTTGTTTTTGTAAAAATGG----TTTTATAACATTGGGGTTACTATAGTTGCACCGGCT 2766
GPH5_vesca                TTTTGTTTTTGTAAAAATGGATCTCTTCATAACATTGGGGTTACTATAGTTGCACCGGCT 2794
                           *************** ***     ** ***********  * ** **************


GPH5_ananassa_clone2      GCGGGAAGGTGTGTGCAACGGCA 2778
GPH5_ananassa_clone7      GCGGGAAGGTGTGTGCAACGGCA 2776
GPH5_viridis              GCGGGAAGGTGTGTGCAACGGCA 2767
GPH5_iinumae              GCGGGAAGGTGTGTGCAACGGCA 2757
GPH5_nilgerrensis         GAGGGAAGGTGTGTGCAACGGCA 2784
GPH5_mandshurica          GCGGGAAGGTGTGTGCAACGGCA 2782
GPH5_nubicola             GCGGGAAGGTGTGTGCAACGGCA 2789
GPH5_vesca                GCGGGAAGGTGTGTGCAACGGCA 2817
                          * *********************
```

GPH23: SNPs other than introduced by DNA polymerase are true. After preliminary sequence alignment, the putative SNPs were verified by observation of unambiguous peaks in the chromatograms.

```
GPH23_iinumae_clone2          CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTGCTAACG 50
GPH23_iinumae_clone5          CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTGCTAACG 50
GPH23_mandshurica_clone3      CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTACTAACG 50
GPH23_ananassa_clone4         CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTACTAACG 50
GPH23_ananassa_clone3         CTTGAGGGCCATCAGCACGTCCCTTCTGCAATACCATCTTAGTACTAACG 50
                              ******************************************* ******

GPH23_iinumae_clone2          ACCTTTACAGTGAGAGTGTGACCAGAGGTGCCTGGGCGGAGCTGCCCAAC 100
GPH23_iinumae_clone5          ACCTTTACAGTGAGAGTGTGACCAGAGGTGCCTGGGCGGAGCTGCCCAAC 100
GPH23_mandshurica_clone3      ACCTTTACAGTGAGAGTGTGACCAGAGGTGCCTGGGCGGAGCTGCCCAAC 100
GPH23_ananassa_clone4         ACCTTTACAGTGAGAGTGTGACCAGAGGTGCCTGGGCGGAGCTGCCCAAC 100
GPH23_ananassa_clone3         ACCTTTACAGTGAGAGTGTGACCAGAGGTGCCTGGGCGGAGCTGCCCAAC 100
                              **************************************************

GPH23_iinumae_clone2          CTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTTGAGTCTGCCATTTGAT 150
GPH23_iinumae_clone5          CTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTTGAGTCTGCCATTTGAT 150
GPH23_mandshurica_clone3      CTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTCGTGTCTGCCATTTGAT 150
GPH23_ananassa_clone4         CTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTCGAGTCTGCCATTTGAT 150
GPH23_ananassa_clone3         CTTTGTGAAGGTTGGTTTCCTCAGGGCTTGCTTCGAGTCTGCCATTTGAT 150
                              ********************************* * **************

GPH23_iinumae_clone2          AAAAGACCTGCCAGAATCCACGCCACCAAAC-TCTTTAGCACTAATCCAA 199
GPH23_iinumae_clone5          AAAAGACCTGCCAGAATCCACGCCACCAAAC-TCTTTAGCACTAATCCAA 199
GPH23_mandshurica_clone3      AAAAGACCTGCCAGAATCCACACCACCAAAC-TCTTTAGCACCAATCCAA 199
GPH23_ananassa_clone4         AAAAGACCTGCCAGAATCCACACCACCAAAC-TCTTTAGCACTAATCCAA 199
GPH23_ananassa_clone3         AAAAGACCTGCCAGAATCCACGCCACCAAACATCTTTAGCACTAATCCAA 200
                              ********************* ********* ********** *******

GPH23_iinumae_clone2          TCCATAACAACTTCATAAAACACACATAGCATCAACATGCAATAATGTGG 249
GPH23_iinumae_clone5          TCCATAACAACTTCATAAAACACACATAGCATCAACATGCAATAATGTGG 249
GPH23_mandshurica_clone3      TCCATAACAACTTCATAAAACACACATAGCATCAACATGCAATAATGTGG 249
GPH23_ananassa_clone4         TCCATAACAACTTCATAAAACACACATAGCATCAACATGCAATGATGTGG 249
GPH23_ananassa_clone3         TCCATAACAACTTCATAAAACACACATAGCATCAACATGCAATAATGTGG 250
                              ******************************************* ******
```

```
GPH23_iinumae_clone2            GTCCATAAGAACCATGAGTATGACATA-GAGTCTTCAAGCTTCGATTTCC 298
GPH23_iinumae_clone5            GTCCATAAGAACCATGAGTATGACATA-GAGTCTTCAAGCTTCGATTTCC 298
GPH23_mandshurica_clone3        GTCCACAAGAACCATGAGTATGAC-CA-TAGTCTTCAAGCTTCGATTTCC 297
GPH23_ananassa_clone4           GTCCATAAGAACCATGAGTATGACATA-GAGTCTTCAAGCTTCGATTTCC 298
GPH23_ananassa_clone3           GTCCATAAGAACCATGTGCATGACATAAGATTCCTCAAGCTTCGATTTCC 300
                                ***** ********** * *****  *  * ** ****************


GPH23_iinumae_clone2            TTATTTGCTTCGAAAGAAGCAAGTTCAGAGTCACACAAACCAGAATATAG 348
GPH23_iinumae_clone5            TTATTTGCTTCGAAAGAAGCAAGTTCAGAGTCACACAAACCAGAATATAG 348
GPH23_mandshurica_clone3        TTATTCGCTTCGAAAGAAGCAAGTTCAGAGTCACACAAACCAGAACATAG 347
GPH23_ananassa_clone4           TTATTTGCTTCGAAAGAAGCAAGTTCAGAGTCACACAAACCAGAATATAG 348
GPH23_ananassa_clone3           TAATTTGCTTCAAAAGAAGTAAGTTCAGAGTCACTCAAACCCTAATATAG 350
                                * *** ***** ******* ************** ******  ** ****


GPH23_iinumae_clone2            ATCTCAAA--TTTAATGAAACATATTCCTAAGAACCTAAAGCAATATAAA 396
GPH23_iinumae_clone5            ATCTCAAA--TTTAATGAAACATATTCCTAAGAACCTAAAGCAATATAAA 396
GPH23_mandshurica_clone3        ATCTCAAAAATTTAATGAAACATATTCCTAAGAACCTAAAGAAATATAAA 397
GPH23_ananassa_clone4           ATCTCAAA--TTTAATGAAACATATTCCTAAGAGCCTAAAGAAATATAAA 396
GPH23_ananassa_clone3           ATCTCAAA--TTTAATGAAACATATTCCTAAGAGCCTACACAAATATAAA 398
                                ********  *********************** **** *  ********


GPH23_iinumae_clone2            ATCGTAACTGGACTTAATCTGAAATTGTCGTATAAATTGTAAATCGATCA 446
GPH23_iinumae_clone5            ATCGTAACTGGACTTAATCTGAAATTGTCGTATAAATTGTAAATCGATCA 446
GPH23_mandshurica_clone3        ATCGTAACTGAACTTATTCTGAAATTGTCGTATAAATTGTAAACCGATCA 447
GPH23_ananassa_clone4           ATCGTAACTGAACTTAATCTGAAATTGTCGTATAAATTGTAAATCGATCA 446
GPH23_ananassa_clone3           ATCGTAACTGAACTTAATCTGAAACTGTCGTATAAATTGTAAATCGATCA 448
                                ********** ***** ******* ****************** ******

GPH23_iinumae_clone2            AAAACAAACTTCAAGTTCAGATTCACAGAGCAGATCAGAGATAGCATACA 496
GPH23_iinumae_clone5            AAAACAAACTTCAAGTTCAGATTCACAGAGCAGATCAGAGATAGCATACA 496
GPH23_mandshurica_clone3        ACA-CAAACTTCAAGTTCAGATTCACAGAGCAGATCAGAGATAGCATACC 496
GPH23_ananassa_clone4           AAAACAAACTTCAAGTTCAGATTCAGAGAGCAGATCAGAGATAGCGTACA 496
GPH23_ananassa_clone3           AAACCAAACTTCATGTTCAGATTCACAGACCGTATCAGAGATAGCATACA 498
                                * * ********* *********** *** *  ************ ***

GPH23_iinumae_clone2            AGTGACCTAAGAAACAAAACAACATTCTAACAAGATCGCAAACATTGGAG 546
GPH23_iinumae_clone5            AGTGACCTAAGAAACAAAACAACATTCTAACAAGATCGCAAACATTGGAG 546
GPH23_mandshurica_clone3        AGTGACCTAAGAAACAAAACAACCTTCTAACAAGATCGCAAACATTGGAG 546
GPH23_ananassa_clone4           AGTGACCTAAGAAACAAAACAAAATTCCAACAAGATCGCAAACATTCGAG 546
GPH23_ananassa_clone3           AGTGACCTCTGAAACAAAACATAATTCCAACAAGATCGCAAACATTCGAG 548
                                ********  ***********   *** ****************** ***

GPH23_iinumae_clone2            ATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAA 596
GPH23_iinumae_clone5            ATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAA 596
GPH23_mandshurica_clone3        ATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAA 596
GPH23_ananassa_clone4           ATTAAATACGATGAGCTATGAGACAACTTTTCCATGCAAATCTAACAAAA 596
GPH23_ananassa_clone3           ATTAAATACGATGAGCTATGAGACAACTATTCCATGCAAATCTAACAAAA 598
                                **************************** *********************

GPH23_iinumae_clone2            GAAACTAAAGGGATCTGGAGAATTAGGGGTTAGAGGTCACCTTAAGAGTT 646
GPH23_iinumae_clone5            GAGACTAAAGGGATCTGGAGAATTAGGGGTTAGAGGTCACCTTAAGAGTT 646
GPH23_mandshurica_clone3        GAGAATAAAGGGATCTGGAGAATTAGGGGTTAGAGGTCACCTTAAGAGTT 646
GPH23_ananassa_clone4           GAGAATAAAGGGATCTGGAGAATTAGGGGTTAGAGGTGACCTTAAGAGTT 646
GPH23_ananassa_clone3           AAGAATAAAGGGATCTGGAGAATTATGGGTTAGAGGTGACCTTCAGAGTT 648
                                 * * ******************** *********** ***** ******

GPH23_iinumae_clone2            TCGGTGAAACACAACACAACTGGGGAGACAGAGACAGAGGAGGAACTGCG 696
GPH23_iinumae_clone5            TCGGTGAAACACAACACAACTGGGGAGACAGAGACAGAGGAGGAACTGCG 696
GPH23_mandshurica_clone3        TGGGTGAAACACAAC----CTGGGGAGACAGAGACAGAGGAGGAACTGCC 692
GPH23_ananassa_clone4           TGGGTGAAACACAAC----TGGGGAAGACAGAGACAGAGGAGGAACTGCG 692
GPH23_ananassa_clone3           TGGGTGAAACACAAC----TGGGGACGACCTACACCGAGGAGGAACTGCC 694
                                * *************      ***  ***  * ** *************
```

```
GPH23_iinumae_clone2          AATATCTATCTGAA-ACCAAACAAAGTAAAAAGGGTTTAGCTGTCAGTAA 745
GPH23_iinumae_clone5          AATATCTATCTGAA-ACCAAAACAAGTAAAAAGGGTTTAGCTGTCAGTAA 745
GPH23_mandshurica_clone3      AAGATCTATCTGACCACCAAACCAAGTAAAAAGGGTTAAGCTATCAGTAA 742
GPH23_ananassa_clone4         AAGATCTATCTGAA-ACCAAACAAAGTAAAAAGGGTTTAGCTGTCAGTAA 741
GPH23_ananassa_clone3         AAAATCTATCTGAA-ACCTAAC-AAATAAAAAGGGTCTATCTGTCAATAA 742
                              ** **********  *** **  ** **********  * ** *** ***


GPH23_iinumae_clone2          -CGAGCTCCTAACCGTCCATCTCCAATCTTGTCAGGGGTGATCCT-ACGC 793
GPH23_iinumae_clone5          -CGAGCTCCTAACCGTCCATCTCCAATCTTGTCAGGGGGGATCCT-ACGC 793
GPH23_mandshurica_clone3      -CGAGCTCCTAACCGTCCATCTCCAATCTTGTCAGGGGTGATCCT-ACGC 790
GPH23_ananassa_clone4         -CGAGCTCCTAACCGTCCATCTCCAATCTTGTCAGGGGTGATCCT-ACGC 789
GPH23_ananassa_clone3         ACGAGCTCCCTATCGTCCATCTCCAATCTTGTAAGGGGTGATCCTTACGC 792
                               ********  * ******************* ***** ****** ****


GPH23_iinumae_clone2          GTCCACTTGTCCTCTCCCAGTTCTAACTATGTACACGCTAGCTGCGGATT 843
GPH23_iinumae_clone5          GTCCACTTGTCCTCTCC-AGTTCTAACTATGTACACGCTAGCTGCGGATT 842
GPH23_mandshurica_clone3      GTCCACTTGTCCTCTCCCAGTTCTAACTGTGTAGACGCTAGCTGCGGATT 840
GPH23_ananassa_clone4         GTCCACTTGTCCTTTCCCAGTTCTAACTATGTAGACGCTAGCTGCGGATT 839
GPH23_ananassa_clone3         TTCCCTTTGTCCTCTCCCCCATCTA-CTAAGTAGACGCTAGCTGCGGATT 841
                               ***  ******* ***     **** **  *** ****************


GPH23_iinumae_clone2          GTTATTATGTTTTGGATAGAATAGAATACCTTTGCAAAATAGGAAGCTCC 893
GPH23_iinumae_clone5          GTTATTATGTTTTGGATAGAATAGAATACCTTTGCAAAATAGGAAGCTCC 892
GPH23_mandshurica_clone3      GTTATTATTTTTTGGATAG-----AATACCTTTGCAAAATAGGAAGCTCC 885
GPH23_ananassa_clone4         GTTATTATGTTTTGGATAG-----AATACCTTTGCAAAATAGGAAGCTCC 884
GPH23_ananassa_clone3         GTTATTATGTTTTGGATAG-----AATACCTTTGCAAAATTGGAAGCTCC 886
                              ******** **********     **************** *********


GPH23_iinumae_clone2          ------TCCTTGTTTTTCGGCAAGAGAAAGGCCAAAATATCTGACCATTC 937
GPH23_iinumae_clone5          ------TCCTTGTTTTTCGGCAAGAGAAAGGCCAAAATATCTGACCATTC 936
GPH23_mandshurica_clone3      ------TCCTTGTTTTTCGGCAAGAGAAAGGCCAAAATATCTGACCATTC 929
GPH23_ananassa_clone4         ------TCCTTGTTTTTCTGCAAGAGAAAGGCCAAAATATCTGACCATTC 928
GPH23_ananassa_clone3         AGCTCCTCCTTGTTTTTCGGCAAGAGAAAGGCCAAAATATCAGACCGTTC 936
                                    ************ ********************** **** ***

GPH23_iinumae_clone2          CGACGCCGGAGCTTCCTCAGAAAGCCGGTTCCGTCCGCAACATCGATGCC 987
GPH23_iinumae_clone5          CGACGCCGGAGCTTCCTCAGAAAGCCGGTTCCGTCCGCAACATCGATGCC 986
GPH23_mandshurica_clone3      CGACGCCGGAGCTTCCTCAGAAAGCCAGTACTATCCGCCACATCGATGCC 979
GPH23_ananassa_clone4         CGACGCCGGAGCTTCCTCAGAAAGCCAGTACCATCCGCAACATCGATGCC 978
GPH23_ananassa_clone3         CGACGCCGGAGCTTCCTCAGAAAGCC--TACCATCCGCAACATCGTTGCC 984
                              **************************  * *  ***** ****** ****

GPH23_iinumae_clone2          AGGCCTTGCGAGGTTTGCCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTT 1037
GPH23_iinumae_clone5          AGGCCTTGCGAGGTTTGCCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTT 1036
GPH23_mandshurica_clone3      AGGCCTTGCGAGGTTTGCCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTT 1029
GPH23_ananassa_clone4         AGGCCTTGCGAGGTTTGCCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTT 1028
GPH23_ananassa_clone3         AGGCCTTGCGAGGTTTGCCTCCGCTTCTTTGGATTGTGTTTTTCGTGGTA 1034
                              *************************************************

GPH23_iinumae_clone2          TAGGAGATTGTTGAACAAAAAAGAAAAACATAACATATGATGAATGAATT 1087
GPH23_iinumae_clone5          TAGGAGATTGTTGAACAAAAAAGAAAAACATAACATATGATGAATGAATT 1086
GPH23_mandshurica_clone3      TAGGAGATTGCTGAACAAAAAAGAAAAACATAACATATGATGAATGAATT 1079
GPH23_ananassa_clone4         TAGGAGATTGTTGAACAAAAAAGAAAAACATA---TATGATAAATGAATT 1075
GPH23_ananassa_clone3         TAGGAGATTGTTGAACAAAAAAGAAAAACATAACATATGATGAATGAATC 1084
                              ********** *********************   ****** ******* 

GPH23_iinumae_clone2          ATCAAATTAATTAACCAAGGTGACAATACAAGATGAGAACACCAAGGGTT 1137
GPH23_iinumae_clone5          ATCAAATTAATTAACCAAGGTGACAATACAAGATGAGAACACCAAGGGTT 1136
GPH23_mandshurica_clone3      ATCAAATTAATTAATCAAGGTGACAATACAAGATGAGAACACCAAAGATT 1129
GPH23_ananassa_clone4         ATCAAATTAATTAATCAAGGTGACGATACAAGATGAGAACACCAAAGGTT 1125
GPH23_ananassa_clone3         ATCAAATTAATTAATCAAGGTGAC-------------------------- 1108
                              ************** *********
```

```
GPH23_iinumae_clone2          CAATAGTGTGTACTCTCAAGCCTAATACTAACACAACAAAGAAAGATTTC 1187
GPH23_iinumae_clone5          CAATAGTGTGTACTCTCAAGCCTAATACTAACACAACAAAGAAAGATTTC 1186
GPH23_mandshurica_clone3      CAATAGTGTGTACTCTCAAGCCTAATACTAACACAACAAAGAAAGATT-C 1178
GPH23_ananassa_clone4         CAATAGTGTGTACTCTCAAGCCTAATACTAACACAACAAAGAAAGATT-C 1174
GPH23_ananassa_clone3         -------------------------------------AAAGAAAGATT-A 1120
                                                                   ***********


GPH23_iinumae_clone2          TATCCTTCCATTCCCAAATCAAAAACCACTACAATGTACCGTC------- 1230
GPH23_iinumae_clone5          TATCCTTCCATTCCCAAATCAAAAACCACTACAATGTACCGTC------- 1229
GPH23_mandshurica_clone3      TATCGTTCCATTCCCAAATCAAAAACCATTACAATAGACCGCC------- 1221
GPH23_ananassa_clone4         TATCCTTCCATTCCCAGATCAAAAACCACTCTAATGTACCGCCGGCAAAG 1224
GPH23_ananassa_clone3         TATCCTTCCATTCCTAAGTCAAAAACCATTACAATGTACCGCCGGCAAAA 1170
                              **** ********* *  ********** *  ***  **** *


GPH23_iinumae_clone2          ------TAATTGAATTGACATTGTAAGTGAGAGATAGTGTCACGAGCTGC 1274
GPH23_iinumae_clone5          ------TAATTGAATTGACATTGTAAGTGAGAGATAGTGTCACGAGCTGC 1273
GPH23_mandshurica_clone3      ----------------GGCA----AAGTG-----------------CTGC 1234
GPH23_ananassa_clone4         TGCTGCTAATAGAATTGACATTGTAAGTAGGGGATAGTGTCACGAGCAGC 1274
GPH23_ananassa_clone3         TGCTGCTAATAGAATTGACATTGTAAGTGGGGGATAGTGTCACGAGCTGC 1220
                                              * **    ****                  * **


GPH23_iinumae_clone2          ATTGGGAGATAGTGTCATCGTGACACTCTATGAAGGGG-ATGCTTAAGAG 1323
GPH23_iinumae_clone5          ATTGGGAGATAGTGTCATCGTGACACTCTATGAAGGTG-ATGCTTAAGAG 1322
GPH23_mandshurica_clone3      --TAGGAGGTAGTGTCATCGTGACACTCTATGAAGGTG-ATGCTTAAGAG 1281
GPH23_ananassa_clone4         TCTAGGAGGTAGTGTCATCGTGACACTTTATTGGGGTGGATGCTAAGGGG 1324
GPH23_ananassa_clone3         AATAGGAGGTAGTGTCATCGTGACACTCTA--GAGGTG-ATGCTTAAGAG 1267
                                * **** ****************** **    ** * ***** * * *


GPH23_iinumae_clone2          GGTCGC---ATCAATGACAA-ACATGAGGGCAAATA-GAAGGTCTACTGG 1368
GPH23_iinumae_clone5          GGTCGC---ATCAATGACAA-ACATGAGGGCAAATA-GAATGTCTACTGG 1367
GPH23_mandshurica_clone3      GGACAC---GTCAATAGCAA-GTATGAGGGAAAAAAAGGATGTTTACTGG 1327
GPH23_ananassa_clone4         GTTCAG---GTTATGAGCAA-GCATGGGGGGTAAGGGGGATATCTACTGG 1370
GPH23_ananassa_clone3         GGTCACAAGGTCAATGGCAAAGCATGAGGGTTAAAGAGGATGTTTACTGA 1317
                              *  *      * *    ***   *** ***  **   * *  * *****

GPH23_iinumae_clone2          CATGTCGAATGACAATGTCGTAATTAGTTAAGTGAAACTTATATTTCAAG 1418
GPH23_iinumae_clone5          CATGTCGAATGACAATGTCGTAATTAGTTAAGTGAAACTTATATTTCAAG 1417
GPH23_mandshurica_clone3      CATGTCGAATGATAATTTCGTAATTAGTTAAGTGAAACTTATATTTCAAG 1377
GPH23_ananassa_clone4         CATTTTGAATGACAATGTTGTAAATGA-------AAACTTATATTTCAAG 1413
GPH23_ananassa_clone3         CATGTTGAAAGACAATGTCGTAATTAGTT-----AAAGT--------GAA 1354
                              *** * *** ** *** * **** *         *** *         *

GPH23_iinumae_clone2          GTACTTTGAC-TTAGTATTTAGAAAAACTGTAACATCGAAAGGAGTTCAA 1467
GPH23_iinumae_clone5          GTACTTTGAC-TTAGTATTTAGAAAAACTGTAACATCGAAAGGAGTTCAA 1466
GPH23_mandshurica_clone3      GTACTTTGAC-CTAGTATTTGGAAAAACCGTAACATCTAAA--------- 1417
GPH23_ananassa_clone4         GTATTTTGAT-TTAATATTTAGAAAAACTGTAACATCAAAAGGGGTTCAA 1462
GPH23_ananassa_clone3         GTACTGTGAAGTTAGTATTTCGAAAAACTGTAACATCGGAAGGGGTTCAA 1404
                              *** * ***   ** ***** ******* ********  **

GPH23_iinumae_clone2          TACATTTGACGACATATTTTTATGAGGTTTCTATAAATTAGTTATGAGAG 1517
GPH23_iinumae_clone5          TACATTTGACGACATATTTTTATGAGGTTTCTATGAATTAGTTATGAGAG 1516
GPH23_mandshurica_clone3      TACACTTGACGACATATTTTTATGAGATTTCTATGAATTAGTTATGAGAG 1467
GPH23_ananassa_clone4         TACATTTGCCGACATATTTTTATGAGGTTTTTATGAATTAGTTATGAGAG 1512
GPH23_ananassa_clone3         TACATTTGACGACATATTTTTATGAAGTTATTAGGAATTAGTTACGAGAG 1454
                              **** *** ****************  **  **  ********* *****

GPH23_iinumae_clone2          ATGGTTTTC-CTTGGACTATTTTGATTTTG-ATGTTTCCTTAACACACAT 1565
GPH23_iinumae_clone5          ATGGTTTTC-CTTGGACTATTTTGATTTTG-ATGTTTCCTTAACACACAT 1564
GPH23_mandshurica_clone3      ATGGTTTTC-TTTAGACTATTTTGATTTTGGATGTTTCCTTAACACACAT 1516
GPH23_ananassa_clone4         ATGGTTTTC-CTTGGACTATTTAGATTTTG-ATGTTTCCTTAACACACAT 1560
GPH23_ananassa_clone3         ATGGTTTTTTCTTAGAATATTTTGATTTTG-ATGTTTCCTTGACACACAT 1503
                              ********   ** ** ***** ******* ********** ********
```

```
GPH23_iinumae_clone2          TATATTTA--------------------AAGTAATTTTCCGTATCACTTA 1595
GPH23_iinumae_clone5          TATATTTA--------------------AAGTAATTTTCTGTATCACTTA 1594
GPH23_mandshurica_clone3      TATATTTCTCCATTTTCTTTTTATGTATAAGTAATTTTCTGTATCACTTA 1566
GPH23_ananassa_clone4         TATATTTCTCCATTTTCTT-------GTAAGTAATTTTCTGTATCACTTA 1603
GPH23_ananassa_clone3         TATATTTCTCCATGTTCTTCT-ATGTATAAGTAATTTTCTGTATCACTTA 1552
                              *******                     *********** **********


GPH23_iinumae_clone2          AAAACATTTCTTACTCTTTCCAGAAGCATCTCCAAACATCTCC------- 1638
GPH23_iinumae_clone5          AAAACATTTCTTACTCTTTTCCAAACCATCTCCAAACATCTCC------- 1637
GPH23_mandshurica_clone3      AAAACATTTCTTACTCTTTCCAGAAGCATCTCTA----TCCCC------- 1605
GPH23_ananassa_clone4         AAAACATTTCTTACTCTTCCCAGAAACATCTCCAAACATCCCTA--AACC 1651
GPH23_ananassa_clone3         GAAACATTTCTTACTCTTTCCAGAAGCATCTCCAAACATCCCCCTAAACC 1602
                               *****************  *  ** ****** *    ** *


GPH23_iinumae_clone2          -------CTAA-ATGTCAATGTC-------AATAGATGAAAGATCAACCT 1673
GPH23_iinumae_clone5          -------CCAA-ATGTCCATGGCCA----TAATAGATGAAAGATCAACCT 1675
GPH23_mandshurica_clone3      -------CTAA-A-----------------------------------CC 1612
GPH23_ananassa_clone4         GATAGCTCTAACATGTCAATGTC-------AATAGATGAAAGATCAACCT 1694
GPH23_ananassa_clone3         AATAGCCCTAACATGTCAATGTCACATGTCAATAGATGAAAGATCAACCT 1652
                                     * ** *                                   *


GPH23_iinumae_clone2          AAATGGTACCATATGTCCATACATAAAAAGACCCAAAAAGA--------- 1714
GPH23_iinumae_clone5          AAATGGTACCATATGTCCATACATAAAAAGACCCAAAAAGA--------- 1716
GPH23_mandshurica_clone3      AAATACACTAACATGTCAATGTCCA-------CCCAAAAGA--------- 1646
GPH23_ananassa_clone4         AAATGGTACCATATGTCCATACATAAAAAGACCCAAAAAGA--------- 1735
GPH23_ananassa_clone3         AAATGGTACCATATGTCCATACATAAAAGGACCCAAAAAAATAAATAAAG 1702
                              ****      * ***** **    *       ** **** *


GPH23_iinumae_clone2          AA-TAAATAAGCACCTTCATTTTTAAGCGCCATAAAAAGTAGAGAAGAAT 1763
GPH23_iinumae_clone5          AA-TAAATAAGCACCTTCATTTTTAAGCGCCATAAAAAGTAGAGAAGAAT 1765
GPH23_mandshurica_clone3      AA-TAAATAAGCACCTTCATTTTTAAGCGCCATAAAAAATAGAGAAGAAT 1695
GPH23_ananassa_clone4         AAATAAATAAGCACCTTCATTTTTAAGCGCCATAAAAAGTAGAGAAGAAT 1785
GPH23_ananassa_clone3         AAATAAATATGCACCTTCATTTTTAAGCGCCAGAAAAAGTAGAGAAGAAT 1752
                              ** ****** ********************** ***** ***********

GPH23_iinumae_clone2          ATAAGGTTTGAAGTGATCAAGGGGATAAGCAGTTTAAGGTCGACTTGTTC 1813
GPH23_iinumae_clone5          ATAAGGTTTGAAGTGATCAAGGGGATAAGCAGTTTAAGGTCGACTTGTTC 1815
GPH23_mandshurica_clone3      ATAAGGTTTGAAGTGAACAAGGGGATAAGCAGTTTAAGGTCGACTTGTTC 1745
GPH23_ananassa_clone4         ACAAGGTTTGAAGTGATCAAGGGGATAAGCAGTTTAAGGTCGACTTGTTC 1835
GPH23_ananassa_clone3         ATAAGGTTTGAAGTGATCAAGGGGATAAGCAGTTTAAGGTCGACTTGTTT 1802
                              * ************** ********************************

GPH23_iinumae_clone2          GGAAACAATGCTAACCACCACCACTGCCAGTCTCAGCT---GCTCCTCCT 1860
GPH23_iinumae_clone5          GGAAACAATGCTAACCACCACCACTGCCAGTCTCAGCT---GCTCCTCCT 1862
GPH23_mandshurica_clone3      GGAAACAATGCTGACCACCACCACTGCCACTCTCAGTCTCAGCTCCTCCT 1795
GPH23_ananassa_clone4         GGAAACAATGCTAACCACCACCACTGCCACTCTCAGCT---GCTCCTCCT 1882
GPH23_ananassa_clone3         GGAAACAATGCTAACCACCACCACTGCCACTCACAGTCTCAGCTCCTCCT 1852
                              ************ **************** ** ***     *********

GPH23_iinumae_clone2          CCTCTGCTTCCCAACTCCCACCACTCTTCCACTCTCTATCACCAAACCCA 1910
GPH23_iinumae_clone5          CCTCTGCTTCCCAACTCCCACCACTCTTCCACTCTCTATCACCAAACCCA 1912
GPH23_mandshurica_clone3      CCTCTACTTCCCAACTCCCACCACTCTTCCACTCTCTAACACCAAACCCA 1845
GPH23_ananassa_clone4         CCTCTGCTTCCCAACTCCCACCACTCTTCCACTCTCTATCACCAAACCCA 1932
GPH23_ananassa_clone3         CCTCTGCTTCCCAACTCCCATCGCTCTTCCACTCTCTATCACAAAACCCA 1902
                              ***** ************** * *************** *** *******

GPH23_iinumae_clone2          ATCTCCCTCAGATTCTCCTCCACATTACAGCTAACCAAAACCAGAACCAG 1960
GPH23_iinumae_clone5          ATCTCCCTCAGATTCTCCTCCACATTACAGCTAACCAAAACCAGAACCAG 1962
GPH23_mandshurica_clone3      ATCTCCCTCAGATTCTCCTCCACATTACAGCTAACTAAAACCAGAAC--- 1892
GPH23_ananassa_clone4         ATCTCCCTCAGATTCTCCTCCACATTACAGCTAACCAAAACCAGAAC--- 1979
GPH23_ananassa_clone3         ATCTCCCTCAGATTCTCCTCCACATTAAAGCTAACCAAAACCAGAAC--- 1949
                              *************************** ******* ***********
```

```
GPH23_iinumae_clone2         AACCAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGCCAGCTCCCTG 2010
GPH23_iinumae_clone5         AACCAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGCCAGCTCCCTG 2012
GPH23_mandshurica_clone3     ---CAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGCCAGCTCCCTG 1939
GPH23_ananassa_clone4        ---CAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGCCAGCTCCCTG 2026
GPH23_ananassa_clone3        ---CAGACCAACCCTTAAAACTCTCACTCGCCAAAAATGCCAGCTCCCTG 1996
                                ***********************************************

GPH23_iinumae_clone2         CTCTGAGAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCC 2060
GPH23_iinumae_clone5         CTCTGAGAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCC 2062
GPH23_mandshurica_clone3     CTCTGAGAGTGTCTGCTAAATACGAAGCTTCCCCTGCCACAGCTGAGGCC 1989
GPH23_ananassa_clone4        CTCTGAGAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCC 2076
GPH23_ananassa_clone3        CTCTGAGAGTGTCTGCTAATTACGAAGCTGCCCCTGCCACGGCTGAGGCC 2046
                             ******************* ********* ********** *********

GPH23_iinumae_clone2         TCCACGGTGCCGTTGGAGATGAAGGCGTGGGTGTA 2095
GPH23_iinumae_clone5         TCCACGGTGCCGTTGGAGATGAAGGCGTGGGTGTA 2097
GPH23_mandshurica_clone3     TCCACGGTGCCGTCGGAGATGAAGGCGTGGGTGTA 2024
GPH23_ananassa_clone4        TCCACGGTGCCGTCGGAGATGAAGGCGTGGGTGTA 2111
GPH23_ananassa_clone3        TCCACAGTGCCGTCGGAGATGAAGGCGTGGGTGTA 2081
                             ***** ******* *********************
```

Gene Pairs Detected Through Prediction from Genomic Sequence

GPH10

```
GPH10_ananassa_clone2      GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGC 50
GPH10_ananassa_clone20     GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGC 50
GPH10_ananassa_clone18     GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGC 50
GPH10_ananassa_clone19     GGCTTCTTCTTGTCCGGCAGCCTCTTCAGCCACTCGTCCTCCGGCGCCGC 50
                           **************************************************

GPH10_ananassa_clone2      CGATACCTCCTCCGCCTCCGACGACTTCGAACACAGCGGAATCGCTAGCC 100
GPH10_ananassa_clone20     CGATACCTCCTCCGCCTCCGACGACTTCGAACACAGCGGAATCGCTAGCC 100
GPH10_ananassa_clone18     CGATACCTCCTCCGCCTCCGACGACTTCGAACACAGCGGAATCGCTAGCC 100
GPH10_ananassa_clone19     CGATACCTCCTCCGCCTCCGACGACTTCGAACACAGCGGAATCGCTAGCC 100
                           **************************************************

GPH10_ananassa_clone2      TCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGT 150
GPH10_ananassa_clone20     TCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGT 150
GPH10_ananassa_clone18     TCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGT 150
GPH10_ananassa_clone19     TCCTTATCGGAGACCGAACGAGCCGAAACGGCGTCGCTTTAGGCGAGAGT 150
                           **************************************************

GPH10_ananassa_clone2      GAATAGCGAACTGAGTAGTTTGGATTTGAGAAGAGGATGTAATTGGTAAC 200
GPH10_ananassa_clone20     GAATAGCGAACTGAGTAGTTTGGATTTGAGAAGAGGATGTAATTGGTAAC 200
GPH10_ananassa_clone18     GAATAGCGAACTGAGTGGTTTGGATTTGAGAAGAGGATGAAATTGGTAAC 200
GPH10_ananassa_clone19     GAATAGCGAACTGAGTGGTTTGGATTTGAGAAGAGGATGAAATTGGTAAC 200
                           **************** ********************** **********

GPH10_ananassa_clone2      GGAGAAGAAGACTGTCGACATTTTTGGAGAAAGCTTTCATCTTTGAAGTG 250
GPH10_ananassa_clone20     GGAGAAGAAGACTGTCGACATTTTTGGAGAAAGCTTTCGGCTTTGAAGTG 250
GPH10_ananassa_clone18     GGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAAGTG 250
GPH10_ananassa_clone19     GGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTTAGCTTTGAAGTG 250
                           ************************* ***********   **********

GPH10_ananassa_clone2      GAGTGTAGGATAATAACAAACTCGTTATC--------------------- 279
GPH10_ananassa_clone20     GAGTGTAGGATAATAACAAACTCGTGATTAAAAGACAGGATTAATGTCAG 300
GPH10_ananassa_clone18     GAGTGTAGGATAATAACAAACTCGTTATC--------------------- 279
GPH10_ananassa_clone19     GAGTGTAGGATAATAACAAACTCGTTATC--------------------- 279
                           ************************* **
```

```
GPH10_ananassa_clone2       --------------------------------------------------
GPH10_ananassa_clone20      TGAGGTTTGGTTGGTTAAGGTGTTAACTGATAAATTTAAGGTCATAGGTT 350
GPH10_ananassa_clone18      --------------------------------------------------
GPH10_ananassa_clone19      --------------------------------------------------


GPH10_ananassa_clone2       --------------------------------------TAAAAGGCAGGT 291
GPH10_ananassa_clone20      CAAACCTCACGACATATGTAGGGTGTATGAATTATTAATAAAAGACAAAT 400
GPH10_ananassa_clone18      --------------------------------------TAAAAGACAGGT 291
GPH10_ananassa_clone19      --------------------------------------TAAAAGACAGGT 291
                                                                  ****** **  *

GPH10_ananassa_clone2       TTAATATCAGCCGTTAGATCATATTACGGCCCTGATCACTCGACATATGT 341
GPH10_ananassa_clone20      TTAATATCAGCCGTTAGATCATATTACGGCC-TGATCACTCGACATATGT 449
GPH10_ananassa_clone18      TTAATATCGGCCGTTAGATCACATTACGGCCCTGATCACTCGACATATGT 341
GPH10_ananassa_clone19      TTAATATCAGCCGTTAGATCATATTACGGCCCTGATCACTCGACATATGT 341
                            ******** ************ ********* ******************

GPH10_ananassa_clone2       TGATATACGCCCAACTCAAATTCGATATATATTTTCGATATACAT----- 386
GPH10_ananassa_clone20      TGATATACGCCCAACTCAAATTCGATATATATTTTCGATATACGT----- 494
GPH10_ananassa_clone18      TGATATACGCCTAACTCAAATTCGATATATATTTTCGATATACATTTTTT 391
GPH10_ananassa_clone19      TGATATACGCCTAACTCAAATTCGATATATATTTTCGATATACATTTTTT 391
                            *********** ******************************* *

GPH10_ananassa_clone2       ----------------------------------------ATATTTTATT 396
GPH10_ananassa_clone20      ----------------------------------------ATATTTTATT 504
GPH10_ananassa_clone18      TTTTAAGTAACTAAATGACTATTCGATATATATTTTCGATATACATTTTT 441
GPH10_ananassa_clone19      TTTTAAGTAACTAAATGACTATTCGATATATATTTTCGATATACATTTTT 441
                                                                    ***  ** **

GPH10_ananassa_clone2       TTTTTAAAGTAACTAAATGACTATGTACATCGTTTAACAAAAGAAACAAT 446
GPH10_ananassa_clone20      TTTTTAAAATAATTAAATAACTATTTACGTTGTTTAACAAAAGAAACAAT 554
GPH10_ananassa_clone18      TTTTTAAAGTAACTAAATGACTATTTACGTCGGTTAATAAAAGAAACAAT 491
GPH10_ananassa_clone19      TTTTTAAAGTAACTAAATGACTATTTACGTCGGTTAATAAAAGAAACAAT 491
                            ******** *** ***** ***** *** * * **** ************

GPH10_ananassa_clone2       TGAAGTTAAATTAAGAGCACCATAACAGCTGAGAAAGAGTACGAGAACAA 496
GPH10_ananassa_clone20      TGAAGTTAAATTAAGAGCACCGTAACAGCTGAGCAAGAGTACGAGAACAA 604
GPH10_ananassa_clone18      TGAAGTTAAATTAAGAGCACCATGACAG---------AGTACGAGAACAA 532
GPH10_ananassa_clone19      TGAAGTTAAATTAAGAGCACCATGACAG---------AGTACGAGAACAA 532
                            ********************* * ****         *************

GPH10_ananassa_clone2       AAGTATGAGCTAAAACAA--------------------ATAGAGAAAATA 526
GPH10_ananassa_clone20      AAGTATGAGCTACATCATTTG------------TTCATATAGAGAAAATA 642
GPH10_ananassa_clone18      AAGTATGAGCTACATTGTTTGCTCGTCGGTTTGTTCATATGGAGAAAATG 582
GPH10_ananassa_clone19      AAGTATGAGCTACATTGTTTGCTCGTCGGTTTGTTCATATGGAGAAAATA 582
                            ************ *                        ** ********

GPH10_ananassa_clone2       TAGAGGCGATGTTGTAGAAATAATTGAACATTAGAAAATTAAATTACCTA 576
GPH10_ananassa_clone20      TAGAGGCGATGTTGTAGAAATAATTGAACATTAGAAAATTAAATTACCTA 692
GPH10_ananassa_clone18      TAGAGGCGATGTTGTAGAAATAATTGAACATTAGAAAATTAAATTACCTA 632
GPH10_ananassa_clone19      TAGAGGCGATGTTGTAGAAATAATAGAACATTAGAAAATTAAATTACCTA 632
                            ************************ *************************

GPH10_ananassa_clone2       AAAGCCGATGAGTAAAATAATAACGAACTCGTAACCTAAAAGCGGCTTCA 626
GPH10_ananassa_clone20      AAAGCCGATGAGTAAAATAATAACGAACTCGTAACCTAAAAGCGGCTTCA 742
GPH10_ananassa_clone18      AAAGCCGATGAGTAAAATAATAACAAACTCGTAACCTAAAAGCGGCTTCA 682
GPH10_ananassa_clone19      AAAGCCGATGAGTAAAATAATAACAAACTCGTAACCTAAAAGCGGCTTCA 682
                            ************************ *************************

GPH10_ananassa_clone2       TATCATCCGCTTGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAA 676
GPH10_ananassa_clone20      TATCATCCGCTTGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAA 792
GPH10_ananassa_clone18      TATCATCCACTGGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAA 732
GPH10_ananassa_clone19      TATCATCCACTGGATCATATATGCGGGTGTGATTCGAAAACCAAAGTTAA 732
                            ******** ** **************************************
```

```
GPH10_ananassa_clone2       CCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAACAAAACCC 726
GPH10_ananassa_clone20      CCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAATAAAACCC 842
GPH10_ananassa_clone18      CCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAACAAAACCC 782
GPH10_ananassa_clone19      CCCGCCAAAGCCTAATTCCCAATTTTCATTTCCCACCAAAAACAAAACCC 782
                            ****************************************** *******

GPH10_ananassa_clone2       ACACGACGCCGTTTTGCTCCAATCCCCC-----TTTCTTCTTCAACCCCA 771
GPH10_ananassa_clone20      ACACGACGCCGTTTTGCTCCAATCCCCC-----TTTCTTCTTCAACCCCA 887
GPH10_ananassa_clone18      ACACGACGCCGTTTTGCTCCAATCCCCCCCCCCTTTCTTCTTCAACCCCA 832
GPH10_ananassa_clone19      ACACGACGCCGTTTTGCTCCAATCCCCCCCCC-TTTCTTCTTCAACCCCA 831
                            ****************************     *****************

GPH10_ananassa_clone2       TAGTCGCCTC---AGCTCAGTTCCATTTGTCTCAGATGCGATGGCCTCCG 818
GPH10_ananassa_clone20      TAGTCGCCTC---AGCTCAGTTCCATTTGTCTCAGATGCGATGGCCTCCG 934
GPH10_ananassa_clone18      TAGTCGCCTCCTCAGCTCAGTTCCATTTGTCTCA--TGCGATGGCTTCCG 880
GPH10_ananassa_clone19      TAGTCGCCTCCTCAGCTCAGTTCCATTTGTCTCA--TGCGATGGCTTCCG 879
                            **********   *********************  ********* ****

GPH10_ananassa_clone2       ------------GCGACCCAATCTCCGACTACACCCAAACACATCGCATT 856
GPH10_ananassa_clone20      ------------GCGACCCAATCTCCGACTACACCCAAACACATCGCATT 972
GPH10_ananassa_clone18      ACTCGAATTCCGGCGACCCAATCTCCTCCTACACCCAAACCCATCGCATC 930
GPH10_ananassa_clone19      ACTCAAATTCCGGCGACCCAATCTCCTCCTACACCCAAACCCATCGCATC 929
                                        **************  ************ ******** 

GPH10_ananassa_clone2       GTCCTTCTAATCGACCTCAACCCACTCCTCCATCTCCAAGATCCAACCCA 906
GPH10_ananassa_clone20      GTCCTTCTAATCGACCTCAACCCACTCCTCCATCTCCAAGATCCAACCCA 1022
GPH10_ananassa_clone18      GTCCTTCTAATCGACCTCAACCCACTCCTCAATCTCCAAGATCCAACCCA 980
GPH10_ananassa_clone19      GTCCTTCTAATCGACCTCAACCCACTCCTCAATCTCCAAGATCCAACCCA 979
                            ****************************** *******************

GPH10_ananassa_clone2       ATTCCTCACCTCTGTCCTCTCCTCAATCAAAACCCTAACCTCCTTCCCTT 956
GPH10_ananassa_clone20      ATTCCTCACCTCTGTCCTCTCCTCAATCAAAACCCTAACCTCCTTCCCTT 1072
GPH10_ananassa_clone18      ATTCCTCACCCCTGTCCTCTCCTCAATCAAAACCCTAACCTCCTTCCCTT 1030
GPH10_ananassa_clone19      ATTCCTCACCCCTGTCCTCTCCTCAATCAAAACCCTAACCTCCTTCCCTT 1029
                            ********** ***************************************

GPH10_ananassa_clone2       CTCTCTCTTC----------GCCGTCAGGCCCTTCTTCTCGTCTCTCTCT 996
GPH10_ananassa_clone20      CTCTCTCTTCCTCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCTCT 1122
GPH10_ananassa_clone18      CTCTCTCTTCATCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCTCT 1080
GPH10_ananassa_clone19      CTCTCTCTTCATCTCTCTTCGCCGTCAGGCCCTTCTTCTCGTCTCTCTCT 1079
                            **********          ******************************

GPH10_ananassa_clone2       CCTCTCCTCTCC-----GCCTCCAAGCTCCCGTCTTCGTCTCTAACGATC 1041
GPH10_ananassa_clone20      CCTCTCCTCTCCTCTCCGCCTCCAAGCTCCCGTCTTCGTCTCTAACGATC 1172
GPH10_ananassa_clone18      CCTCTCCTCTCC-----GCCTCCAAGCTCCCGTCTTCGTCTCTAACGATC 1125
GPH10_ananassa_clone19      CCTCTCCTCTCC-----GCCTCCAAGCTCCCGTCTTCGTCTCTAACGATC 1124
                            ************     *********************************

GPH10_ananassa_clone2       TCTTTCAACTCGCCGGAAGACACATATCGATCCCTATCTCAAACCCTGGC 1091
GPH10_ananassa_clone20      TCTTTCAACTCGCCGGAAGACACATATCGATCCCTATCTCAAACCCTGGC 1222
GPH10_ananassa_clone18      TCTTTCAACTCGCCGGAAGACACTTATCGATCCCTATCTCAAACCCTGGC 1175
GPH10_ananassa_clone19      TCTTTCAACTCGCCGGAAGACACTTATCGATCCCTATCTCAAACCCTGGC 1174
                            *********************** **************************

GPH10_ananassa_clone2       GTCTCTCTCGTTTGACCGGAAGTTGACCGGGTCCGATTCGCCGCGGGGAA 1141
GPH10_ananassa_clone20      GTCTCTCTCGTTTGACCGGAAGTTGACCGGGTCCGATTCGCCGCGGGGAA 1272
GPH10_ananassa_clone18      GTCTCTCTCTTTTGACCGCAAGTTGGCCGGGTCCGATTCGCCGCGGGGAA 1225
GPH10_ananassa_clone19      GTCTCTCTCTTTTGACCGCAAGTTGGCCGGGTCCGATTCGCCGCGGGGAA 1224
                            ********* ******** ****** ************************

GPH10_ananassa_clone2       CGCTTGTTGCGGCTGCGATGCGGCAGCTGGTGCATGATTACGCTTGGGAG 1191
GPH10_ananassa_clone20      CGCTTGTTGCGGCTGCGATGCGGCAGCTGGTACATGATTACGCTTGGGAG 1322
GPH10_ananassa_clone18      CGCNTGTTGCGGCGGCGATGCGGCAGCTGGTGCATGATTACGCTTGGGAG 1275
GPH10_ananassa_clone19      CGCTTGTTGCGGCGGCGATGCGGCAGCTGGTGCATGATTACGCTTGGGAG 1274
                            *** ********* ***************** ******************
```

```
GPH10_ananassa_clone2      CAGGTGATCTGCGACGCCGTGGCGGCGGAGACAGATACGTTTTCGAATTG 1241
GPH10_ananassa_clone20     CAGGTGATCTGCGACGCCGTGGCGGCGGAGACAGGTACGTTTTCGAATTG 1372
GPH10_ananassa_clone18     CCGGTGATCTGCGACGCCGCGGCGGCGGAGACCGGTACGTTATCGAATTG 1325
GPH10_ananassa_clone19     CCGGTGATCTGCGACGCCGCGGCGGCGGAGACCGGTACGTTATCGAATTG 1324
                           * ***************** ************ * ****** ********

GPH10_ananassa_clone2      CTGTGGTTTGAGGTCTAATTTGGCTGTTGTGTTTTCACCGGCGTGTCAAT 1291
GPH10_ananassa_clone20     CTGTGGTTTGAGGTCTAATTTGGCTGTTGTGTTTTTACCGGCGTGTCAAT 1422
GPH10_ananassa_clone18     CTGTGGTTTGAGGTCTAATTTGGCTGTTGTGTTTTCACCGGCGTGTCAAT 1375
GPH10_ananassa_clone19     CTGTGGTTTGAGGTCTAATTTGGCTGTTGTGTTTTCACCGGCGTGTCAAT 1374
                           *********************************** **************

GPH10_ananassa_clone2      TTGTGAATGAGTTCTTGAATTGTGAG------------GGTTTGGAGGAT 1329
GPH10_ananassa_clone20     TTGTGAATGAGTTCTTGAATTGTGAGTTGAATTGTGAGGGTTTGGAGGAT 1472
GPH10_ananassa_clone18     TTGTGAATGAGTTCTTGAATTGTGAG------------GGTTTGGAGGAT 1413
GPH10_ananassa_clone19     TTGTGAATGAGTTCTTGAATTGTGAG------------GGTTTGGAGGAT 1412
                           **************************            ************

GPH10_ananassa_clone2      TTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGA 1379
GPH10_ananassa_clone20     TTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGA 1522
GPH10_ananassa_clone18     TTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGA 1463
GPH10_ananassa_clone19     TTCAATGTGTTTTGTGAGAGGTTTCGAGGGTTTTTCGAGAATGTGGATGA 1462
                           **************************************************

GPH10_ananassa_clone2      GGCATATGTGTATAGAGATATTCAATTGAGTCGGGTTGATGTGAGGTATG 1429
GPH10_ananassa_clone20     GGCATATGTGTATAGAGATATTCAATTGAGTTGGGTTGATGTGAGGTATG 1572
GPH10_ananassa_clone18     GGCATTTGTGTGTAGAGATATTCAATTGAGTTGGGTTGATGTGAGGTATG 1513
GPH10_ananassa_clone19     GGCATTTGTGTGTAGAGATATTCAATTGAGTTGGGTTGATGTGAGGTATG 1512
                           ***** ***** ******************* ******************

GPH10_ananassa_clone2      GATTCGATAGCGGTGAGGATGAGGTAGTTGGATTGAAATGTGGTGTTTTC 1479
GPH10_ananassa_clone20     GATTCGATAGCGGTGAGGATGAGGTAGTTGGATTGAAATGTGGTGTTTTC 1622
GPH10_ananassa_clone18     GATTCGATAGCGGTGAGGATGAGGTAGTTGGATTGAAATGTGGTGTTTTC 1563
GPH10_ananassa_clone19     GATTCGATAGCGGTGAGGATGAGGTAGTTGGATTGAAATGTGGTGTTTTC 1562
                           **************************************************

GPH10_ananassa_clone2      GAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGAT 1529
GPH10_ananassa_clone20     GAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGAT 1672
GPH10_ananassa_clone18     GAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGAT 1613
GPH10_ananassa_clone19     GAGAGGGGGGTTAGGAGTTTAGGGTGGGGGTTTTGCTCATCTGATTCGAT 1612
                           **************************************************

GPH10_ananassa_clone2      TGTGCTTGGTTCGGCTCTTGTTCCATTTGGTTTGATTTATCCAGAGATTG 1579
GPH10_ananassa_clone20     TGTGCTTGGTTCGGCTCTTGTTCCATTTGGTTTGATTTATCCAGAGATTG 1722
GPH10_ananassa_clone18     TGTGCTTGGTTCGGCTCTTGTTCCATTTGGTTTGATTTATCCAGAGATTG 1663
GPH10_ananassa_clone19     TGTGCTTGGTTCGGCTCTTGTTCCATTTGGTTTGATTTATCCAGAGATTG 1662
                           **************************************************

GPH10_ananassa_clone2      GGGTGTCATCTAGGATTTTCGGGTGTAATGATCGATATAAGAAGGTTAGA 1629
GPH10_ananassa_clone20     GGGTGTCATCTAGGATTTTCGGGTGTAATGATCGATATAAGAAGGTTAGA 1772
GPH10_ananassa_clone18     GGGTGTCATCTAGGATTTTCGGGTGTAATGATCGATATAAGAAGTTTAGA 1713
GPH10_ananassa_clone19     GGGTGTCATCTAGGATTTTCGGGTGTAATGATCGATATAAGAAGTTTAGA 1712
                           ******************************************** *****

GPH10_ananassa_clone2      GCGCATTTGAGTCTTGAGATATCGGATGTAAAGGGGATGCCTTTGGAGTG 1679
GPH10_ananassa_clone20     GCGCATTTGAGTCTTGAGATATCAGATGTAAAGGGGATGCCTTTGGAGTG 1822
GPH10_ananassa_clone18     GCGCATTTGAGTCTTGAGATATCGGATGGAAAGGGGATGCCTTTGGAGTG 1763
GPH10_ananassa_clone19     GCGCATTTGAGTCTTGAGATATCGGATGGAAAGGGGATGCCTTTGGAGTG 1762
                           *********************** **** *********************

GPH10_ananassa_clone2      CAAGTTTTGTGATCTTGAGTTGGCTGATTTGAAAATGTTGTGTAGGAGTA 1729
GPH10_ananassa_clone20     CAAGTTTTGTGATCTTGAGTTGGCTGATTTGAAAATGTTGTGTAGGAGTA 1872
GPH10_ananassa_clone18     CAAGTTTTGTGATCTTGAGTTGGCTGATTTGAAAATGTTGTGTAGGAGTA 1813
GPH10_ananassa_clone19     CAAGTTTTGTGATCTTGAGTTGGCTGATTTGAAAATGTTGTGTAGGAGTA 1812
                           **************************************************
```

```
GPH10_ananassa_clone2       GAGGTGATGATCGCTTGTTTTCGGTGGAAGGCATGAGCTCGCAGACAAGA 1779
GPH10_ananassa_clone20      GAGGTGATGATCGCTTGTTTTCGGTGGAAGGCATGAACTCGCAGACAAGA 1922
GPH10_ananassa_clone18      GAGGTGATGATGGCTTGTTTTCGGTGGAAGGCATGAACTCGCAGACAAGA 1863
GPH10_ananassa_clone19      GAGGTGATGATGGCTTGTTTTCGGTGGAAGGCATGAACTCGCAGACAAGA 1862
                            *********** ************************ *************

GPH10_ananassa_clone2       GGTCATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAATGGAGTGTC 1829
GPH10_ananassa_clone20      GGTCATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAATGGAGTGTC 1972
GPH10_ananassa_clone18      GGTCATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAACGGAGTGTT 1913
GPH10_ananassa_clone19      GGTCATGAGGTGAAGAGGCTGTTTTGGGGAAGTGTTGGCAACGGAGTGTT 1912
                            ***************************************** *******

GPH10_ananassa_clone2       GAAGATTCAGGTTAAGGCTTTGCAGAAGGATAGTGAGTTTGGGAAATTTA 1879
GPH10_ananassa_clone20      GAAGATTCAGGTTAAGGCTTTGCAGAAGGATAGTGAGTTTGGGAAATTTA 2022
GPH10_ananassa_clone18      GAAGATTCAGGTTAAGGCTTTGCAGAAGGATAGTGAGTTTGGGAAATTTA 1963
GPH10_ananassa_clone19      GAAGATTCAGGTTAAGGCTTTGCAGAAGGATAGTGAGTTTGGGAAATTTA 1962
                            **************************************************

GPH10_ananassa_clone2       AGGGGGAATTGTCGGATCTGATTCTGGTCTATGAAGTTTCAGGAAAAGAT 1929
GPH10_ananassa_clone20      AGGGGGAATTGTCGGATCTGATTCTGGTCTATGAAGTTTCAGGAAAAGAT 2072
GPH10_ananassa_clone18      AGGGGGAATTGTCGGATCCGATTCTGGTCTATGAAGTTTCAGGAAAAGAT 2013
GPH10_ananassa_clone19      AGGGGGAATTGTCGGATCCGATTCTGGTCTATGAAGTTTCAGGAAAAGAT 2012
                            ****************** *******************************

GPH10_ananassa_clone2       GGAAAAGAAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAATGCT 1979
GPH10_ananassa_clone20      GGAAAAGAAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAGTGCT 2122
GPH10_ananassa_clone18      GGAAAAGAAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAATGCT 2063
GPH10_ananassa_clone19      GGAAAAGAAGTTTCTGGTGGTTTGTTTGTAGATAAGGTTCTTGAAATGCT 2062
                            ********************************************* ****

GPH10_ananassa_clone2       ATCA-GTGGAATTGGGTGAGTTTGTACCGAGGAAATTGCCACCTGTTTGG 2028
GPH10_ananassa_clone20      ATCAAGTGAAATTGGGTGAGTTTGTACCGAGGAAATTGCCACCTGTTTGG 2172
GPH10_ananassa_clone18      ATCAAGTGGAATTGGGTGAGTTTGTACCAAGGAAATTGCCACCTGTTTGG 2113
GPH10_ananassa_clone19      ATCA-GTGGAATTGGGTGAGTTTGTACCAAGGAAATTGCCACCTGTTTGG 2111
                            **** *** ******************* *********************

GPH10_ananassa_clone2       CAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTAGTTTC 2078
GPH10_ananassa_clone20      CAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTAGTTTC 2222
GPH10_ananassa_clone18      CAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTAGTGTC 2163
GPH10_ananassa_clone19      CAGATTCTCTTGAGTTTTATATACAGGGAGGGTTGCTGGGCATTAGTGTC 2161
                            *********************************************** **

GPH10_ananassa_clone2       TATTTCAAATGATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTA 2128
GPH10_ananassa_clone20      TATTTCAAATGATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTA 2272
GPH10_ananassa_clone18      TATTTCAAATGATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTA 2213
GPH10_ananassa_clone19      TATTTCAAATGATAGTGGTGTATCACATACTGGAATCCTTAAGCCTTTTA 2211
                            **************************************************

GPH10_ananassa_clone2       CAGTTTCTTCAGCTCTTATTTTTGTTATGGATGAAGGAATTCACCCTCAT 2178
GPH10_ananassa_clone20      CAGTTTCTTCAGCTCTTATTTTTGTTATGGATGAAGGAATTCACCCTCAT 2322
GPH10_ananassa_clone18      CAGTTTCTTCAGCTCTTATTTTTGTTATGGATGAAGGAATTCACCCTCAT 2263
GPH10_ananassa_clone19      CAGTTTCTTCAGCTCTTATTTTTGTTATGGATGAAGGAATTCACCCTCAT 2261
                            **************************************************

GPH10_ananassa_clone2       AAAAAAGGGCATGGCATTGGTGCAGTGAATAAGGGTCAGTCTCGTCCAAA 2228
GPH10_ananassa_clone20      AAAAAAGGGCATGGCATTGGTGCAGAGAATAAGGGTCAGTCTCGTCCAAA 2372
GPH10_ananassa_clone18      AAAAAAGGGCATGTCATTGGTGCAGTGAATAAGGGTCAGTCTCGTCCAAA 2313
GPH10_ananassa_clone19      AAAAAAGGGCATGTCATTGGTGCAGTGAATAAGGGTCAGTCTCGTCCAAA 2311
                            ************* *********** ************************

GPH10_ananassa_clone2       GATGAAGAATGAGATGTGCAAACCTGATGCTGATTTGAACGACTTTTGTG 2278
GPH10_ananassa_clone20      GATGAAGAATGAGATGTGCAAACCTGATGCTGATTTGAACGACTTTTGTG 2422
GPH10_ananassa_clone18      GATGAAGAATGAGATGTGCAAACCTGATGCTGATTTGAACGACTTTTGTG 2363
GPH10_ananassa_clone19      GATGAAGAATGAGATGTGCAAACCTGATGCTGATTTGAACGACTTTTGTG 2361
                            **************************************************
```

```
GPH10_ananassa_clone2       GGTCGCAAACTGGGCCTTCACCATCTAATAAGCATTCTGCTGAGATTGAT 2328
GPH10_ananassa_clone20      GGTCGCAAACTGGGCCTTCACCATCTAATAAGCATTCTGCTGAGATTGAT 2472
GPH10_ananassa_clone18      GGTCGCAAACTGGGCCTTCACCATCTAATAAGCATTCTGCTGAGATTGAT 2413
GPH10_ananassa_clone19      GGTCGCAAACTGGGCCTTCACCATCTAATAAGCATTCTGCTGAGATTGAT 2411
                            **************************************************

GPH10_ananassa_clone2       GGAAAGAAAAAAAGTAGCAAAAGAAGTTCACATTCACTCAAAGATCTCAC 2378
GPH10_ananassa_clone20      GGAAAGAAAAAAAGTAGCGAAAGAAGTTCACATTCACTCAAAGATCTCAC 2522
GPH10_ananassa_clone18      GGAAAGAAAAAAAGTAGCAAAAGAAGTTCACATTCACTCAAAGATCTCAC 2463
GPH10_ananassa_clone19      GGAAAGAAAAAAAGTAGCAAAAGAAGTTCACATTCACTCAAAGATCTCAC 2461
                            ****************** *******************************

GPH10_ananassa_clone2       CTGGAGTTCTTTCTGTAAGGCAGCATTCGAATTTTCAGACTTACATTTGG 2428
GPH10_ananassa_clone20      CCGGAGTTCTTTCTGTAAGGCAGCATTCGAATTTTCAGACTTACATTTGG 2572
GPH10_ananassa_clone18      CTGGAGTTCTTTCTGTAAGGCAGCATTCGAATTTTCAGACTTACATTTGG 2513
GPH10_ananassa_clone19      CTGGAGTTCTTTCTGTAAGGCAGCATTCGAATTTTCAGACTTACATTTGG 2511
                            * ************************************************

GPH10_ananassa_clone2       AAGAGGTTTACTTTGCCAGGCAACGTAGCAGCTCAAAAAAGTTGAAATTT 2478
GPH10_ananassa_clone20      AAGAGGTTTACTTTGCCAGGCAACGTAGCAGCTCAAAAAAGTTGAAATTT 2622
GPH10_ananassa_clone18      AAGAGGCTTACTTTGCCAGGCAACGTAGCAGCTCAAAAAAGTTGAAATTT 2563
GPH10_ananassa_clone19      AAGAGGTTTACTTTGCCAGGCAACGTAGCAGCTCAAAAAAGTTGAAATTT 2561
                            ****** *******************************************

GPH10_ananassa_clone2       CTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGA 2528
GPH10_ananassa_clone20      CTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGA 2672
GPH10_ananassa_clone18      CTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGA 2613
GPH10_ananassa_clone19      CTAAAATGCTGGATGAAACAGATTAAAAAACTGAAGTATCCAATAACGGA 2611
                            **************************************************

GPH10_ananassa_clone2       GGAATCTAAGGTGCACCAGGAAAAACAAAAGGAGATGAGCAATAGGTTGG 2578
GPH10_ananassa_clone20      GGAGTCTAAGGTGCACCAGGAAAAACAAAAGGAGATGAGCAATAGGTTGG 2722
GPH10_ananassa_clone18      GGAGTCTAAGGTGCACCAGGAAAAACAAAAGGAGATGAGCAATAGGTTGG 2663
GPH10_ananassa_clone19      GGAGTCTAAGGTGCACCAGGAAAAACAAAAGGGGATGAGCAATAGGTTGG 2661
                            *** **************************** *****************

GPH10_ananassa_clone2       ATTTGTTGCACCAAGAGAGCGAACAGCCAATGTCGTCATCTGGTTCAGCT 2628
GPH10_ananassa_clone20      ATTTGTTGCACCAAGAGAGCGAACAGCCAATGTCGTCATCTGGTTCAGCT 2772
GPH10_ananassa_clone18      ATTTGTTGCACCAAGAGAGCGAACAGCCAATGTCATCATCTGGTTCAGCT 2713
GPH10_ananassa_clone19      ATTTGTTGCACCAAGAGAGCGAACAGCCAATGTCATCATCTGGTTCAGCT 2711
                            ********************************** ***************

GPH10_ananassa_clone2       GGAGAAATTTCTTTCCCTGTGGCCTTTGGAGTACAGGATGAAGCTGCTCA 2678
GPH10_ananassa_clone20      GGAGAAATTTCTTTCCCTGTCGCCTTTGGAGTACAGGATGAAGCTGCTCA 2822
GPH10_ananassa_clone18      GGAGAAATTTCTTTCTCTGCGGCCTTTGGAGTACAGGATGAAGCTGCTCA 2763
GPH10_ananassa_clone19      GGAGAAATTTCTTTCTCTGCGGCCTTTGGAGTACAGGATGAAGCTGCTCA 2761
                            *************** ***  *****************************

GPH10_ananassa_clone2       GGAACATAGATTACAAACCTCAGAAGATTTTTTCTGTAATTTCTCTGATA 2728
GPH10_ananassa_clone20      GGAACATAGATTACAAACCTCAGAAGATTTTTTCTGTAATTTCTCTGATA 2872
GPH10_ananassa_clone18      GGAACATAGATTACAAACCTCAGAAGATTTTTTCTGTAATTTCTCTGATA 2813
GPH10_ananassa_clone19      GGAACATAGATTACAAACCTCAGAAGATTTTTTCTGTAATTTCTCTGATA 2811
                            **************************************************

GPH10_ananassa_clone2       AGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTGGGGGCATTCGCA 2778
GPH10_ananassa_clone20      AGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTGGGGGCATTCACA 2922
GPH10_ananassa_clone18      AGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTGGGGGCATTCGCA 2863
GPH10_ananassa_clone19      AGATCCAACAAGGGCTAGAATCTGAAGTAGTAGACTTGGGGGCATTCGCA 2861
                            *********************************************** **

GPH10_ananassa_clone2       CATCGGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAACATAGCAC 2828
GPH10_ananassa_clone20      CATCGGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAACATAGCAC 2972
GPH10_ananassa_clone18      CATCGGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAGCATAGCTC 2913
GPH10_ananassa_clone19      CATCGGCTTTTGAGTCAATCAATATATTTTTTGACTCAAAAGCATAGCTC 2911
                            ***************************************** ****** *
```

```
GPH10_ananassa_clone2       AACAACCCCTTCAGAAGATCAAACTCCTGTAAAATATGACAATCTTGATG 2878
GPH10_ananassa_clone20      AACAACCCCTTCAGAAGATCAAACTCCTGTAAAATCTGACAATCTTGATG 3022
GPH10_ananassa_clone18      AACAACCCCTTCAGAAGATCAAACTCCTGTAAAATCTGACAATCTTGATG 2963
GPH10_ananassa_clone19      AACAGCCCCTTCAGAAGATCAAACTCCTGTAAAATCTGACAATCTTGATG 2961
                            **** ****************************** **************

GPH10_ananassa_clone2       ATTTGGTTACTGCTGAGCTGTTAAAACTTTTACTCAGAGATCCCAAGGAT 2928
GPH10_ananassa_clone20      ATTTGGTTACTGCTGAGCTGTTAAAACTTTTACTCAGAGATCCCAAGGAT 3072
GPH10_ananassa_clone18      ATTTGGTTACTGCTGAGCTGTTAAAACTTT-ACTCAGAGATCCCAAGGAT 3012
GPH10_ananassa_clone19      ATTTGGTTACTGCTGAGCTGTTAAAACTTTTACTCAGAGATCCCAAGGAT 3011
                            ****************************** *******************

GPH10_ananassa_clone2       ATGGTTGCCAGGCACAAAAGCTATGATTCATCTTCTCAAGCATCTGATCC 2978
GPH10_ananassa_clone20      ATGGTTGCCAGGCACAAAAGCTATGATTCATCTTCTCAAGCATCTGATCC 3122
GPH10_ananassa_clone18      ATGGTTGCCAGGCACAAAAGCTATGATCCATCTTCTCAAGCATCTGATCC 3062
GPH10_ananassa_clone19      ATGGTTGCCAGGCACAAAAGCTATGATCCATCTTCTCAAGCATCTGATCC 3061
                            *************************** **********************

GPH10_ananassa_clone2       TGGATGTGAAGGCTTTACTTCAGAAATAATAGTTCGAGAGTATCCTTTCA 3028
GPH10_ananassa_clone20      TGGATGTGAAGGCTTTACTTCAGAAATAATAGTTCGAGAGTATCCTTTCA 3172
GPH10_ananassa_clone18      TGGATGTGATGGCTTTACTTCAGAAATAATAGTTCGAGAGTATCCTTTCA 3112
GPH10_ananassa_clone19      TGGATGTGATGGCTTTACTTCAGAAATAATAGTTCGAGAGTATCCTTTCA 3111
                            ********* ****************************************

GPH10_ananassa_clone2       TTTCTCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAATTCT 3078
GPH10_ananassa_clone20      TTTATCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAATTCT 3222
GPH10_ananassa_clone18      TTTATCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAATTCT 3162
GPH10_ananassa_clone19      TTTATCAGTTGATCGTTTTATTTTCTTTTATACTATGCATAATCAATTCT 3161
                            *** **********************************************

GPH10_ananassa_clone2       ACTTTAATGCTATGTAAACTTTGCCCCTTGTTAGTGTTACACTTTTCCTT 3128
GPH10_ananassa_clone20      ACTTTAATGCTATGTAAACTTTGCCCCTTGTTACTGTTACACTTTTCCTT 3272
GPH10_ananassa_clone18      ACTTTAATGCTATGTAAACTTTGCCCCCTGTTACTGTTACACT--TCCTT 3210
GPH10_ananassa_clone19      ACTTTAATGCTATGTAAACTTTGCCCCCTGTTACTGTTACACT--TCCTT 3209
                            *************************** ***** *********  *****

GPH10_ananassa_clone2       CACTAGCACAAAGATATGAATTACAGATACTTTTCCGGATGGAGATTTTA 3178
GPH10_ananassa_clone20      CACTAGCACAAAGATATGAATTACAGATACTTTTCCGGATGGAGATTTTA 3322
GPH10_ananassa_clone18      CACTAGCACAAAGATATGAATTACAGATACTTTTCCGGATGGAGATTTTA 3260
GPH10_ananassa_clone19      CACTAGCACAAAGATATGAATTACAGATACTTTTCCGGATGGAGATTTTA 3259
                            **************************************************

GPH10_ananassa_clone2       CAATCAGAAGTTGGAGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGT 3228
GPH10_ananassa_clone20      CAATCAGAAGTTGGAGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGT 3372
GPH10_ananassa_clone18      CAATCAGAAGTTGGAGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGT 3310
GPH10_ananassa_clone19      CAATCAGAAGTTGGGGCAAGTATCAAAGATGCTGTGAAACAGAAGTTTGT 3309
                            ************** ***********************************

GPH10_ananassa_clone2       GAAACATATTTGCACGCTTTTGGAGACCATTCG-TGCTCGGTGTCATCTG 3277
GPH10_ananassa_clone20      GAAACATATTTGCACGCTTTTGGAGACCATTCG-TGCTCGGTGTCATCTG 3421
GPH10_ananassa_clone18      GAAACATATTTGCACGCTTTTGGAGACCATTCG-TGCTCAGTGTCATCTG 3359
GPH10_ananassa_clone19      GAAACATATTTGCACGCTTTTGGAGACCATTCGGTGCTCAGTGTCATCTG 3359
                            ********************************* ***** **********

GPH10_ananassa_clone2       GAGGGAGGCTTCTTTGGTGACTGGACCCTAGAAAATTATGCTGGAAAGAT 3327
GPH10_ananassa_clone20      GAGGGAGGCTTCTTTGGTGACTGGACCCTAGAAAATTATGCTGGAAAGAT 3471
GPH10_ananassa_clone18      GAGGGAGGCTTCTTTGGTGACTGGACCCTAGAAAATTATGCTGGAAAGAT 3409
GPH10_ananassa_clone19      GAGGGAGGCTTCTTTGGTGACTGGACCCTAGAAAATTATGCTGGAAAGAT 3409
                            **************************************************

GPH10_ananassa_clone2       TATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAAC 3377
GPH10_ananassa_clone20      TATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAAC 3521
GPH10_ananassa_clone18      TATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAAC 3459
GPH10_ananassa_clone19      TATAAAAAGCAGGTAGATGAGTCACATGTATAAATCTAATTACCCATAAC 3459
                            **************************************************
```

```
GPH10_ananassa_clone2       TATTATTTTCTAATGAAATT-GTATTCATGAACACTGAAATGGTAGATAC 3426
GPH10_ananassa_clone20      TATTATTTTCTAATGAAATTTGTATTCATGAACACTGAAATGGTAGATAC 3571
GPH10_ananassa_clone18      TATTATTTTCTAATGAAATTTGTATTCATGAACACTGAAATGGTAGATAC 3509
GPH10_ananassa_clone19      TATTATTTTCTAATGAAATTTGTATTCATGAACACTGAAATGGTAGATAC 3509
                            ******************** *****************************

GPH10_ananassa_clone2       TCAGTTATTTACAATGAAACTCCAATATATGTTTATGGTTTGCCTGTTAA 3476
GPH10_ananassa_clone20      TCAGTTATTTACAATGAAACTCCAATATATGTTTATGGTTTGCCTGTTAA 3621
GPH10_ananassa_clone18      TCAGTTATTTACAATGAAACTCCAGTATATGTTTATGTTTTGCCTGTTAA 3559
GPH10_ananassa_clone19      TCAGTTATTTACAATGAAACTCCAATATATGTTTATGTTTTGCCTGTTAA 3559
                            ************************ ************* ************

GPH10_ananassa_clone2       TGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTAT 3526
GPH10_ananassa_clone20      TGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTAT 3671
GPH10_ananassa_clone18      TGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTAT 3609
GPH10_ananassa_clone19      TGATACTTTTATCAGTACTTCGATGAAACATATAGTGTTGAAACAATTAT 3609
                            **************************************************

GPH10_ananassa_clone2       GTGATTGATTTGTATGCCCTCCCAAAAGGCCTTTGGGGGTAGTATGAAGA 3576
GPH10_ananassa_clone20      GTGATTGATTTGTATGCCCTCCCAAAAGGCCTTTGGGGGTAGTATGAAGA 3721
GPH10_ananassa_clone18      GTGATTGATTTGTATGCCCTCCCAAATGGCCTTTGGGGGTAGTAAGAAGA 3659
GPH10_ananassa_clone19      GTGATTGATTTGTATGCCCTCCCAAATGGCCTTTGGGGGTAGTAAGAAGA 3659
                            ************************** ***************** *****

GPH10_ananassa_clone2       AGGGAGACATTGACAGTCAAAAATATTATCTCCTTATTTTACGTACAAAA 3626
GPH10_ananassa_clone20      AGGGAGACATTGACCGTCAAAACTATTATCTCCTTATTTTACGTACAAAA 3771
GPH10_ananassa_clone18      AGGGAGACATTGACAGTCAAAAATATTATCTCCTTATTTTACGTACAAAA 3709
GPH10_ananassa_clone19      AGGGAGACATTGACAGTCAAAAATATTATCTCCTTATTTTACGTACAAAA 3709
                            ************** ******* ***************************

GPH10_ananassa_clone2       TTGATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAA 3676
GPH10_ananassa_clone20      TTGATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAA 3821
GPH10_ananassa_clone18      TTGATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAA 3759
GPH10_ananassa_clone19      TTGATGACTCCTCATCAGGCTGTTGAAGGCAGGGTTGACAGAGAACAGAA 3759
                            **************************************************

GPH10_ananassa_clone2       AAGCTAAATACCTCCTTTGCATAATTTCATATGACTTAAGTGACTTTCCT 3726
GPH10_ananassa_clone20      AAGCTAAATACCTCCTTTGCATAATTTCATATGACTTAAGTGACTTTCCT 3871
GPH10_ananassa_clone18      AAGCTAAATACCTCCTTTGCATAATTTCATATGACTTAAGTGACTTTCCT 3809
GPH10_ananassa_clone19      AAGCTAAATACCTCCTTTGCATAATTTCATATGACTTAAGTGACTTTCCT 3809
                            **************************************************

GPH10_ananassa_clone2       TATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGTATGCATACA 3776
GPH10_ananassa_clone20      TATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGTATGCATACA 3921
GPH10_ananassa_clone18      TATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGTATTCATACA 3859
GPH10_ananassa_clone19      TATTAATCTAGATTTGCAACCTTGTTTTTCTGACACTATGTATTCATACA 3859
                            ******************************************* ******

GPH10_ananassa_clone2       ACTTTTGCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTC 3826
GPH10_ananassa_clone20      ACTTTTGCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTC 3971
GPH10_ananassa_clone18      ACTTTTGCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTC 3909
GPH10_ananassa_clone19      ACTTTTGCAATTGTATTCTGTATGTTGCAATAGTTCATTCCTTTGTTTTC 3909
                            **************************************************

GPH10_ananassa_clone2       CAGACCCAAAAAAAACTGCCCAAATTTATGTGAAAACACTGCATTTATGT 3876
GPH10_ananassa_clone20      CAGACCCAAAAAAAACTGCCCAAATTTATGTGAACACACTGCATTTATGT 4021
GPH10_ananassa_clone18      CAGACCCAAAAAAAACGGCCCAAATTTATGTGAAAACACTGCATTTATGT 3959
GPH10_ananassa_clone19      CAGACCCAAAAAAAACGGCCCAAATTTATGTGAAAACACTGCATTTATGT 3959
                            **************** ***************** ***************

GPH10_ananassa_clone2       TTGAAGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCAAATTT 3926
GPH10_ananassa_clone20      TTGAAG---TAGGATTAGGCAGGTAGACTGATGATTCAATTCCCAAATTT 4068
GPH10_ananassa_clone18      TTGAAGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCAAATTT 4009
GPH10_ananassa_clone19      TTGAAGAAGTAGGATTAGGCAGGTAGACTGATGATTCAATTCCCAAATTT 4009
                            ******   *****************************************
```

```
GPH10_ananassa_clone2       TCAGGTACTGTCAGACTCTTGAAGACGTGGTTCATAAAATCTACACAAAA 3976
GPH10_ananassa_clone20      TCAGGTACTGTCAGACTCTTGAAGACGTGGTTCATAAAATCTACACAAAA 4118
GPH10_ananassa_clone18      TCAGGTACTGTCAGACTCTTGAAGACGTGGTTCATAAAATCTACACAAAA 4059
GPH10_ananassa_clone19      TCAGGTACTGTCAGACTCTTGAAGACGTGGTTCATAAAATCTACACAAAA 4059
                            **************************************************

GPH10_ananassa_clone2       ATGGATTTGTTACTGTTTGATGATGAGGAAGAACTCCCTAATAATTTATT 4026
GPH10_ananassa_clone20      ATGGATTTGTTACTGTTTGATGATGAGGAAGAACTCCCTAATAATTTATC 4168
GPH10_ananassa_clone18      ATGGATTTGTTACTGTTTGATGATGAGGAAGAACTCCCTAATAATGTATT 4109
GPH10_ananassa_clone19      ATGGATTTGTTACTGTTTGATGATGAGGAAGAACTCCCTAATAATGTATT 4109
                            ********************************************* ***

GPH10_ananassa_clone2       CAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGATG 4076
GPH10_ananassa_clone20      CAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGATG 4218
GPH10_ananassa_clone18      CAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGATG 4159
GPH10_ananassa_clone19      CAACAGCGAGGATAGCAGTCATTCATACAAAGAAAAACCAGGGAAAGATG 4159
                            **************************************************

GPH10_ananassa_clone2       AGGTGGGTGAAAATAGTAGAATGAAGAAATTGGTATCAGCAGAAGATGAA 4126
GPH10_ananassa_clone20      AGGTGGGTGAAAATAGTAGAATGAAGAAATTGGTATCAGCAGAAGATGAA 4268
GPH10_ananassa_clone18      AGGTGGGTGAAAATAGTAGAATGAAGAAATTGGTATCAGCAGAAGATGAA 4209
GPH10_ananassa_clone19      AGGTGGGTGAAAATAGTAGAATGAAGAAATTGGTATCAGCAGAAGATGAA 4209
                            **************************************************

GPH10_ananassa_clone2       TCCCCTGATCCACAGAAACATTACAATGGAAGGCCAAGTGCTCAAGTACT 4176
GPH10_ananassa_clone20      TCCCCTGATCCACAGAAACATTACAATGGAAGGCCAAGTGCTCAAGTAGT 4318
GPH10_ananassa_clone18      TCCCCTGATCCACAGAAACATTACAATGGAAGGCCAAGTGCTCAAGTAGT 4259
GPH10_ananassa_clone19      TCCCCTGATCCACAGAAGCATTACAATGGAAGGCCAAGTGCTCAAGTAGT 4259
                            ***************** ****************************** *

GPH10_ananassa_clone2       TAAACAAGAAGAGCATGCTCGCAAGTTGATGAAAGCTCAAGAGAGTAGAG 4226
GPH10_ananassa_clone20      TAAACAAGAAGAGCATGCTCGCAAGTTGATGAAAGCTCAAGAGAGTAGAG 4368
GPH10_ananassa_clone18      TAAACAAGAAGAGCATGCTCGCAAGTTGATGGAAGCTCAAGAGAGTAGAG 4309
GPH10_ananassa_clone19      TAAACAAGAAGAGCATGCTCGCAAGTTGATGGAAGCTCAAGAGAGTAGAG 4309
                            ******************************* ******************

GPH10_ananassa_clone2       AGAGGGCTTGGAGAATTGCTTCTTTCACAAGTCGGGTAGCTGATTTGCAG 4276
GPH10_ananassa_clone20      AGAGGGCTAGGAGAATTGCTTCTTTCACAAGTCGGGTAGCTGATTTGCAG 4418
GPH10_ananassa_clone18      AGAGAGCTAGGAGAATTGCTTCTTTTACAAGTCGGGTAGCTGATTTGCAG 4359
GPH10_ananassa_clone19      AGAGAGCTAGGAGAATTGCTTCTTTTACAAGTCGGGTAGCTGATTTGCAG 4359
                            **** *** ***************** ************************

GPH10_ananassa_clone2       CGAG 4280
GPH10_ananassa_clone20      CGAG 4422
GPH10_ananassa_clone18      CGAG 4363
GPH10_ananassa_clone19      CGAG 4363
                            ****
```

Polymorphic fragments of GPH10: clones from octoploid *versus* clones from diploids. SSRs detected are magenta-colored.

```
10PPR1AB22_nubicola              AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAA 50
10PPR1AB22_mandshurica           AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAA 50
10PPR1AB22_vesca                 AACGGAGAAGAAGACTGTCGACATTTTAAGAGAAAGCTTTCAGCTTTGAA 50
10PPR1AB22_viridis               AACGGAGAAGAAGACTGTCGACATTTCTAGAGAAAGCTTTCAGCTTTGAA 50
10PPR1AB22_nilgerrensis          AACGGAGAAGAAGACTGTCGACATATCTAGAGAAAGCTTTCAGCTTTGAA 50
10PPR1AB22_ananassa_clone18      AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAA 50
10PPR1AB22_ananassa_clone20      AACGGAGAAGAAGACTGTCGACATTTTTGGAGAAAGCTTTCGGCTTTGAA 50
10PPR1AB22_ananassa_clone19      AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTTAGCTTTGAA 50
10PPR1AB22_ananassa_clone2       AACGGAGAAGAAGACTGTCGACATTTTTGGAGAAAGCTTTCATCTTTGAA 50
10PPR1AB22_iinumae               AACGGAGAAGAAGACTGTCGACATTTTTAGAGAAAGCTTTCAGCTTTGAA 50
                                 ************************ *   ***********   *******
```

```
10PPR1AB22_nubicola              GT-----G-----TAGGATAATAACAAAGAAACTCGTTATCTGAAAGACA 90
10PPR1AB22_mandshurica           GT-----G-----TAGGATAATAACAAAGAAACTCGTTATCTGAAAGACA 90
10PPR1AB22_vesca                 GT-----G-----TAGGATAATAACAAAGAAACTCGTTATCTGAAAGACA 90
10PPR1AB22_viridis               GT-----G-----TAGGATAATAACAAAGAAACTCGTTATCTGAAAGACA 90
10PPR1AB22_nilgerrensis          GT-----GGAGTGTAGGATAATAAC----AAACTCGTTATCTGAAAGACA 91
10PPR1AB22_ananassa_clone18      GT-----GGAGTGTAGGATAATAAC----AAACTCGTTATCTAAAAGACA 91
10PPR1AB22_ananassa_clone20      GT-----GGAGTGTAGGATAATAAC----AAACTCGTGAT-TAAAAGACA 90
10PPR1AB22_ananassa_clone19      GT-----GGAGTGTAGGATAATAAC----AAACTCGTTATCTAAAAGACA 91
10PPR1AB22_ananassa_clone2       GT-----GGAGTGTAGGATAATAAC----AAACTCGTTATCTAAAAGGCA 91
10PPR1AB22_iinumae               CTTTGAAGTAGTGTAGGATAATAAC----AAACTCGTTATCTAAAAGACA 96
                                  *     *     ************    ******** ** * **** **

10PPR1AB22_nubicola              GGTTTAATATCAGC-------------------CGTTGGATCATA---TT 118
10PPR1AB22_mandshurica           GGTTTAATATCAGC-------------------CGTTGGATCATA---TT 118
10PPR1AB22_vesca                 AGTTTAATATCAGC-------------------CGTTGGATCATA---TT 118
10PPR1AB22_viridis               AGTTTAATACCAGC-------------------CGTTGGATCATA---TT 118
10PPR1AB22_nilgerrensis          GGTTTAATATCAGC-------------------CGTTGGATTATA---TT 119
10PPR1AB22_ananassa_clone18      GGTTTAATATCGGC-------------------CGTTAGATCACA---TT 119
10PPR1AB22_ananassa_clone20      GGATTAATGTCAGTGAGGTTTGGTTGGTTAAGGTGTTAACTGATAAATTT 140
10PPR1AB22_ananassa_clone19      GGTTTAATATCAGC-------------------CGTTAGATCATA---TT 119
10PPR1AB22_ananassa_clone2       GGTTTAATATCAGC-------------------CGTTAGATCATA---TT 119
10PPR1AB22_iinumae               GGTTTAATATCAGC-------------------CGTTAGATCCTA---TT 124
                                  * *****  * *                     ***   *   *

10PPR1AB22_nubicola              ACGGCCCTG-------ATCGCTCGACATA--------------------- 140
10PPR1AB22_mandshurica           ACGGCCCTG-------ATCGCTCGACATA--------------------- 140
10PPR1AB22_vesca                 ACGGCCCTG-------ATCGCTCGACATA--------------------- 140
10PPR1AB22_viridis               ACTGCCCTG-------ATCGCTCGACATA--------------------- 140
10PPR1AB22_nilgerrensis          CCGGCCCTG-------ATCTCTCGACATA--------------------- 141
10PPR1AB22_ananassa_clone18      ACGGCCCTG-------ATCACTCGACATATGTTGA-TATACGCCTAACT- 160
10PPR1AB22_ananassa_clone20      AAGGTCATAGGTTCAAACCTCACGACATATGTAGGGTGTATGAATTATTA 190
10PPR1AB22_ananassa_clone19      ACGGCCCTG-------ATCACTCGACATATGTTGA-TATACGCCTAACT- 160
10PPR1AB22_ananassa_clone2       ACGGCCCTG-------ATCACT---------------------------- - 134
10PPR1AB22_iinumae               ---------------------------------------------------

10PPR1AB22_nubicola              -------------------------------------------------A 141
10PPR1AB22_mandshurica           -------------------------------------------------A 141
10PPR1AB22_vesca                 -------------------------------------------------A 141
10PPR1AB22_viridis               -------------------------------------------------A 141
10PPR1AB22_nilgerrensis          -------------------------------------------------T 142
10PPR1AB22_ananassa_clone18      --------CAAATTCGATAT-------------ATATT------------ 177
10PPR1AB22_ananassa_clone20      ATAAAAGACAAATTTAATATCAGCCGTTAGATCATATTACGGCCTGATCA 240
10PPR1AB22_ananassa_clone19      --------CAAATTCGATAT-------------ATATT------------ 177
10PPR1AB22_ananassa_clone2       --------------------------------------------------
10PPR1AB22_iinumae               --------------------------------------------------

10PPR1AB22_nubicola              TTCGATATATATATATA------TTATTTTTTTCTAAA----------AA 175
10PPR1AB22_mandshurica           TTCGATATATATATA--------TTATTTTTTTCTAAA----------AA 173
10PPR1AB22_vesca                 TTCGATATATATATATATATATTTTTTTTTTTTCTAAA----------AA 181
10PPR1AB22_viridis               TTCGATATATATATATATA----TTATTTTTTTCTAAA----------AA 177
10PPR1AB22_nilgerrensis          GTTGATATACGC---------------------CTGAC----------TC 161
10PPR1AB22_ananassa_clone18      TTCGATATACA------TTTTTTTTTTAAGTAACTAAATGACTATTCGAT 221
10PPR1AB22_ananassa_clone20      CTCGACATATGTTGATATAC-------GCCCAACTCAA-----ATTCGAT 278
10PPR1AB22_ananassa_clone19      TTCGATATACA------TTTTTTTTTTAAGTAACTAAATGACTATTCGAT 221
10PPR1AB22_ananassa_clone2       --CGACATATGTTGATATAC-------GCCCAACTCAA-----ATTCGAT 170
10PPR1AB22_iinumae               --------------------------------------------------
```

```
10PPR1AB22_nubicola             AAAAAA--TCGATATACAGTATATT----TTTTTTGAATTAATTAAATAA 219
10PPR1AB22_mandshurica          AAAAAAAATCGATATACAGTATATT---TTTTTTTGAATTAATTAAATGA 220
10PPR1AB22_vesca                AAAAAA--TCGATATACAGTATATT----TTTTTTGAATTAATTAAATGA 225
10PPR1AB22_viridis              AAATAAA-TCGATATACAGTATATT---TTTTTTTGAAGTAATTAAATGA 223
10PPR1AB22_nilgerrensis         AAAT----TCTATATACA------------TTTTCGAAAGAATT--TTTG 193
10PPR1AB22_ananassa_clone18     ATATATTTTCGATATACA-TTT------TTTTTTTAAAGTAACTAAATGA 264
10PPR1AB22_ananassa_clone20     ATATATTTTCGATATACG-TATATTTTATTTTTTTAAAATAATTAAATAA 327
10PPR1AB22_ananassa_clone19     ATATATTTTCGATATACA-TTT------TTTTTTTAAAGTAACTAAATGA 264
10PPR1AB22_ananassa_clone2      ATATATTTTCGATATACA-TATATTTTATTTTTTTAAAGTAACTAAATGA 219
10PPR1AB22_iinumae              --------------------------------------------------


10PPR1AB22_nubicola             GTATTTAGATCGCTTAAAAAGA-TAAACAATTGAAGTTGGTTTAGAAGCA 268
10PPR1AB22_mandshurica          GTATTTAGATCGCTTAAAAAAAATAAACAATCGAAGTTGAATTAGGAGCA 270
10PPR1AB22_vesca                GTATTTAGATCGCTTAAAAAGA-TAAACAATCGAAATTGGTTTAAGAACA 274
10PPR1AB22_viridis              TTATTTAAATCGCTTAAAAAGA-TAAACAA--GAAGTTGGTTTAGGAGCA 270
10PPR1AB22_nilgerrensis         TTGTTGAAGTAACT--AAATGACTATACGATTGAAGATAGATTAAGAGAA 241
10PPR1AB22_ananassa_clone18     CTATTTACGTCGGTTAATAAAA-GAAACAATTGAAGTTAAATTAAGAGCA 313
10PPR1AB22_ananassa_clone20     CTATTTACGTTGTTTAACAAAA-GAAACAATTGAAGTTAAATTAAGAGCA 376
10PPR1AB22_ananassa_clone19     CTATTTACGTCGGTTAATAAAA-GAAACAATTGAAGTTAAATTAAGAGCA 313
10PPR1AB22_ananassa_clone2      CTATGTACATCGTTTAACAAAA-GAAACAATTGAAGTTAAATTAAGAGCA 268
10PPR1AB22_iinumae              --------------------------------------------------


10PPR1AB22_nubicola             TCATAGGAGC------AAGAGTACGAGAACAAAAGTATGAGCTACACTGT 312
10PPR1AB22_mandshurica          CCATAGGAGC------AAGAGTATGAGAACAAAAGTATGAGCTACATTGT 314
10PPR1AB22_vesca                CCATAGGAGC------AAGAGTATGAGAACAAAAGTATGAGCTACACTGT 318
10PPR1AB22_viridis              CCATAGGAGC------AAGAGTATGAGAACAAAAGTATGAGCCACACTGT 314
10PPR1AB22_nilgerrensis         ACATAGCAACTGAGTAAAAAGTATGAGAACAAAAGTATGAGCTACATTGT 291
10PPR1AB22_ananassa_clone18     CCATGACAG----------AGTACGAGAACAAAAGTATGAGCTACATTGT 353
10PPR1AB22_ananassa_clone20     CCGTAACAGCTGAGC-AAGAGTACGAGAACAAAAGTATGAGCTACATCAT 425
10PPR1AB22_ananassa_clone19     CCATGACAG----------AGTACGAGAACAAAAGTATGAGCTACATTGT 353
10PPR1AB22_ananassa_clone2      CCATAACAGCTGAGA-AAGAGTACGAGAACAAAAGTATGAGCTAAAACAA 317
10PPR1AB22_iinumae              --------------------------------------------------


10PPR1AB22_nubicola             TTGCTCGTCGGTTTATTTATATGGAGAAAATATCAAGGTGATGTTGTATA 362
10PPR1AB22_mandshurica          TTGCTCGTCGGTTTATTTATATGGAGAAAATATCAAGGTGATGTTGTATA 364
10PPR1AB22_vesca                TTGCTCATCGGTTTATTTATATGGAGAAAATATCAAGGTGATGTTGTATA 368
10PPR1AB22_viridis              TTGCTCTTCGGTTTGTTTATACAGAGAAAATATAAAGTGATGTTGTAGA 364
10PPR1AB22_nilgerrensis         TTGCTCCTCGGTCTGTTTATATGGAGAAAATATAAATGTGATGTTGTAAA 341
10PPR1AB22_ananassa_clone18     TTGCTCGTCGGTTTGTTCATATGGAGAAAATGTAGAGGCGATGTTGTAGA 403
10PPR1AB22_ananassa_clone20     ------------TTGTTCATATAGAGAAAATATAGAGGCGATGTTGTAGA 463
10PPR1AB22_ananassa_clone19     TTGCTCGTCGGTTTGTTCATATGGAGAAAATATAGAGGCGATGTTGTAGA 403
10PPR1AB22_ananassa_clone2      --------------------ATAGAGAAAATATAGAGGCGATGTTGTAGA 347
10PPR1AB22_iinumae              --------------------------------------------------


10PPR1AB22_nubicola             AACAATTAAACATTACAAAA--TCAAATTACTTAACAATGAACCATCTTC 410
10PPR1AB22_mandshurica          AACAATTAAACATTACAAAA--TCAAATTACTTAACAATAAACCATCTTC 412
10PPR1AB22_vesca                AACAATTAAACATTACAAAA--TCAAATTACCTAACAATGAACCATTTTC 416
10PPR1AB22_viridis              AACAATTGAACACTAAAAAA--TCAAATTACCTAACAACGAACCATCTTC 412
10PPR1AB22_nilgerrensis         AATAATTGAACATTAAAAAAAATTAAATTATCTAATAACGAATCATCTTT 391
10PPR1AB22_ananassa_clone18     AATAATTGAACATTAGAAAA--TTAAATTACCTAA--------------- 436
10PPR1AB22_ananassa_clone20     AATAATTGAACATTAGAAAA--TTAAATTACCTAA--------------- 496
10PPR1AB22_ananassa_clone19     AATAATAGAACATTAGAAAA--TTAAATTACCTAA--------------- 436
10PPR1AB22_ananassa_clone2      AATAATTGAACATTAGAAAA--TTAAATTACCTAA--------------- 380
10PPR1AB22_iinumae              --------------------------------------------------
```

```
10PPR1AB22_nubicola              AGACATGT----AAAATCAGAAAGTTAAAAGGTTCGAGTCGCATATGAGT 456
10PPR1AB22_mandshurica           AGACATGT----AAAATCAAAAAGTTAAAAGGTTCGAGTCGCATATGAGT 458
10PPR1AB22_vesca                 AGACATGT----AAAATCATAAAATTAAAAGGTTCGAGTCGCATATGAGT 462
10PPR1AB22_viridis               AGACATAC----AAGATCAGAAAGTTAAGAGGTTCGAGTCGCACATGAGT 458
10PPR1AB22_nilgerrensis          AGACGTACGTACAAAATCAGAGAGTTAAGAGATTCGAGTTGC-------- 433
10PPR1AB22_ananassa_clone18      --------------------------------------------------
10PPR1AB22_ananassa_clone20      --------------------------------------------------
10PPR1AB22_ananassa_clone19      --------------------------------------------------
10PPR1AB22_ananassa_clone2       --------------------------------------------------
10PPR1AB22_iinumae               --------------------------------------------------


10PPR1AB22_nubicola              TTGTCGAGCTGATCAAATACCACAGTTTACTTGACTGAACAAACTTACGT 506
10PPR1AB22_mandshurica           TTGTCGAGCTGATCAAATACCACAGTTTACTTGACTGAACAAACTTACGT 508
10PPR1AB22_vesca                 TTGTCGAGCTGATCAAATACCACAGTTTACTTGACTGAACAAACTTACGT 512
10PPR1AB22_viridis               TTCTTGAGGCGATCAAATACCACAGTTTACTTGACTCAACAACTTTACGC 508
10PPR1AB22_nilgerrensis          ------------TCAAATACCATATTTTACTTGACTTAACAAT------- 464
10PPR1AB22_ananassa_clone18      --------------------------------------------------
10PPR1AB22_ananassa_clone20      --------------------------------------------------
10PPR1AB22_ananassa_clone19      --------------------------------------------------
10PPR1AB22_ananassa_clone2       --------------------------------------------------
10PPR1AB22_iinumae               --------------------------------------------------


10PPR1AB22_nubicola              A-ACGAGTCAAACGAGCTAAAAACGAGTCGAATAAAAATCGGGCACCATC 555
10PPR1AB22_mandshurica           A-ACGAGTCAAACGAGCTAAAAACGAGTCGAATAAAAATCGGGCACCATC 557
10PPR1AB22_vesca                 A-ACGAGTCAAACGAGCTAAAAACGAGTCGAATAAAAATCGGGCACCATC 561
10PPR1AB22_viridis               ATACGAGTCAAACGAGCTAAAAACGAGTCGAATAAAAATCGGGCACCATC 558
10PPR1AB22_nilgerrensis          --ACGAGTCAAACGAGCTAAAAACGAGTCGATTAAA-------------- 498
10PPR1AB22_ananassa_clone18      --------------------------------------------------
10PPR1AB22_ananassa_clone20      --------------------------------------------------
10PPR1AB22_ananassa_clone19      --------------------------------------------------
10PPR1AB22_ananassa_clone2       --------------------------------------------------
10PPR1AB22_iinumae               --------------------------------------------------


10PPR1AB22_nubicola              TATATCGAGACTATGTAAGAGCCGAGGAGTAAAA--TAATAACAAACTCG 603
10PPR1AB22_mandshurica           AATATCGAGACTATGTAAAAGCCGAGGAGTAAAA--TAATAACAAACTCG 605
10PPR1AB22_vesca                 AATATCGAGACTATGTAAGAGCCGAGGAGTAAAA--TAATAACAAACTCG 609
10PPR1AB22_viridis               AATATCGAGACTATGTAAGAGCCGAGGAGTAAAAAATAATAACAAACTCG 608
10PPR1AB22_nilgerrensis          -------------TCTAAGAGCCGAGGAGTAAAG--TAATAACAAACTCG 533
10PPR1AB22_ananassa_clone18      ------------------AAGCCGATGAGTAAAA--TAATAACAAACTCG 466
10PPR1AB22_ananassa_clone20      ------------------AAGCCGATGAGTAAAA--TAATAACGAACTCG 526
10PPR1AB22_ananassa_clone19      ------------------AAGCCGATGAGTAAAA--TAATAACAAACTCG 466
10PPR1AB22_ananassa_clone2       ------------------AAGCCGATGAGTAAAA--TAATAACGAACTCG 410
10PPR1AB22_iinumae               ----------------AAGAGCCGAGGAGTAAAA--TAATAACAAAGTCG 156
                                                     ***** *******   ******* ** ***


10PPR1AB22_nubicola              TTATCTAAAAGACAGGTTTAATATCAGCCCTTGGACCATATGTACGGGTG 653
10PPR1AB22_mandshurica           TTATCTAAAAGACAGGTTTAATATCAGCCGTTGGACCATATGTACAGGTG 655
10PPR1AB22_vesca                 TTATCTAAAAGACAGGTTTAATATCAGCCGTTGGACCATATGTACAGGTG 659
10PPR1AB22_viridis               TTATCTAAAAGACAGGTTTAATATCAGCCGTTGGACCATATGTACAGGTG 658
10PPR1AB22_nilgerrensis          TTATCTAAAATACAGGTTTAATATCAGCCGTTGGATCATATATACAGGTG 583
10PPR1AB22_ananassa_clone18      TAACCTAAAAGC--GGCTTCATATCATCCACTGGATCATATATGCGGGTG 514
10PPR1AB22_ananassa_clone20      TAACCTAAAAGC--GGCTTCATATCATCCGCTTGATCATATATGCGGGTG 574
10PPR1AB22_ananassa_clone19      TAACCTAAAAGC--GGCTTCATATCATCCACTGGATCATATATGCGGGTG 514
10PPR1AB22_ananassa_clone2       TAACCTAAAAGC--GGCTTCATATCATCCGCTTGATCATATATGCGGGTG 458
10PPR1AB22_iinumae               TAACCTAAAAGC--GGCTTCATATCATCTACTGGATCATATATGCGGGTG 204
                                 * * ******    ** ** ****** *   * ** ***** * * ****
```

```
10PPR1AB22_nubicola           TGATTCGAAAACCGAAGTTAACCCGCCAAACCCTCCTTCC-AATTTTCAT 702
10PPR1AB22_mandshurica        TGATTCGAAAACCGAAGTTAACCCGCCAAACCCTCATTCCCAATTTTCAT 705
10PPR1AB22_vesca              TGATTCGAAAACCGAAGTTAACCCGCCAAACCCTCATTCCCAATTTTCAT 709
10PPR1AB22_viridis            TGATTCGAAAACCGA-GTTAACCCGCCAAACCCTCCTTCCCAATTTT-AT 706
10PPR1AB22_nilgerrensis       TGATTCGAAAACCGAAGTTAACCCGCCAAACCCTCATTCCCAATTTTCAT 633
10PPR1AB22_ananassa_clone18   TGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCAT 564
10PPR1AB22_ananassa_clone20   TGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCAT 624
10PPR1AB22_ananassa_clone19   TGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCAT 564
10PPR1AB22_ananassa_clone2    TGATTCGAAAACCAAAGTTAACCCGCCAAAGCCTAATTCCCAATTTTCAT 508
10PPR1AB22_iinumae            TGATTCGAAAACCAAAGTTAACCCGCCAAA--------CCCAATTTTCAT 246
                              ************* * **************        ** ****** **

10PPR1AB22_nubicola           TTCCCACCAAAAACAAAACC 722
10PPR1AB22_mandshurica        T-CCCACCAAAAACAAAACC 724
10PPR1AB22_vesca              T-CCCACCAAAAACAAAACC 728
10PPR1AB22_viridis            TTCCCACCAAAAACAAAACC 726
10PPR1AB22_nilgerrensis       T-CCCACCAAAAACAAAACC 652
10PPR1AB22_ananassa_clone18   TTCCCACCAAAAACAAAACC 584
10PPR1AB22_ananassa_clone20   TTCCCACCAAAAATAAAAC- 643
10PPR1AB22_ananassa_clone19   TTCCCACCAAAAACAAAACC 584
10PPR1AB22_ananassa_clone2    TTCCCACCAAAAACAAAACC 528
10PPR1AB22_iinumae            TTCCCACCAAAAACAAAACC 266
                              * *********** *****
```

11D02

```
11D02_viridis       GAGCTGCTGTGTGA-CCAAATGGGTACAGAAGAAGCCNGTTTGCCAAACCTACCCATGAT 59
11D02_nubicola      GAGCTGCTGTGTGA-CCAAATGG-TACAGA-GAAGCCNGTTTGCCAAACCTACCCATGAT 57
11D02_vesca         GAGCTGCTGTGTGAACCAAATGG-TACAGA-GAAGCCG-TTTGCCAAACCTACCCATGAT 57
11D02_iinumae       GAGCTGCTGTGTGAACCAAATGG-TACAGA-GAAGCCG-TTTGCCAAACCTACCCATGAT 57
                    ************** ******** ****** ******  *********************

11D02_viridis       CCAATCAA-TGCATAAACTTTAAGAACTCAAATACCAAACGATCAAACATAATGACTGAA 118
11D02_nubicola      CCAATCAAATGCATAAACTTTAAGAACTCAAATACCAAACGATCAAACATAATGACTGAA 117
11D02_vesca         CCAATCAAATGCATAAACTTTAAGAACTCAAATACCAAACGATCAAACATAATGACTGAA 117
11D02_iinumae       CCAATCAAATGCATAAACTTTAAGAACTCAAATACCAAACGATCAAACATAATGACTGCA 117
                    ******** ************************************************* *

11D02_viridis       ATGAACAAAAATCAAATGAGCAAAGACTAAATGAGAAAACAGACCTCTTTGTAAACTGGG 178
11D02_nubicola      ATGAACAAAAATCAAATGGGCAAAGACTAAATGAGAAAACAGACCTCTTTGTAAACTGGG 177
11D02_vesca         ATGAACAAAAATCAAATGGGCAAAGACTAAATGAGAAAACAGACCTCTTTGTAAACTGGG 177
11D02_iinumae       ATGAACAAAAATCAAATGGGCAAAGACTAAATGAGAAAACAGACCTCTTTGTAAACTGGG 177
                    ****************** *****************************************

11D02_viridis       TTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCTGCCATTCCTGTT 238
11D02_nubicola      TTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCAGCCATTCCTGTT 237
11D02_vesca         TTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCAGCCATTCCTGTT 237
11D02_iinumae       TTTTGGGGTTTAAAGCCATGGGCACCATATGAGCCCAACCCAAGAGCTGCCATTCCTGTT 237
                    *********************************************** ************

11D02_viridis       TATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCA----AACACAA 294
11D02_nubicola      TATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCAGACAAACACAA 297
11D02_vesca         TATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCAGACAAACACAA 297
11D02_iinumae       TATCACAAAATCACAAATTGGGTCCTCTCAGATTGATGCAAAATCACCAGACAAACACAA 297
                    *************************************************    *******

11D02_viridis       TTTGATACAGAAGTCTTCACACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAA 354
11D02_nubicola      TTTCATACAGAAGTCTTCCCACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAA 357
11D02_vesca         TTTCATACAGAAGTCTTCCCACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAA 357
11D02_iinumae       TTTGATACAGAAGTCTTCACACAGAGAATATGACATTTGTAATTAAACACAGAATAAAAA 357
                    *** ************** *****************************************
```

```
11D02_viridis       TGATAACTTTTCAATAGTATAAGGAGAAGATGAGGACAGTACCAGAGACTGCAGCTACTT 414
11D02_nubicola      TGATAACTTTTCAATAGTATAAGAAGGAGATGAGGACAGTACCAGAGACTGCAGCTACTT 417
11D02_vesca         TGATAACTTTTCAATAGTATAAGAAGGAGATGAGGACAGTACCAGAGACTGCAGCTACTT 417
11D02_iinumae       TGATAACTTTTCAATACTATAAGAAGGAGATGAGGACAGTACCAGAGACTGCAGCTACTT 417
                    **************** ****** ** *********************************

11D02_viridis       TGTGCCATAGCATAGGATTCATTGCTGTATTCTTCCCCTAACTTGGTTTCCTTTCAGTGT 474
11D02_nubicola      TGTGCCACAGCATAGGATTCATTGCTGTATTCTTCCCTTTGCTTGGTTTCCTTTCAGCCT 477
11D02_vesca         TGTGCCACAGCATAGGATTCATTGCTGTATTCTTCCCTTTGCTTGGTTTCCTTTCAGTCT 477
11D02_iinumae       TGTGCCATAGCATAGGATTCATTGCTGTATTCTTCCCTTAACTTGGTTTCC--------- 468
                    ******* ***************************** *  **********

11D02_viridis       CTTCGACCTTCTTCTAAAACGACGGAGTCGGTGAAACTGTGCAAGTCTTCTTGTGA---- 530
11D02_nubicola      CTTCGACTTTCTTCTAAAACGACGTAGTCGGTGCAACTGTGCAAGTCTTCTTGTGATGCA 537
11D02_vesca         CTTCGACTTTCTTCTAAAACGACGTAGTCGGTGCAACTGTGCAAGTCTTCTTGTGATGCA 537
11D02_iinumae       --------------------GACGGAGTCGGTGCAACTGTGCAAGTCTTCTTGTGATGCA 508
                                        **** ******** **********************

11D02_viridis       ATTTTCTTTTCTAGGTGATTTTTTTTTTCTTTTATAATTAATTTGGTTTTATTTTTCCAA 590
11D02_nubicola      ATTTTCTTTTCTAGGTGATTTTTTTT--CTTTTATAATTAATTTGGTTTTATTTTTCCAA 595
11D02_vesca         ATTTTCTTTTCTAGGTGATTTTTTTT--CTTTTATAATTAATTTGGTTTTATTTTTCCAA 595
11D02_iinumae       ATTTTCTTTTCTAGGTGTTTTTTTTT--CTTTTATAATTAATTTGGTTTTATTTTTCCAA 566
                    ***************** ********  ********************************

11D02_viridis       ATAATACCTGAAAGACTTTTTTTTTTTTTTTTTGATAGAAATACCTAAAAGACTTCATAA 650
11D02_nubicola      ATAATACCTGAAAGACTTTTTTTTC--------GATAG---------------------- 625
11D02_vesca         ATAATACCTGAAAGACTTTTTTTTT-------CGATAGGA-------------------- 628
11D02_iinumae       ATAATACCTGAAAGACTTTTTTTTTTT-----CGATAGAAATACCTGTTAAGACT----- 616
                    ************************         *****

11D02_viridis       AAGCTGTTAAGGCTTCATTTAGGATTGCAGTAATTTTTTTTGGACAGTATTACGGGACAC 710
11D02_nubicola      ----------------------GATTGCAGTAATTTTTTTTGGACAGTATTACGGGACAC 663
11D02_vesca         ------------------------TTGCAGTAATTTTTTTTGGACAGTATTACGGGACAC 664
11D02_iinumae       -------TAAGACTTCATTTAGTATTGCAGTAATTTTTTT-GGACAGTATTACGGGACAC 668
                                            **************** *******************

11D02_viridis       TGTGACAGCTT--GAGTTTGAATCTTAGGTGGGATGATTTAAGTATCTTAGTTGAATGGA 768
11D02_nubicola      TGTGACAGCTTTAGAGTTTGAATCTTAGGTTGGATGATTTAAGTATCTTAGTTGAATGGA 723
11D02_vesca         TGTGACAGCTTTAGAGTTTGAATCTTAGGTTGGATGATTTAAGTATCTTAGTTGAATGGA 724
11D02_iinumae       TG--ACAGCTTTAGAGTTTGAATCTTAGGTTGGATGATTTAAGTATCTTAGTTGAACGGA 726
                    **  *******  ***************** ************************* ***

11D02_viridis       TGTTATGACATATTGGTGATTAGTATTAGAGTTATGAGA------AAATAAAATGAAAAT 822
11D02_nubicola      TGTTATGACATATTGGTCATTAGTATTAGAGTTATGAGAAAGAGAAAATAAAATGAAAAT 783
11D02_vesca         TGTTATGACATATTGGTGATTAGTATTAGAGTAATGAGAAAGAGAAAATAAAATGAAAAT 784
11D02_iinumae       TGTTATGACATATTGGTACTTAGTATTAGAGTTATGAGAA----AAGAAAAAATGAAAAT 782
                    *****************  ************* ******      * * ***********

11D02_viridis       ACAGTACTGGCAATAAACACAATACGGTGGAGCAATCAACAAAGCAATAGATTGACAA-G 881
11D02_nubicola      ACAGTACTGGCAATAAACACAATACGGTGGAGCAATCAACAATGCAATAGATTGACAAAG 843
11D02_vesca         ACAGTACTGGCAATAAACACAATACGGTGGAGCAATCAACAATGCAATAGATTGACAAAG 844
11D02_iinumae       ACAGTACTGGCAATAAACACAATTCGGAGGAGCAATCAACAATGCAATAGATTGGCAA-G 841
                    *********************** *** ************** *********** *** *

11D02_viridis       AAATGAAGACCTAAAAAAAACCATTGCATTAATGCAATAGTGTTGATTTTCCAATCTCTC 941
11D02_nubicola      AAATGAAGACCTAAAAAAA-CCATTGCATTAATGCAATAGTGTTGATATTCCAATCTCTC 902
11D02_vesca         AAATGAAGACCTAAAAAAA-CCATTGCATTAATGCAATAGTGTTGATATTCCAATCTCTC 903
11D02_iinumae       AAATGAAGACCTAAAAAAA-CCATTGCATTAATGCAATAGTGTCGATTTTCCAATCTCTC 900
                    ******************* *********************** *** ************

11D02_viridis       CTGAATAGTATTACAACTCTCCTGGACAAGTCATAACTGTGGGGGGTAATGGTGTAAGCA 1001
11D02_nubicola      CTGAATAGTATTACAACTCTCCTGGACAAGTCGTAACTGTGGGGGGTAATGGTGTAAACA 962
11D02_vesca         CTGAATAGTATTACAACTCTCCTGGACAAGTCATAACTGTGGGGGGTAATGGTGTAAACA 963
11D02_iinumae       CTGAATAGTATTACAACTCTCCTGGACAAGTCATACCTGTGGGGGGTAATGGTGTAAACA 960
                    ******************************** ** ********************* **
```

```
11D02_viridis        AACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTTGCTGGGCAGACATAGCACCCCA 1061
11D02_nubicola       AACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTTGCTGGGCAGACATAGCACCCCA 1022
11D02_vesca          AACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTTGCTGGGCAGACATAGCACCCCA 1023
11D02_iinumae        AACAGTCACTAGAATCGAAATTGTTTGTCACAAGTTTTGCTGGGCAGACGTAGCACCCCC 1020
                     ************************************************* *********

11D02_viridis        TAAATCATATCAGATGGGGTTAATGCTACCCAGGTGTGACATATTTGTACAGTTAAACCT 1121
11D02_nubicola       TAAATCATATCAGGTGGGGTTAATGCTACCCAGGTGTGACATATTTGTACAGTTAAACCT 1082
11D02_vesca          TATATCATATCAGATGGGGTTAATGCTACCCAGGTGTGACATATTTGTACAGTTAAACCT 1083
11D02_iinumae        TAAATCATATCAGATGGGGTTAATACTACCCAGGTGTGACATATTTGTACAGTTAAACCT 1080
                     ** ********** ********** ***********************************

11D02_viridis        AATTTTGTCTAAAGAATGCTAAAATCGAA--CTCCCAAGCAACCAAATCTTCTGTTCCCC 1179
11D02_nubicola       AATTTTGTCTAAAGAATGCTAAAATCGAACACTCCCAAGCAACCGAATCTTCTGTTCCCC 1142
11D02_vesca          AATTTTGTCTAAAGAATGCTAAAATCGAA--CTCCCAAGCAACCGAATCTTCTGTTCCCC 1141
11D02_iinumae        AATTTTGTCTAAAGAATGCTAAAATCGAA--CTCCCAAGCAACCGAATCTTCTGTTCCTC 1138
                     *****************************  ************* ************* *

11D02_viridis        TGCTTTAGTATGTTGTGATTATGCCTCTGCTTCCCCAGCAGCATGAATCCGCTCGTCTGG 1239
11D02_nubicola       TGCTTTAGTATGTTGTGGTTATGCCTCAGCTTCCCCAGCAGCATGAATCCGCTCGTCTGG 1202
11D02_vesca          TGCTTTAGTATGTTGTGGTTATGCCTCAGCTTCCCCAGCAGCATGAATCCGCTCGTCTGG 1201
11D02_iinumae        TGCTTTAGTATGTTGTGGTTATGCCTCAGCTTCCCCAGCAGCATGAATCCGCTCGTCTGG 1198
                     ***************** ********* ********************************

11D02_viridis        AGTTACAGCATGAAGCAGTTCGTCTCT---------------------------TGTTGC 1272
11D02_nubicola       AGTTACAGCATGAAGCAGCTCGTCTCT---------------------------TGTTGC 1235
11D02_vesca          AGTTACAGCATGAAGCAGCTCATCTCT---------------------------TGTTGC 1234
11D02_iinumae        AGTTACAGCATGAAGCAGCTCGTCTCTAGTTGCAGCATGAGGTAGCTCGTCTCTTGTTGC 1258
                     ****************** ** *****                           ******

11D02_viridis        AGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGTAGCTCGTCTGGCATTGCAGCATG 1332
11D02_nubicola       AGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGTAGCTCATCTGGCATTGCAGCATG 1295
11D02_vesca          AGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGTAGCTCATCTGGCATTGCAGCATG 1294
11D02_iinumae        AGCATGAGGTAGCTCGTCTCTTGTTGCAGTTTGAGGTAGCTCGTCTGGCATTGCAAGATG 1318
                     ****************************************** ************  ***

11D02_viridis        AAGCTGCTCGTCTGGAGTTGCAGCATTAAGTAGTCCTTCTGGAGTTGCAGCAGGATCTAG 1392
11D02_nubicola       AAGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1355
11D02_vesca          AAGCTGCTCGTCTGGAGTTGCAGCATTAAGTAGTCCTTCTGGAGTTGCAGCAGGATCCAG 1354
11D02_iinumae        AAGCTGCTCGTCTGGAGTTGCAGCATTAAGTAGTCCTTCTGGAGTTGCAGCAGGATCCAG 1378
                     ****

11D02_viridis        GTCCCAACACTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1452
11D02_nubicola       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1415
11D02_vesca          GTCCCAACACTTACCAGGTAGGTTAGTCTCTTCTGCGTCGAGTAACCATGCGGGCACCTG 1414
11D02_iinumae        GTCCCAACACTTACCAGGTAGGTTAGTCTCTTCTGCGTCGAGTAACCATGCGGGTACCTG 1438


11D02_viridis        NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1512
11D02_nubicola       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1475
11D02_vesca          GTGAGAAAAGCGTAACATCTCTCTTCTCGGAATCCATAGAATGGCGCTTCTGTCCGTATC 1474
11D02_iinumae        GTGGGAAAAGCGTAACATCTCTCTTCTCGGAATCCATAGAATGGCGCTTCTGTCCGTATC 1498


11D02_viridis        NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1572
11D02_nubicola       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1535
11D02_vesca          AGTCCGGTATACTGACCTAAATCCAGCCAACTTCACAAGGGGTGAGACACAAACACCAAT 1534
11D02_iinumae        AGTCCGGTATACTGACCTAAATCCAGCCAACTTCACAAGGGGTGAGACACAA-CACCAAT 1557


11D02_viridis        NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1632
11D02_nubicola       NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1595
11D02_vesca          CTCTTCAGAGTAATCATCAAGAACCTCCACCATTTGATATTGGTGCCTCACTTCATCTGG 1594
11D02_iinumae        CTCTTCAGAGTAATCATCAAGAACCTCCACCATTTGATAT-GGTGCCTCACTTCATCTGG 1616
```

```
11D02_viridis       NNNNNNNN 1640
11D02_nubicola      NNNNNNNN 1603
11D02_vesca         AGTTGAAC 1602
11D02_iinumae       AGTNNNNN 1624
```

17O22

```
17O22_vesca          AAAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCT 60
17O22_mandshurica    -AAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCT 59
17O22_viridis        TAAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCT 60
17O22_nubicola       -AAATGGGTTGCACGAGTTCGTGAACGTACAATTTACGACCCAAAGCGTCCAATACTGCT 59
17O22_iinumae        AAAATGGGTTG-ACGAGTTCGTGAACA-ACACTTTACGACCCAAAGCGTCCAATACTTCT 58
                      ********** **************  *** ************************* **

17O22_vesca          TAATTTGACAACAGACATAGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACT 120
17O22_mandshurica    TAATTTGACAACAGACATAGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACT 119
17O22_viridis        TAATCTGACAACGGACATAGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACT 120
17O22_nubicola       TAATTTGACAACGGACATAGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACT 119
17O22_iinumae        TAATTTGACAACGGACATAGTAGAGGAAAACAGGTACCTCCAATGCAAGGAATCGGCACT 118
                     **** ******* ***********************************************

17O22_vesca          AGAGACTGCATTTCTTATAAAGGCAATGGAATCGTAGAGACTGCATTTCTTACTCAGTAC 180
17O22_mandshurica    AGAGACTGCATTTCTTATAAAGGCAATGGAATCGTAGAGACTGCATTTCTTACTCAGTAC 179
17O22_viridis        AGAGACTGCATTTCTTATAAAGGCAATGGAATCGTAGAGACTGCATTTCTTACTCAGTAC 180
17O22_nubicola       AGAGACTGCATTTCTTATAAAGGCAATGGAATCGTAGAGACTGCATTTCTTACTCAGTAC 179
17O22_iinumae        AGAGACTGCTTTTCTTATAAAAGCAATGGAATCGTAGAGACTGCATATCTTACTCAGCAC 178
                     ********* *********** ************************ ********** **

17O22_vesca          TGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTTCCCGAAATT 240
17O22_mandshurica    TGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTTCCCGAAATT 239
17O22_viridis        TGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTTCCCGAAATT 240
17O22_nubicola       TGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTTTAACTTCCCGAAATT 239
17O22_iinumae        TGAATCTGTTGGTCAGCAACACAGAAACTAGCTGTGGGCAATGTATAACTTCCCGAAATT 238
                     ******************************************** ***************

17O22_vesca          CAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACTCTACATTCATTCGATC 300
17O22_mandshurica    CAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACC--ACATTCATTCGATC 297
17O22_viridis        CAACAGCCATCAGAGTTCATCTGCCAATCAAGGCAAATATGACTCTACATTCATTCGATC 300
17O22_nubicola       CAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACTCTACATTCATTCGATC 299
17O22_iinumae        CAACAGCCATCAGAGTTCATCTGCCAATCAGGGCAAATATGACTCTACATTCATTCGATC 298
                     ****************************** ************   **************

17O22_vesca          CCCTTATCACTGTAGGGCTTCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATA 360
17O22_mandshurica    CCCTTATCACTGTAGGGCTTCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATT 357
17O22_viridis        CCCTTATCACTGTAGGGCTCCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATT 360
17O22_nubicola       CCCTTATCACTGTAGGGCTCCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATT 359
17O22_iinumae        CCCTTATCACTGTAGGGCTCCATTGGAAACGCTTTGGTCAGCGCAAGACTGATGTTGATT 358
                     ******************* ***************************************

17O22_vesca          GTAGCCTAGTTTAGTTTCTTATGCTGAAGCAAAATATGTAATCACCTACGCTACAGAATA 420
17O22_mandshurica    GTAGCCTAGTTTAGTTTCTTATGCTGAAGCAAAATATGTAATCACCTAGGCTACAGAATA 417
17O22_viridis        GTAGCCTAGTCTAGTTTCTTATGCTGAAGCAAAATATGTAATCACCTAGGCTACAGAATA 420
17O22_nubicola       GTAGCCTAGTTTAGTTTCTTATGCTGAAGCAAAATATGTAATCACCTAGGCTACAGAATA 419
17O22_iinumae        GTAGCCTAGTTTAGTTTCTTATGCTGAAGCAAAATATGTAATCACCTAGGCTACAGAAGA 418
                     ********** ************************************* ********* *

17O22_vesca          GTGTT-ACTTGTTACCGGACATGTTCACAATCTTTGAAGATGAAGAACGGTACCAGTTAC 479
17O22_mandshurica    GTGTT-ACTTGTTACCGGACATGTTCACAATCTTTGAAGATGAAGAACGGTACCAGTTAC 476
17O22_viridis        GTGTT-ACTTGTTACCGGACATGTTCACAATCTTTGAAGATAAAGAACGGTACCAGTTAC 479
17O22_nubicola       GTGTTTACTTGTTACCGGACATGTTCACAATCTTTGAAGATGAAGAACGGTACCAGTTAC 479
17O22_iinumae        GTGTT-ACTTATTACCGGACATGTTCACAATCTTTGAAGATGAAGAACGGTACCAGTTAC 477
                     ***** **** ****************************** ******************
```

```
17O22_vesca            CCAACATAATCATAGTTATTTTGGCCTATTGATATTTTGATTAACGTGTAATTGATCGCT 539
17O22_mandshurica      CCAACATAATCATAGTTATTTTGGCCTATTGATATTTTGATTAATGTGTAATTGATCGCT 536
17O22_viridis          CCAACATAATCATACTTGTTTTGGCCTATTGATATTTTGATTAATATGTAATTGATCGCT 539
17O22_nubicola         CCAACATAATCATAGTTATTTTGGCCTA-----ATTTTGATTAATATGTAATTGATCGCT 534
17O22_iinumae          CCAACATAATCATAGTTATTTCGGAATATTGATATTTTGATTAATATGTAATTGATCGCT 537
                       ************** ** *** **  **     ***********  **************


17O22_vesca            ACTTGAATGATGTATATTAT-----GAATGGCACTATTTAATATTTTGGGCTGCTACCTA 594
17O22_mandshurica      ACTTGAATGATGTATATTAT-----GAATGGCACTATTTAATATTTTGTGCTGCTACCTA 591
17O22_viridis          ACTGGAATGATGTATATTATATTATGAATGGCACTATTTAATATTTTGGGCTGCTACCTA 599
17O22_nubicola         ACTTGAATGATGTATATTAT-----GAATGGCACTATTTAATATTTTGGGCTGCTACCTA 589
17O22_iinumae          ACTTGACTGATGTATATTAT-----GAATGTCACTATTTAATATTTTGGGCTGCTACCTA 592
                       *** ** *************     ***** ***************** ***********


17O22_vesca            CTCTTCAACAAACTCTAATTAATTAACCAAACATCAGTGTCACAAGTCACACCAACCTAG 654
17O22_mandshurica      CTCTTCAACAAACTCTAATTAATTAACCAAACATCAGTGTCACAAGTCACACCAACCTAG 651
17O22_viridis          CTCTTCAACAAACTCTAATTAATTAACCAAACATCAGTGTCACAGGTCACACCAACCTAG 659
17O22_nubicola         CTCTACAACAAACTCTAATTAATTAACCAAACATCAGTGTCACAAGTCACACCAACCTAA 649
17O22_iinumae          CTCTTCAACAAACTTTAATTAATTAACCAAACATCAGTGTCACAAGTCACACCAACCTAG 652
                       **** ********* ***************************** **************


17O22_vesca            TTAAACTTTCCATTATAAGTAGCTTTCCCAATAACCTACCTCCCAAAAATAGTTACTTTA 714
17O22_mandshurica      TTAAACTTTCCATTATAAGTAGCTTTCCCAATAACCTACCTCCCAAAAATAGTTACTTTA 711
17O22_viridis          TTAAACTTTCCATTATAAGTAGCTTTCCCAATAACCTACCTCCCAAAAATAGTTACTTTA 719
17O22_nubicola         TTAAACTTTCCATTACAAGTAGCTTTCCCAATAACCTACCTCCCAAAAATAGTTACTTTA 709
17O22_iinumae          TTAAACTTTCCATTATAAGTAGCTTTCCCAATAACATACCTCCCAAAAATAGTTACTTTA 712
                       *************** ******************* ************************


17O22_vesca            AA-GCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATTCTGGATTTG 773
17O22_mandshurica      AA-GCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATTCTGGATTTG 770
17O22_viridis          AAAGCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATTCTGGATTTG 779
17O22_nubicola         AA-GCTAGTTCTTGTCAAATAGTGAACCACCATCAACTCTTGCCTATAATTCTGGATTTG 768
17O22_iinumae          AA-GCTGGTTCTTGTCAAATAGTGAACCACCATCAACTCTTCCCTATAATTCTGGATTTG 771
                       ** *** ********************************** ******************


17O22_vesca            TTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTCTGGTCAT 833
17O22_mandshurica      TTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTTGGTCGT 830
17O22_viridis          TTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTTGGTCAT 839
17O22_nubicola         TTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTTGGTCAT 828
17O22_iinumae          TTACTCGCTAGTATGTGTTGAACTTTGTTTCTTTTACAAAGACAAAAGGACTTCGGTCAT 831
                       ****************************************************  **** *


17O22_vesca            CAGTGTCAAACTAGAAGAACCGTGAATTGCGAC--------------------------- 866
17O22_mandshurica      CAGTGTCAAACTAGAAGAACCGTGAATTGCGACTATACCAGGATGCCTTTGGTCACTTAC 890
17O22_viridis          CAGTGTCAAACTAGAAGAACTGTGAATTGCGACTACACCAGGATGCCTTTGGTCACTTAC 899
17O22_nubicola         CAGTGTCAAACTAGAAGAACCGTGAATTGCGACTATACCAGGATGCCTTTGGTCACTTAC 888
17O22_iinumae          CAGTGTCAAACAAGAAGAACCGTGAATTGCGACTATACCAGGATGCCTTTGGCCACTTGC 891
                       *********** ******** ************


17O22_vesca            -------------------CCCTCAGAATGTCAAAATGAGATCACTGTGATTCCTTTTAA 907
17O22_mandshurica      CAACCTCAAGAAAAGGAC-CCCTCAGAATGTCACAATGAGATCACTGTGATTCCTTTTAA 949
17O22_viridis          CAACCTCAAGAAAAGGACCCCCTCAGAATGTCAAAATGAGATCACTGTGATTCCTTTTAA 959
17O22_nubicola         CAACCTCAAGAA-----------------TGTCAAAATGAGATCTCTGTAATTCCTTTTAA 932
17O22_iinumae          CAACCTCAAGAAAAGGAC-CCCTCAGAATGTCAAGATGAGATCACTGTGATTCCTTTTAA 950
                                                   *****  ******** **** ***********
```

```
17O22_vesca            AATTTTAACAGCGATTCTTCTACAAAAGATG-GACTAAATTCCACCTTGTACTGTACAAA 966
17O22_mandshurica      AATTTTAACAGTGATTCTTCTACAAAAGATAAGACTAAATTCCACCTTGTACTGTACAAA 1009
17O22_viridis          AATTTTAACAGTGATTCTTCTACAAAAGA-----CTAAATTCCACTTTGTACTGTACAAA 1014
17O22_nubicola         AATTTTAACAGTGATTCTTCTACAAAAGATG-GACTAAATTCCACCTTGTACTGTACAAA 991
17O22_iinumae          AATTTTGACAGTGATTCTTCTACAAAAGATG-GACCAAATTCCACCTTGTACTGTACAAA 1009
                       ****** **** *****************     * ********* **************

17O22_vesca            AAACGAGTTTGAGTAGTGGGAATCGTTCCAATAT-ATTTCTGCTCTGTTTACCAATTGCC 1025
17O22_mandshurica      AAACGAGTTTGAGTAGTGGGAATCGTTCCAATAT-ATTTCTGCTCTGTTTACCAAGTGCC 1068
17O22_viridis          AAACGAGTTTGAGTAGTGGGAATCGTTCCAATAT-ATTTCTGCTCTGTTTACCAATTGCC 1073
17O22_nubicola         AAACGAGTTTGAGTAGTGGGAATCGTTCCAATATTATTTCTGCTCTGTTTACCAATTGCC 1051
17O22_iinumae          AAACGAGTTTGAGCAGTGGGAATCGTTCCAATAT-ATTTCTGCTCTGTTTACCAATTGCC 1068
                       ************* ******************** ******************** ****

17O22_vesca            AGGATGATACAAACATCTAAACTCTACAGGAACCCTTTTCTAGCAAAAGAA-TGAGAAGA 1084
17O22_mandshurica      AGGATGATACAAACATCTAAACTCTACAGGAACCATCTTCTAGCAAAAAAA-TGAGAAGA 1127
17O22_viridis          AGGATGATACAAACATCTAAACTCTACAGGAACCCTTTTCTAGCAAAAAAA-TGAGAAGA 1132
17O22_nubicola         AGGATGATTCAAACATCTAAACTCTACAGGAACCCTTTTCTAGCAAAAAAA-TGAGAAGA 1110
17O22_iinumae          AGGATGATACAAACATCTAAACTCTACAGGAACCCTTTTCTAGCAAAAAAAATGAGAAGA 1128
                       ******** ************************* * *********** ** ********

17O22_vesca            AAGAACTCTACAAGAATCCAAAGCGCGAAAACAAAATCAGAACTAAGACTAGACATGAAC 1144
17O22_mandshurica      AAGAACTCTACAAGAATC-AAAGCGCGAAAACAAAATCAGAACTAAGACTAGACATGAAC 1186
17O22_viridis          AAGAACTCTACAAGAATCCAAAGTGCGAAAACAAAATCAGAACTAAGACTAGACATGAAC 1192
17O22_nubicola         AGGAACTCTACAAGAATCCAAAGCGCGAAAACAAAATCAGAACTAAGACTAGACATGAAC 1170
17O22_iinumae          AAGAACTCTACAAGAATCCAAAGCGCGAAAACAAAATCAGAACTAAGACTAGACATGAAC 1188
                       * **************** **** ************************************

17O22_vesca            AAATTTGCTGCAGCCTCCACTGATGAGCTTCTCCAGCAAGAACAAAAGAATCAAACCAGA 1204
17O22_mandshurica      AAATTTGCTGCAGCCTCCACTGAGGAGCATCTCCAGCAAGAACAAAAGAATCAAACCAGA 1246
17O22_viridis          AAATTTGCTGCAGCCTCCACTGATGAGCTTCTCCAGCAAGAACAAAAGAATCAAACCAGA 1252
17O22_nubicola         AAATTTGCTGCAGCCTCCACTGATGAGCTTCTCCAGCAAGTACAAAAGAATCAAACCAGA 1230
17O22_iinumae          AAATTTGCTGCAGCCTCCACTGATGAGCTTCTCCAGCAAGAACAAAAGAATCAAACCAGA 1248
                       *********************** **** *********** *******************

17O22_vesca            TAAAATGGAAAATCTCCTCTCACGTTGGAACAATATCATTGATTTCAGATTTTGTCTCAG 1264
17O22_mandshurica      TAAAATGGAAAATCTCCTCTCACGTTGGAACAATATCATTGATTTCAGATTTTGTCTCAG 1306
17O22_viridis          TAAAATGGAAAATCTCCTCTCACGTTGGAACAATATCATTGATTTCAGATTTTGTCTCAG 1312
17O22_nubicola         TAAAATGGAAAATCTCCTCTCACGTTGGAACAATATCATTGATTTCAGATTTTGTCTCAG 1290
17O22_iinumae          TAAAATGGAAAATCTCCTCTCACGTCGGAACAATATCATTGATTTCAGATTTTGTCTCAG 1308
                       ************************* **********************************

17O22_vesca            ATTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATGGATTCAGAAAATTTG 1324
17O22_mandshurica      ATTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATGGATTCAGAAAATTTG 1366
17O22_viridis          ATTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATGGATTCAGAAAATTTG 1372
17O22_nubicola         ATTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGAAGGATTCAGAAAATTTG 1350
17O22_iinumae          ATTCTTCGTCAACAGTAGATAGTCCGCCTTCTCTGATGAAGGATGGATTCAGAAAATTTG 1368
                       ******************************************* ****************

17O22_vesca            CTACAAAAGCCCATAACTTGTAAA-CATCATCGAAGT-TTGTGAGGAAACCC 1374
17O22_mandshurica      CTACAAAAGCCCATAACTTGTAAGGCATCATCGAAGTATTGTGAGGAAACCC 1418
17O22_viridis          CTACAAAAGCCCATAACTTGTAAA-CATCATCGAAGT-TTGTGAGGAAACCC 1422
17O22_nubicola         CTACAAAAGCCCATAACTTGTAAA-CATCATCGAAGT-TTGTGAGGAAACCC 1400
17O22_iinumae          CTACAAAAGCCCATAACTTGTAAA-CATCATCGAAGT-TTGTGAGGAAACCC 1418
                       ***********************  ************ **************
```

27F10

```
27F10_vesca            CCTGCAGGGTTTTTCATCATGTAAGGACCTCCATTGT-CAGTAGCTTTATGCATATCATC 59
27F10_mandshurica      CCTGCAGGGCTTTTTATCATGTAAGGACCTCCATTGT-CAGTAGCTTTATGCATATCATC 59
27F10_nubicola         CCTGCGGG--TTTTTATCATGTAAGGACCTCCATTGT-CAGTAGCTTTATGCATATCATC 57
27F10_iinumae          CCTGCAGG-TTTTTCATCATGTAAGGACCTCCATTGT-CAGTAGCTTTATGCATATCATC 58
27F10_ananassa         CCTGCAGG-TTTTT-ATCATGTAAGGACCTCCATTGT-CAGTAGCTTTATGCATATCATC 57
27F10_viridis          --TGCGGG-TTTTTCATCATGTAAGGACCTCCATTGTTCGGTAGCTTTATGCATATCATC 57
                         *** **  **** ********************** * ********************
```

```
27F10_vesca            TTCATCACAACAGCTGAAGCAGCTCATG-ATTCCTTTAAACACACACAAAAAAA---CCC 115
27F10_mandshurica      TTCATCACAACAGCTGAAGCAGCTCATG-ATTCCTTTAAACACACACAAAAAAA---CCC 115
27F10_nubicola         TTCATCACAACAGCTGAAGCAGCTCATG-ATTCCTTTAAACACACACAAAAAAAA--CCC 114
27F10_iinumae          TTCATCACAACAGCTGAAGCAGCTCATG-ATTCCTTTAAACACACA--AAAAAAA--CCC 113
27F10_ananassa         TTCATCACAACAGCGGAAGCAGCTCATGGACTCCTTTAAACACACA--AAAAAAA--CCC 113
27F10_viridis          TTCATCACAACAGCTGAAGCAGCTCATG-ATTCCTTTAAACACACAAAAAAAAAACACCC 116
                       ************** ************* * ***************  ******   ***

27F10_vesca            ACAGTCAAAATGAGGAAATGAACAATACCCAAGTCATGAACACACAAAATTCAGTAAAAA 175
27F10_mandshurica      ACAATCAAAATGAGGAAATGAACAATACCCAAGTCATGAACACACAAAATTCAGTAAAAA 175
27F10_nubicola         ACAATCAAAATGAGAAAATGAACAATACCCAAGTCATGAACACACAAAATTCAGTAAAAA 174
27F10_iinumae          ACAATCAAAATGAGGAAATGAACAATACCCAAGTCATGAACGCACAAAATTCAGTAAAAA 173
27F10_ananassa         ACGATCAAAATGAGGAAATGAACAATACCTAAGTCATGAACACACAAAATTCAGTAAAAA 173
27F10_viridis          ATAATCAAAATGAGGAAATGAACAATACCCTAGTCATGAACACACAAAATTCAGTAAAAA 176
                       *   ********** **************  ********** ******************

27F10_vesca            AGTAAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCAGACCTTTGGAGACAAAT 235
27F10_mandshurica      AGTAAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCAGACCTTTGGAGACAAAT 235
27F10_nubicola         AG-AAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCAGACCTTTGGAGACAAAT 233
27F10_iinumae          AGAAAAAAGGGATCCGCTTCAATACAATCCCATCAAACTTGCAGACCTTTGGAGACAAAT 233
27F10_ananassa         AGAAAAAAGGGATCCGCTTCAATCCAATCCCATCAAACTTGCACCCCTTTGGAGACAAAT 233
27F10_viridis          AGAAAAAAGGGATCCGCTTCAAGCCAATCCCATCAAACTTGCAGACCTTTGGAGACAAAT 236
                       ** *******************  *******************  ***************

27F10_vesca            TTCGTTGCTTAATGTAATAAGCAACAAAAAATTCAGCTCAGCTGGATCAAAGCCCAGATG 295
27F10_mandshurica      TTCGTTGCTTAATGTAATAAGCAACAAAAAAATCAGCTCGGCTGGATCAAAGCCCAGATG 295
27F10_nubicola         TTCGTTGCTTAATGTAATAAGCAACAAAAAATTCAGCTCAGCTGGATCAAAGCCCAGATG 293
27F10_iinumae          TTCGTTGCTTAATGTAATAAGCAACAAAAA-TCCAGCTCAGCTGGATCAAAGCCCAGATG 292
27F10_ananassa         TTCGTTGCTTAATGTAATAAGCAACAAAAA-TTCAGCTCAGCTGGATCAAAGCCCAGATG 292
27F10_viridis          TTCGTTGCTTAATGTAATAAGCAACAAAAAATTCAGCTCAGCTGGATCAAAGCCCAGATG 296
                       ******************************   ****** ********************

27F10_vesca            AAAAAGATTAAAACTTTAAACAAGAAAATAAAGATCAGAGAAAGAAAATATGATGGGTAG 355
27F10_mandshurica      AAAAAGATTAAAACTTTAAACAAGAAAATAAAGATCAGAGGAAGAAAATATGATGGGTAG 355
27F10_nubicola         AAAAAGATTAAAACTTCAAACAAGAAAATAAAGATCAGAGGAAGAAAATATGATGGGTAG 353
27F10_iinumae          AAAAAGATTAAAACTTTACCCAAGAAAATAAAGGTCAGAGGAAGAAAATATGATGGGTAG 352
27F10_ananassa         AAAAAGATTAAAACTTTACCCAAGAAAATAAAGGTCAGAGGAAGAAAATATGGTGGGTAG 352
27F10_viridis          AAAAAGATTAAAACTTTAAACAAGAAAATAAAGATCAGAGGAAGAAAATATGATGGGNAG 356
                       **************** *  ************* ****** *********** **** **

27F10_vesca            ATCGGGAGAGATAAAATTACCTGAATCTGAAGTGGGGGAAGTGAGTCAGTGAAGGACTGA 415
27F10_mandshurica      ATCGGGAGAGATAAAATTACCTGAATCTGAAGTGGGGGAAGTGAGTCAGTGAAGGACTGA 415
27F10_nubicola         ATCGGGAGAGATAAAATTACCTGAATCTGAAGTGGGGGAAGTGAGTCAGTGAAGGACTGA 413
27F10_iinumae          ATCGGGAGAGATAAAATTACCAGAATCTGAAGTGGGGGAAGTGAATCAGTGAAGGACTGA 412
27F10_ananassa         ATCGGGAGAGATAAAATTACCAGAATCTGAAGTGGGGGAAGTGAATCAGTGAAGGACTGA 412
27F10_viridis          ATCGGGAGAGATAAAATTACCTGAATCTGAAGTGGGGGAAGTGAGTCCGTGAAGAAGTGA 416
                       ********************* ********************** ** ****** * ***

27F10_vesca            GTTGGTGGAGTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGCG-AAGGATGCGCGG 474
27F10_mandshurica      GTTGGTGGAGTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGCG-AAGGATGCGCGG 474
27F10_nubicola         GTTGGTGGAGTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGCG-AAGGATGCGCGG 472
27F10_iinumae          GTTGCTGGAGTCTTGGGAGATCTGAG-------CTCTAAAGCCGGCG-AAGGATGCGCGG 464
27F10_ananassa         GTTGCTGGAGTCGTGGAAGATCTGAG-------C-------CCGGCG-AAGGATGCGCGG 457
27F10_viridis          GTTGGTGGANTCTTGGGAGATCTGAGATATGAGCTCTAAAGCCGGCGCAAGGATGCCCGG 476
                       **** **** ** *** *********       *       ****** ******** ***

27F10_vesca            CGCAGGATAGGAGGGAAAAGGGTGCGTAGGATAACCCAATCAATGAACCAGATGAGAATA 534
27F10_mandshurica      CGCAGGATAGGAGGGAACAGGGTGCGTAGGATAACCCAATCAATGAACCAAATGAGAATA 534
27F10_nubicola         CGCAGGATAGGAGGGAAAAGGGTGCGTAGGATAACCCAATCAATGAACCAAATGAGAATA 532
27F10_iinumae          CGCAGGATAGGAGGGAAAAGGGTGCGTAGGATAACCCAATCAATGAACCAAATGAGAATA 524
27F10_ananassa         CGCAGGATCGGAGGGAAAAGGGTGCGTAGGATAACCCAACCAATGAACCAAATGAGAACA 517
27F10_viridis          CGCAGGATAGGAGGGAAAAGGGTGCGTAGGATAACCCACTCCANGAACCANATGACAATG 536
                       ******** ******** ********************  * * ****** **** **
```

```
27F10_vesca            CGCTAGTGATTTTGATTATGAATTCTATAAATTCTATAAAAA-TTTATTTCATTTCTTAA 593
27F10_mandshurica      CGCTAGTGATTTTGATTATGAATTCTATAAATTCTATAAAAA-TTTATTTCATTTCTTAA 593
27F10_nubicola         CGCTAGTGATTTTGATTATGAATTCTATAAATTCTACAAAAAATTTATTTCATTTCTTAA 592
27F10_iinumae          CGCTAGTGATTTTGATTATGAATTCTATAAATTCTACAAAAA-TTTATTTCATTTCTTAA 583
27F10_ananassa         CGCTAGTGATTTTGATTATGAATTCTATAAATTCTACAAAAA-TTTATTTCATTTCTTAA 576
27F10_viridis          CNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 596
                       *


27F10_vesca            TTCTTACTCTGTTTCGGTGTTGGCCAGATTTGACTCTTCTGTGCTTCAGT------TTTG 647
27F10_mandshurica      TTCTTACTCTGTTTCGGTGTTGGCCAGATTTGACTCTTCTGTGCTTCAGT------TTTG 647
27F10_nubicola         TTCTTACTCTGTTTCGGTGTTGGCCAGATTTGACTCTTCTGTGCTTCAGT------TTTG 646
27F10_iinumae          TTCTTACTCTGTTTCGGTGTTGGCCAGATTTGACTCTTCTGTGCTTCAGTCATGACTTTG 643
27F10_ananassa         TTCTTACTCTGTTTCGGTGTTGGCCAGATTTGACACTTCTGTGCCTCAGTCATGACTTTG 636
27F10_viridis          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 656


27F10_vesca            ACCATTTACTTTTATAACCTCAGGAAGGGTTCAAGCGCGGCCTGCCACGTGGTGAATTC- 706
27F10_mandshurica      ACCATTTATTTTTATATCCTCAGGAAGGGTTCAAGCGCGGCCTGCCACGTGGTGAATTC- 706
27F10_nubicola         ACCATTTACATTTATAACCCCGGGAAGGGTTCAAGCGCGGCCTGCCACGTGGTGAATTC- 705
27F10_iinumae          ACCATTTACTTTTATAACCCCAGGAAGGGTTCAAGCGCGGCCTGCCACGTGGTGAATTCT 703
27F10_ananassa         GCCATTTACTTTTATAACCCCAGGAAGGGTTCAAGCGCGGCCTGCCACGTGGTGAATTCT 696
27F10_viridis          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 716


27F10_vesca            -----------AAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCC 755
27F10_mandshurica      -----------AAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCC 755
27F10_nubicola         -----------AAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCC 754
27F10_iinumae          GGTTCGTCCGGAAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCC 763
27F10_ananassa         GGTTCGTCCGGAAAAGAGTCTGGAAGCAAAGCCTTGACCTCGTGGAATTCGTCTCTCCCC 756
27F10_viridis          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 776

27F10_vesca            TCCCGGTAACAGTAACTTTATCGACAAAACGCTTCTTATTTTATTTTATTTTTTTTGGCG 815
27F10_mandshurica      TCCCGGTA-CAGTAACTTTATCGTTTTACCGCTAGTATGTCTCTGTCAGTACTCT--GTC 812
27F10_nubicola         TCCCGGTAACAGTAACTTTATCGTTTTACCGCTAGTATGTCTCTGTCT---------GTC 805
27F10_iinumae          TCCGGGTAACAGTAACTTTATCGTTTTACCGCTAGCATGTCTCTGTCTGTC-----GACA 818
27F10_ananassa         TCCCGGCAACAGTAACTTTATCGTTTTACCGCTAGTATGTCTCTGTCT------------ 804
27F10_viridis          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 836

27F10_vesca            AGCAAAACGCTTCTTATTTGTTTTGGGTCTGTACGCTTTTGGGTCTTATTTGTCAAGTTT 875
27F10_mandshurica      GACATAACGCTTCTTATTTGTTTTGGGTCTCTACGCTTTTGGGTCTTATTTGTCAAGTTT 872
27F10_nubicola         GACATGACGCTTCTTATTTCTTTTGGGTCTCTACGCTTTTGGGTCTTATTTGTCAAGTTT 865
27F10_iinumae          TATATAACGCTTCTTATTTGTTTTGGGTCTCTACGCTTTTGGGTCTTATTTGTCAAGTTT 878
27F10_ananassa         --------------TATTTGTTTTGGGTCCCTACGCTTTTGGGTCTTATTTGTCAAGTTT 850
27F10_viridis          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 896

27F10_vesca            CAATCACTAGCAGGA--------AGACTTGCGTAT------------------------- 902
27F10_mandshurica      CAATCACTAGCAGGT--------AGACTTGCGTAT------------------------- 899
27F10_nubicola         CAATCACTAGCAGGA--------AGACTTGCGTAT------------------------- 892
27F10_iinumae          CAATCACTTGAAACT--------------------------------------------- 893
27F10_ananassa         CAATCACTTGAAACATAGCAGGAAGACTTGCATAT------------------------- 885
27F10_viridis          NNNNNNNNNNNNNNN--------NNNNNNNNNTATGCAAAAATACACTCATATTTATGTA 948

27F10_vesca            ------------------------------------------------------------
27F10_mandshurica      ------------------------------------------------------------
27F10_nubicola         ------------------------------------------------------------
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         ------------------------------------------------------------
27F10_viridis          GAAAACGAGAATTGAACCTCTAACCTCTTACAAACAACCTATGAAATGTATAATATATGT 1008
```

```
27F10_vesca            ------------------------------------------------------------
27F10_mandshurica      ------------------------------------------------------------
27F10_nubicola         ------------------------------------------------------------
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         ------------------------------------------------------------
27F10_viridis          AAAGACGATTAAAATATATGTATAATATATAATATATGTATTGTTTTATATTTATAACAT 1068


27F10_vesca            ------------------------------------------------------------
27F10_mandshurica      ------------------------------------------------------------
27F10_nubicola         ------------------------------------------------------------
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         ------------------------------------------------------------
27F10_viridis          ACTATAGATATAAATACAATTAAAATATAAAATGTTCAAATTTTAGCAAGAGGTATGTTT 1128


27F10_vesca            ------------------------------------------------------------
27F10_mandshurica      ------------------------------------------------------------
27F10_nubicola         ------------------------------------------------------------
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         ------------------------------------------------------------
27F10_viridis          CGAAACCATGACTGCTCTGATGGAAAATATGACCACTTACGATCAAAACAAAGCTATCAT 1188


27F10_vesca            ------------------------------------------------------------
27F10_mandshurica      ------------------------------------------------------------
27F10_nubicola         ------------------------------------------------------------
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         ------------------------------------------------------------
27F10_viridis          TGCATTATATTTGTGAAAAAAATTATATTTATCACTTCATTTTTTGGGCCACAATCTAAG 1248


27F10_vesca            ------------------------------------------------------------
27F10_mandshurica      ------------------------------------------------------------
27F10_nubicola         ------------------------------------------------------------
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         ------------------------------------------------------------
27F10_viridis          TTTAGTAGAGGCCTATTACCAACCGTACCAACTAAGTCGGTATACCAACATCGATGGTTG 1308


27F10_vesca            ------------------------------------------------------------
27F10_mandshurica      ------------------------------------------------------------
27F10_nubicola         ------------------------------------------------------------
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         ------------------------------------------------------------
27F10_viridis          GTTTTGATAGAGGATTTTGCCTACCAATCATAAGTTGGTTGGTACATGATATTGGTAAAT 1368


27F10_vesca            -----------------------------------------------------------A 903
27F10_mandshurica      -----------------------------------------------------------A 900
27F10_nubicola         -----------------------------------------------------------A 893
27F10_iinumae          ------------------------------------------------------------
27F10_ananassa         -----------------------------------------------------------A 886
27F10_viridis          AAAGTCGGTATATCTACCAATGCCAGCCCTACTTGAAACTTAGCCGGAAGACTTCATATA 1428


27F10_vesca            ATTGAAATAGCCACTATCTATACT--CTATATGCAAACACAA---GAGTAGAGAAGGAGA 958
27F10_mandshurica      ATTGAAATAGCCACTATCTATACT--CTATATGCAAACACAA---GAG-AGAGAAGGAGA 954
27F10_nubicola         ATTGAAATAGCCATTATCTATACT--CTATATGCAAACACAA---GAG-AGAGAAGGAGA 947
27F10_iinumae          ATTGAAATAGCTGAAATACACACTTACTATATGCAAACACAA---GGG-AGAG-AGGAGA 948
27F10_ananassa         ATTGAAATAGCTGCAATACACACTTGCTATATGCAAACACAACAAGAG-AGAGGAGGAGA 945
27F10_viridis          ATTGAAATAGCTGAGATACACACTTGCTATATGCAAACACAA---GAGTAGAGAAGGAGA 1485
                       ***********    **  * ***  ****************   * * **** ******
```

```
27F10_vesca            ACCAGAATACATTTCCA 975
27F10_mandshurica      ACCAGAATACATTTCCA 971
27F10_nubicola         ACCAGAAT---CATCCA 961
27F10_iinumae          ACCAGATC---ATTCTA 962
27F10_ananassa         ACCAGAAT---CATCCA 959
27F10_viridis          ACCAGAATACATTTCCA 1502
                       ******       ** *
--------------------------------------------------------------------------------------------
```

29G10

```
29G10_vesca             TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGAGCTTTGGAGAGGTAGGAAGAGTTTATT 60
29G10_mandshurica       TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGAGCTTTGGAGAGGTAGGAAGAGTTTATT 60
29G10_nubicola          TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGGGCTTTGGAGAGGTAGGAAGAGTTGATT 60
29G10_nilgerrensis      TGGCCTTGTTTCCTAAACTCTTCAGGGTCTAGGGCTTTGGAGAGTTAGGAAGAGTTGATT 60
                        ******************************** *********** *********** ***

29G10_vesca             TCTAGAGGGAGGCTACCCATTTGAAGTAGAGATTGGACTAAAAACAACTTGAAAG--GAA 118
29G10_mandshurica       TCTAGAGGGAGGCTACCCATTTGAAGTAGAGATTGGACTAAAAACAACTTGAAAG--GAA 118
29G10_nubicola          TCTAGAGGGAGGCTACCCATTTGAAGTAGAGATTG-ACTAAAAACAACTTGAAAGAGGAA 119
29G10_nilgerrensis      TCTAGTGGGAGGCTACCCATTTGAAGTAGAGATTGGACTAAAAACAACTTGAAAG--GAA 118
                        ***** ***************************** *******************  ***

29G10_vesca             GATGGGGAGGATAAATAAAAAGGATAGAAACTGCTCAAGTGCTTAACAATGGTTGTAGAC 178
29G10_mandshurica       GATGGGGAGGATAAATAAAAAGGATAGAAACTGCTCAAGTGCTTAACAATGGTTGTAGAC 178
29G10_nubicola          GATGGGG-------------AGGATAGAAACTGCTCAAGTGCTTAACAATGGTTGTAGAC 166
29G10_nilgerrensis      GATGGGGAGGATAAATAAAAAGGATAGAAACTGCTCAAGTGCTTAACAATGGTTGTGGAC 178
                        *******             ************************************ ***

29G10_vesca             GAGTTGTGTCTTGCTGCATATATTGAAGAGATTATATAGAGGTGCATGTAGGATGAAGAC 238
29G10_mandshurica       GAGTTGTGTCTTGCTGCATATATTGAAGAGATCATATAGAGGTGCATGAAGGATGAAGAC 238
29G10_nubicola          GAGTTGTGTCTTGCTGCATATATTGAAGAGATTATATAGAGGTGCATGTAGGATGAAGAC 226
29G10_nilgerrensis      GAGTTGTGTCTTGCTGCATATATTGAAGAGATTATATAGAGGTGCATGTAGGATGAAGAC 238
                        ******************************** ************** ************

29G10_vesca             GCCGTATCTTAAATTTTGATTTGGTTCTTCTCA-------CACACCAGAGATTGAGTTCG 291
29G10_mandshurica       ACCGTATCTTAAATTTTGATTTGGTTCTTCTCAGTTCTCACACACCAGAGATTGAGTTCG 298
29G10_nubicola          ACCGTATCTTAAATTTTGATTTGGTTCTTCTCA-------CACACCAGAGATTGAGTTCG 279
29G10_nilgerrensis      ACCGTATCTTAAATTTTGATTTGGTTCTTCTCA-------CACACCAGAGATTGAGTTCG 291
                         ********************************       ********************

29G10_vesca             GATCATCGGATCCGAAAAATCAAGTCCTTGTGTATAAAAGCACGTTACGGAGTGATCCCA 351
29G10_mandshurica       GATCATCGGATCCGAAAAATCAAGTCCTTGTGTATAAAAGCACGTTACGGAGTGATCCCA 358
29G10_nubicola          GATCATCGGATTCGAAAAATCAAGTCCTTGTGTATAAAAGCACGTTACGGAGTGATCCCA 339
29G10_nilgerrensis      GATAATCGGATTCGAAAAATCAAGTCCTTGTGTGTAAAAGCACGTTGCGGAGTGATCCCA 351
                        *** ******* ********************* ************ *************

29G10_vesca             CTCATCAATAAGTTATCGGACTTAATTATTGTCACGGTGGACCACGTCAGTCTGGCATAT 411
29G10_mandshurica       CTCATCAATAAGTTATCGGACTTAATTATTGTCACGGTGGACCACGTCAGTCTGGCATAT 418
29G10_nubicola          CTCATCAATAAGTTATCGGACTTAATTATTGTCACGGTGGACCATGTCAGTCTGGCATAT 399
29G10_nilgerrensis      CTCATCAATAAGTTATCGGACTTAATTATTGTCACGGTGGACCACGTCAGTCTGGCATAT 411
                        ******************************************** ***************

29G10_vesca             CGATCATCACTCCCAATCTTGTCGATCATCAATTTGGCATGCATATCAGACCCAAGCCAT 471
29G10_mandshurica       CGATCATCACTCCCAATCTTGTCGATCATCAATTTGGCATGCATATCAGACCGAAGCCAT 478
29G10_nubicola          CGATCATCACTCCCAATCTTGTCGATCATCAATTTGGCATGCATATCAGACCGAAGCCAT 459
29G10_nilgerrensis      CTATCATCACTCTCAATCTTGTCGATCATCAATTTGGC-TACATATCAGACCGAAGCCAT 470
                        * ********** ************************* * *********** *******

29G10_vesca             TACTTGCTTCTATGAACGTATTTATATCATTTCTAATCACCCAGAATTATGGATAATATT 531
29G10_mandshurica       TACTTGCTTCTATGAACGTATTTATATCATTTCTAATCACCCAGAATTATGGATAATATT 538
29G10_nubicola          TACTTGCTTCTATGAATGTATTTATATCATTTCTAATCACCCAGAATTATGGATAATATT 519
29G10_nilgerrensis      TACTTGCTTCTATGAATGTATTTATATCATTACTAATCACCCAGAATTATGGATAATATT 530
                        **************** ************** ****************************
```

```
29G10_vesca             TCTTATTCACAACGACGATTGGCTTCTTGGTGTGTTGCGCTTTGTTAGGAC-AGTTCATT 590
29G10_mandshurica       TCTTATTCACAACGACGATTGGCTTCTTGGTGTGTTGCGCTTTGTTAGGAC-AGTTCATT 597
29G10_nubicola          TCTTATTCACAACGACGATTGGCTTCTTGGTGTGTTGCGCTTTGTTAGGAC-AGTTCATT 578
29G10_nilgerrensis      TCTTATTCACAACGACGATTGGCTTCTTGGTGTGTTGCGCTTTGTTAGGACCAGTTCATT 590
                        *************************************************** ********

29G10_vesca             GAATTTCAGGAATCCACAATTGGGTGCTGCCTTCTTCT 628
29G10_mandshurica       GAATTTCAGGAATCCACAATTGGGTGCTGCCTTCTTCT 635
29G10_nubicola          GAATTTCAGGAATCCACAATTGGGTGCTGCCTCT---- 612
29G10_nilgerrensis      GAATTTCAGGAATCCACAATTGGGTGCTGCCTTCTTCT 628
                        ********************************
```

32L07

```
32L07_vesca        GAGTTGAAAAACGGGTCGAATCCCGGCACCACCGTCCGCGTCGCGTAGGACTTGAATCCT 60
32L07_viridis      GAGTTGAAAAACGGGTCGAATCCCGGCACCACCGTCCGCGTCGCGTAGGACTTGAATCCT 60
                   ************************************************************

32L07_vesca        TCCAAGGTCACCTCCTTGATGTACATAGCTGCCCTCGCCGGAGAGGTGCGGACGCTAATC 120
32L07_viridis      TCCAAGGTCACCTCCTTGATGTACATAGCTGCGCTCGCCGGAGAGGTGTGGACGCTAATC 120
                   ******************************** *************** ***********

32L07_vesca        GGAAGCCGATTTTGGAGAGATTTAGTGTCGGTGATAGATCGGAACCCTAGAAATCTGAGC 180
32L07_viridis      GGTAGCCGATTTTGAAGAGATTTAGGGTCGGTGATAGATCGGAACCCTAGAAA------- 173
                   ** *********** ********** ***************************

32L07_vesca        TTCTGGTTTTTGCTTTCGGAAGTTGAGAGTCTGAAATGACATGGTTCGAATTTCTTTTTG 240
32L07_viridis      ------------------------------------------------------------


32L07_vesca        TTGTTTTCCGCTTTTTTGGTGGGTTCGAATTTTTAGACCAAGGCGGGAGATATTTGGGCC 300
32L07_viridis      ------------------------------------------------------------


32L07_vesca        AGTGATTTATATCTTGGGCTCACTCTGGGACTCATGTCTTTGGGCCTCGTCGACCTCGAG 360
32L07_viridis      ------------------------------------------------------------


32L07_vesca        GTGCTCATGAAGTCCGGCCGTCCTCAGGGTCGGAAACACCGCGGTACTACTGACTACTGT 420
32L07_viridis      ------------------------------------------------------------


32L07_vesca        GTCATCGCTTTAGAATTTCATTAATTGGCTTTGCGAGCTATAAATAATTGTGATTTGGTT 480
32L07_viridis      ------------------------------------------------------------


32L07_vesca        TGAATTTAGGTAAGTTTTAGTATTAGTATTTATCAACGGGAATTGCGGAGATGAGAAAAG 540
32L07_viridis      ------------------------------------------------------------


32L07_vesca        TTGAGGTTGATTTGGGGGAGTGTGGTGTTGTTAGTTAGTTGAATTATTAGAAACGAAAAA 600
32L07_viridis      ------------------------------------------------------------


32L07_vesca        ATAACAGAAGAATATAAATGTGGATGGATTATTGGATTAAGATTTGATTCAACGGAAGAA 660
32L07_viridis      ------------------------------------------------------------


32L07_vesca        GGAGGCGTGGTGTGTGTTTTGATAGTCTAATTTGAACTGTTTTGCTTCTGACAGCTAAAA 720
32L07_viridis      ------------------------------------------------------------


32L07_vesca        TCTATCCGGTGGTGAAAAATCAGCATCGGCTACTATGTACACTTTTAATCGGCAACGCAT 780
32L07_viridis      ------------------------------------------------------------
```

```
32L07_vesca        TAGCGATGGAGGTGACTTGTCTAATTTACTAAGTTTATTTAGGTTGTTACTGGTACATTT 840
32L07_viridis      ------------------------------------------------------------


32L07_vesca        TATGTGTTTATTGCCGTGGATGTAGTTTGTATGGGCCAGTTGACCAGCAGTTTCAAATGG 900
32L07_viridis      ------------------------------------------------------------


32L07_vesca        CAGGCCAATAGGGCCAACCTAGATTGTAGTTGAATTTTGGGAAGGAAAAAAAAAAGCAAA 960
32L07_viridis      ------------------------------------------------------------


32L07_vesca        CCAAAAGACATCACCACGAGCCACTTTGGCCTATCTATATATATTACTTCCTTGCTTAAT 1020
32L07_viridis      -----ATACGTCACCACGAGCCACTTTGGCCTATCTATGT----TACTTCCTTGCTTAAT 224
                        * ** **************************** *    ****************

32L07_vesca        GTGTTGCTCAATTGCTAAACAATATCATCAATGTCTAAAATAACGCGCCTCAAGGCTAAG 1080
32L07_viridis      GTGTTGCTCAATTGCTCAACAATATCATCAATGTCTAAAATAACGCGCCTCGAGGCTAAG 284
                   **************** ********************************** ********

32L07_vesca        GCAAGGGAAGGCGT-GCCTTAGGACGACCTCTGAAAGACATTTGATATCAAAGGTGTGAT 1139
32L07_viridis      GCAAGGGAAGGCGTCACCTTAGGACGACCTCTGAAAGACATTTGATATCAAAGGTGTGAT 344
                   **************  ********************************************

32L07_vesca        TGAGGCGCGCGATCAAGACGACGAGGTCAAGGTGCCTAATACAACATCAGGTTATAGGTT 1199
32L07_viridis      TGAGGCGCGCGATCAAGACGACGAGGTCAAGGTGCCTAATACAACATCAGGTTATAGATT 404
                   ********************************************************* **

32L07_vesca        TGAATCTCACTTTGAGAAATGTGATGGTTTGAACGGTTAAATCTATTGTCTTTTTATATT 1259
32L07_viridis      TGAATCTCACTTTGAGAAATGTGATGGTTTGAACGGTTAAATCTATTGTCTCTTTATATT 464
                   *************************************************** ********

32L07_vesca        GTATGGGCGGTAAAATTAAATGTTAAACTTCGGTAAATTGTCAAATGTTTAATAGTATAA 1319
32L07_viridis      GTATGGGTGGTAAAATTAAATGTTAAATTTCGGTAAATTGTCAAATGTTTAATAGTATAA 524
                   ******* ******************* ********************************

32L07_vesca        GAATCTACATATAGTAGGTGTAAAATAGATACCGAAATGATAATATTTTGTGAATAACGT 1379
32L07_viridis      GAATCTACATATAGTAGGTGTAAAATAGATACCGAAATGATAATATGTTGTGAATAATAT 584
                   ********************************************** **********  *

32L07_vesca        ACGTCATATGATTTAATATTAAGACTTTGTACGATTTAACGTTACACATTAAAATTGTAG 1439
32L07_viridis      ACGTCATATGGTTTAATATTAAGACTTTGTGCGATTTAATGCTACACATTAAAATTGTAG 644
                   ********** ******************* ******** * ******************

32L07_vesca        ATAAAAAGTTTATATCATCATCAACATCGATGTTCGAATAAATTTTATAACGTTCAATGC 1499
32L07_viridis      ATAAAAAATTTATATCATTATCATCATCGATGTTCGAATAAATTTTATAACGTTCAATGC 704
                   ******* ********** **** ************************************

32L07_vesca        GGTACAAATCTCCCAATGACTATTATCGAGTACAACGTCCATATCCGACACATGATATAG 1559
32L07_viridis      GGTACAAATCTCCCAATGACTATAATCGAGTACAACGTCCATATCGGACACATGATATAG 764
                   *********************** ********************* **************

32L07_vesca        GCTATCAAATTATCAAAACCCTTTGATCCGATTCTGTAGCTTTGACGACTATAAGCTTAG 1619
32L07_viridis      GCTATCAAATTATCAAAACCCTTTGATCCGATTCTGTAGCTTTGACGACTATAAGCTTAG 824
                   ************************************************************

32L07_vesca        TTAAGTTTAGTAGGACTCACCGCAATTTCGCACTAGTAGGACAAAAAGATGGTAAGATTC 1679
32L07_viridis      TTAAGTTTAGTAGGACTCACCGCAATTTCGCACTAGTAGGACAGAAAGATGGTAAGATTC 884
                   ******************************************* ****************

32L07_vesca        CTTTCATTTTTCTTCTTTACTATCCTTCTTTTCCTCAATTTTTCCCTAGAATCCTACAAC 1739
32L07_viridis      CTTTTATTTTTCTTCTTTACTATCCTTCTTTTCCTCATTTTTTCCCTAGAATCCTACAAC 944
                   **** ******************************** **********************

32L07_vesca        AAGAAAGGACTTTGGCCCCTTGTGCTCCTTTATCATCTTAAAAGCATCACCACCATCCCC 1799
32L07_viridis      AAGAAAGGACTTTGGCCCCTTGTGCTCCTTTATCATCTTAAAAGCATCACCACCATCCCC 1004
                   ************************************************************
```

```
32L07_vesca        TATATAGATGCATATTCACTATCAAGCTACCCAAGTATGCAAATTTATAGCATCTCATTA 1859
32L07_viridis      TATTTAGATGCATATTCACTATCAAGCTACCCAAGTATGCAAATTAATAGCATCTCATTA 1064
                   *** ***************************************** **************

32L07_vesca        TCTTGTTTCCTCTAGCTATTCTACTCAATGCATATCAACAACCTGACCCAGTTCTCCTAT 1919
32L07_viridis      TCTTGTTTCCTCTAGCTATTCTACTCAATGCATATCAACAACCTGACCCAGTTCTCCTAT 1124
                   ************************************************************

32L07_vesca        AATTGCTGGCAGATAGTAATACCAATTACTCCAGAATCTTCACACCCAGAACTTGAAATT 1979
32L07_viridis      AATTGCTGGCAGATAGTAATACCAATTACTCCAGAATCTTCACACCCAGAACTTGAAATT 1184
                   ************************************************************

32L07_vesca        ACACGACCTCAATACTCCAAACAGTAC---------AAAACAACCCAGATGATCAAAACA 2030
32L07_viridis      ACACGACCTCAATACTCCAAACAGTACTGTCAGTACAAAACAACCCAGATGATCAAAACA 1244
                   ***************************         ************************

32L07_vesca        CATAACATTCTTTATTTCATCTTATTGGGAAAATCTCTATATCTATTATCTTCATTATTC 2090
32L07_viridis      CTTAAAATTCTTTATTTCATCTTATTG---------CTATCTCTATCATCTTCATTATTC 1295
                   * *** *********************         **** ***** *************

32L07_vesca        AATTTTTCTACACTGCATGCTATACATGTTACAAAAGAGAAAGAAAAGACACTAGTCCAT 2150
32L07_viridis      AATTTTTCTACACTGCATGCTATACATGTTACAAAAGAGAAACAAAAGACACTAGCCCAT 1355
                   ****************************************** ************ ****

32L07_vesca        ATCACATAGGCCATGTCCTTCCCAATTCTAACCCAACAATTCAAGGACCACACCCATGAG 2210
32L07_viridis      ATCACATAGGCCATGTCCTTCCCAATTCTAACCCAACAATTCAAGGACCACACCCATGA- 1414
                   ***********************************************************

32L07_vesca        TAGTGGCACTGAATCACTGAATCGTCGCCTTCACAACTACACTACCTATCCAACCCAGAC 2270
32L07_viridis      --GTGGCACTGAATCACTGAATCGTCACCTTCACAACCACACTACCTATCCAACCCAGAC 1472
                     ************************ ********** **********************

32L07_vesca        TCAACACAGATGAAAATTCACAGCAGCTAAGAATATAGTACTAGTTTTGCTCTATCTTTT 2330
32L07_viridis      -----ACAGATGAAAATTCACAGCAGCTAAGAATATAGTACCAATTTTGCTCTATCTTTC 1527
                        ************************************ * ***************

32L07_vesca        TTCTTTACCAAAACAAAAAAAACCCTGTAGTAACCAATATAACCGCTAACAGCTTTTCCC 2390
32L07_viridis      TTCTTTACCAAAACAAAAAAGATCCTGTAGTAACTAATATAACAGCTAACAGCTTTTCCC 1587
                   ******************** * *********** ******** ****************

32L07_vesca        ATCCTGCCCATAACAGCTTTTCCCCTGCAGTATGGGAAACCCTTATCTAAAACCCCCCGA 2450
32L07_viridis      ATCCTGCCCATAACAGCTTTTCCCCTGCAGTATGGGAAACCCTGATCTAAA-TCCCCCGA 1646
                   ******************************************* *******  *******

32L07_vesca        TTTATAGTAACAAAAAAATAAATAAAATAATTTACTTTCCTCATTTACCATTTTACCCTC 2510
32L07_viridis      TTTATAGTAACAAAAAAATAAATAAAATAATTTGCTTTCCTCATTTACCATTTTACCCTC 1706
                   ********************************* **************************

32L07_vesca        ATCTTCTCCTTCATTGCCACTTGAACCCCCACTCTCCATGCTCCTTGAACCTTCTCAACA 2570
32L07_viridis      ATCTTCTCCTTCATTGCCACTTGAACCCCCACTCTCCATGCTCCTTGAACCTTCTCAACA 1766
                   ************************************************************

32L07_vesca        CCCTTTCTAGGGCAATGTCAAAAGCGTCTTTTACCGTCTCCAACCCCTCCTGCGGTTTCG 2630
32L07_viridis      CCCTTTCTAGGGCAATGTCAAAAGCGTCTTTTACCGTCTCCAACCCCTCCTGCGGTTTCG 1826
                   ************************************************************

32L07_vesca        CGTACAGAAAATTCGGTATGTAATCGATAACTTTCTCCCGCATTTTCCT 2679
32L07_viridis      CGTACAGAAAATTCGGTATGTAATCGATAACTTTCTCCCGCATTTTCCT 1875
                   *************************************************
--------------------------------------------------------------------------------------------
```

34D20

```
34D20_vesca             GCAGAAAGAAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTC 60
34D20_mandshurica       NNNNNGAAGAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGAGCAGTTC 60
34D20_nilgerrensis      NGCAGAAAGAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTC 60
34D20_iinumae           NNNNNNNNNAACTGATGTGCTTTCCGGAGGGACTGACAGTAGAAAAGGACAGTGCAGTTC 60
34D20_ananassa          NNNNNNNNNAACTGATGCGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTC 60
34D20_viridis           NNNGAGAAGAACTGATGTGCTTTCCGGAGGGACTGACAGTGGAAAAGGACAGTGCAGTTC 60
34D20_nubicola          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 60


34D20_vesca             AGGGGATAAAGGAAGTATTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCT 120
34D20_mandshurica       AGGGGATAAAGGAAGTATTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCT 120
34D20_nilgerrensis      AGGGGATAAAGGAAGTATTA-TGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCTCT 119
34D20_iinumae           AGGGGATAAAGGAAGTATTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCT 120
34D20_ananassa          ATGGGATAAAGGAAGTATTAATGTTAGGCATCCTAGACGGCATCTGGTTTTGGAGTCCCT 120
34D20_viridis           AGGAGATAAAGGAAGTATTAATGTTAGGCATCCAAGACGGCATCTGGTTTTGGAGTCCCT 120
34D20_nubicola          NNNNNNNNNNNNNNNNNNNNATGTT-GGCATCC-AGACGGCATCTGGTTTTGGAGTCCCT 118
                                             **** ******* *********************** **


34D20_vesca             CTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGAATTTGATTTCTACAAGGTGAGCAA 180
34D20_mandshurica       CTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGAATTTGATTTCTACAAGGTGAGCAA 180
34D20_nilgerrensis      CTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGAATTTGATTTCTACAAGGTGAGCAA 179
34D20_iinumae           CTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGAATTTGATTTCTACAAGGTGAGCAA 180
34D20_ananassa          CTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGAATTTGATTTCTACAAGGTGATCAA 180
34D20_viridis           CTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGAATTTGATTTCTACAAGGTGAGCAA 180
34D20_nubicola          CTCCAAGAAATGGAGCAAGTCCTACTTCCTACGCGAATTTGATTTCTACAAGGTTAGCAA 178
                        ****************************************************** * ***

34D20_vesca             CATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGTTATGTGAAT 240
34D20_mandshurica       CATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGTTATGTGAAT 240
34D20_nilgerrensis      CATGCCGGCAAACTAGTTATATTTTGTTTTTCTTACTATTACAGTGTGTGTTATGTGAAT 239
34D20_iinumae           CATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGTTATGTGAAT 240
34D20_ananassa          CATGCCTGCACACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGTTATGTGAAT 240
34D20_viridis           CATGCCTGCAAACTAGATATATTTTGTTTTTCTTACTATTACAGTGTGTGTTATGTGAAT 240
34D20_nubicola          CA-GCCTGCAAACTAGATATATTTTGTTTTTCTAACTATTACAGTGTGTATTATGTGAAT 237
                        ** *** *** ***** **************** *************** **********

34D20_vesca             CATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGA 300
34D20_mandshurica       CATCTGCAAATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGA 300
34D20_nilgerrensis      CATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGA 299
34D20_iinumae           CATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGA 300
34D20_ananassa          CATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACCAGGAAAGA 300
34D20_viridis           CATCTGCATATTATCTATATCTAACTCTATGGTATAATCATTCAGAACTACAAGGAAAGA 300
34D20_nubicola          CATCTGCATATTATCTATATCTAACTCTATGGTATAATGATTCAGAACTACAAGGAAAGA 297
                        ******** **************************** ************* ********

34D20_vesca             TTATCGGCGAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCT 360
34D20_mandshurica       TTATCGGCGAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCT 360
34D20_nilgerrensis      TTATCGGCGAGAAGGTGTTTTGCATGCAAGCAGCAGCAAATGCTATGGGCCAATTTCCCT 359
34D20_iinumae           TTATCGGCGAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCT 360
34D20_ananassa          TTATCGGCAAGAAGGTGTTTTGCATGCATGCAGCACAAAATGCTATGGGCCAATTTCCCT 360
34D20_viridis           TTATCGGCGAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCT 360
34D20_nubicola          TTATCGGCGAGAAGGTGTTTTGCATGCAAGCAGCAGAAAATGCTATGGGCCAATTTCCCT 357
                        ******** ******************* ******  ***********************

34D20_vesca             TGCAAACACTTGCTATGGTGTAATGACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTT 420
34D20_mandshurica       TGCAAACACTTGCTATGGTGTAATGACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTT 420
34D20_nilgerrensis      TGCAAACACTTGCTATGGTGTAATGACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTT 419
34D20_iinumae           TGCAAACACTTGCTATGGTGTAATGACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTT 420
34D20_ananassa          TGCAAACACTTGCTATGGTGTAATGACTGCAAGTTGCGGGCAATAGGGGCTTCGGGTCTT 420
34D20_viridis           TGCAAACACTTGCTATGGTGTAATGACTGCAAGTTACGGGCAATAGGGGCTTCGGGTCTT 420
34D20_nubicola          TGCAAACACTTGCTATGGTGTAATGACTGCAAGTTGCGGGCAATAGGGGCTTCAGGCCTT 417
                        *********************************** ***************** ** ***
```

```
34D20_vesca             TTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATGGATTTAGTACTTCCA 480
34D20_mandshurica       TTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATGGATTTAGTACTTCCA 480
34D20_nilgerrensis      TTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATGGATTTAGTACTTCCA 479
34D20_iinumae           TTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATGGATTTAGTACTTCCA 480
34D20_ananassa          TTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATGGATTTAGTACTTCCA 480
34D20_viridis           TTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATGGATTTAGTACTTCCA 480
34D20_nubicola          TTCCCTCACAAATGCGTGCTGTGTGACACAGAAGTACAGAAAATGGATTTAGTACTTCCA 477
                        ************************************************************


34D20_vesca             TTAAGTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTCCATGCACAGGAGAACTGCA 540
34D20_mandshurica       TTAAGTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTCCATGCACAGGAGAACTGCA 540
34D20_nilgerrensis      TTAACTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTCCATGCATAGGAGAACTGCA 539
34D20_iinumae           TTAACTAGTAACTGAGGAATCCAATTGCACTCTGTGTTTAC-ATGCACAGGAGAACTGCA 539
34D20_ananassa          TTAACTAGTAACTGAGGAATACAATTGCACTCTGTGTTTTCCATGCACAGGAGAACTGCA 540
34D20_viridis           TTAACTAGTAACTGAGGAATCCAATTGCAATCTGTGCTTTTCATGCACAGGAGAACTGCA 540
34D20_nubicola          TTAACTAGTAACTGAGGAATCCAATTGCAATCTGTGTTTTCCATGCACAGGAGAACTGCA 537
                        **** *************** ******** ****** **   ***** ************


34D20_vesca             GGTGAACCGTATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGT 600
34D20_mandshurica       GGTGAACCGTATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGT 600
34D20_nilgerrensis      GGTGAACCATATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGT 599
34D20_iinumae           GGTGAACTATATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGT 599
34D20_ananassa          GGTGAACTATATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGT 600
34D20_viridis           GGTGAACCATATGTCTATATAGATATGTCGAATGTTAGATAGGATACATAGTATGAGGGT 600
34D20_nubicola          GGTGAACTTTATGTCTATATAGATATGTCGTATGTTAGATAGGATACATAGTATGTGGGT 597
                        *******  ********************* ************************ ****


34D20_vesca             GTGGATGAACTATACGTAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGT 660
34D20_mandshurica       GTGGATGAACTATACGTAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGT 660
34D20_nilgerrensis      GTGGATGAACTATACGTAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGT 659
34D20_iinumae           GTGGATGAACTATAAGTAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGT 659
34D20_ananassa          GTGTATGAACTATAAGTAGAACACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGT 660
34D20_viridis           GTGGATGAACTATACGTAGAGCACCCAGAAAACCAGAAAAAGTAAAGAGGAACTGCGGGT 660
34D20_nubicola          GTGGATGAACTATACGTAGAACACCCAGAAAACCAGAGAAAGTAAAGAGGAACTGCAGGT 657
                        *** ********** ***** **************** ****************** ***

34D20_vesca             TGGGTATGAATCTCCCTCCCGGCCACACTAGACCACACTTTTGAACTGGCGGATTCCATC 720
34D20_mandshurica       TGGGTATGAATCTCCCTCCCGGCCACACTAGACCACACTTTTGAACTGGCGGATTCCATC 720
34D20_nilgerrensis      TGGGTATGAGTCTCCCTCCCGGCCACACTAGACCACACTTTTGAACTGGCGGATTCCATC 719
34D20_iinumae           TGGGCATGAGTCTCCCTCCCGGCCACACTAGACCAC------------------------ 695
34D20_ananassa          TGGGCATGAGTCTCCCTCCCGGCCACACTAGACCACACTTTTGAACTGGCGGATTCCATC 720
34D20_viridis           TGGGCATGAATCTCCCTCCCGGCCACAGTAGACCACACTTTTGAACTGGCGGATTCCATC 720
34D20_nubicola          TGGGCATGAATCTCCCTCCCGGCCACACTAGACCACACTTTTGAACTGGCGGATTCCATC 717
                        **** **** ***************** ********

34D20_vesca             CGTCCTAGATTTTGTGCCGACTATCACAATAGTGTAATTAAGTTGGTCCTCCTAGCCATA 780
34D20_mandshurica       CGTCCTAGATTTTGTGCCGACTATCACAATAGTGTAATTAAGTTGGTCCTCCTAGCCATA 780
34D20_nilgerrensis      CGGCCTAGATTTTGTGCCGACTATCACAATAGTGTA----AGTTGGTCCTCCTAGCTATA 775
34D20_iinumae           ----CTAGATTTTGTGCCGACTATCACAATAGTGAA----AGTTGGTCCTCCTAGCTATA 747
34D20_ananassa          CGGCCTAGATCTTGTGCCGACTATCACAATAGTGTA----AGTTGGTCCTCCTAGCTATA 776
34D20_viridis           CGGCCTAGATTTTGTGCCGACTATCACAATAGTGTA----AGTTGGTCCTCCTAGCTATA 776
34D20_nubicola          CGGCGTAGATTTTGTGCCGACTATCACAATAGTGTA----AGTTGGTCCTCCTAGCTATA 773
                             ***** *********************** *    **************** ***

34D20_vesca             GTTTCTAGTACTATTCTACTGATATCATGTATTGCCTCAGCTTTTGACAATGGAATATGA 840
34D20_mandshurica       GTTTCTAGTACTATTCTACTGATATCATGTATTGCCTCAGCTTTTGACAATGGAATATGA 840
34D20_nilgerrensis      GTTTCTAGTACTATTCTACTGATATCATGTTTTGTCTCAGCTTTTGACAATGGAATATGA 835
34D20_iinumae           GTTTCTAGTACTATTCTACTGATATCATGTTTCGTCTCAGCTTTTGACAATGGAATATGA 807
34D20_ananassa          GTTTCTAGTACTATTCTACTGATATCATGTTTCGTCTCAGCTTTTGACAATGGAATATGA 836
34D20_viridis           GTTTCTAGTACTATTCTACTGATATCATGTTTTGTCTCAGCTTTTGACAATGGAATATGA 836
34D20_nubicola          GTTTCTAGTACTATTCTACTGATATCATGTTTTGTCTCAGCTTTTGACAATGGAATATGA 833
                        ****************************** * * *************************
```

```
34D20_vesca             TGAATTTGGAATGAATACAAAAACTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAA 900
34D20_mandshurica       TGAATTTGGAATGAATACAAAGACTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAA 900
34D20_nilgerrensis      TGAATATGGAAC-------AAAGCTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAA 888
34D20_iinumae           TGAATATGGAATGA--ACAAAACCTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAA 865
34D20_ananassa          TGGATATGGAATGA--ACAAAACCTGCTTTGTCCATCTATTAGCATTTTCTGAAACCCAA 894
34D20_viridis           TGAATATGGAATGA--ACAAAAGCTGCTTTGTCCATCTTTTAGCATTTTCTGAAACCCAA 894
34D20_nubicola          TGAATATGGAATGA--ACAAAAGCTGCTTTGTCCATCTGTTAGCATTTTCTGAAACCCAA 891
                        ** ** *****        **  *************** *********************


34D20_vesca             AAGATGGGTACATGTTTGCTTATTCTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTC 960
34D20_mandshurica       AAGATGGGTACATGTTTGCTTATTCTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTC 960
34D20_nilgerrensis      AAGATGGGTACATGTTTGCTTATTCTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTC 948
34D20_iinumae           AAGATGGGTACATGTTTGCTTATTCTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTC 925
34D20_ananassa          AAGATGGGTACATGTTTGCTTATTTTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTC 954
34D20_viridis           AAGATGGGTACATGTTTGCTTATTCTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTC 954
34D20_nubicola          AAGATGGGTACATGTTTGCTTATTCTCTTTATCTAGTGCATCATGTGAGTTATCAAGTTC 951
                        *********************** ************************************


34D20_vesca             ATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCAGTACTTGTATAGTTGTATTGA 1020
34D20_mandshurica       ATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCAGTACTTGTATAGTTGTATTGA 1020
34D20_nilgerrensis      ATGTCTATGCATTCTGCTGATTTAGGAATAAGGATTGCAGTACTTGTATAGTTGTATTGA 1008
34D20_iinumae           ATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCACTACTTGTGTAGTTGTATTGA 985
34D20_ananassa          ATGTTTATGCATTCTGCTGATTTAGGATTTAGGATTGCACTACTTGTATAGTTGTATTGA 1014
34D20_viridis           ATGTTTATGTATTCTGCTGATTTAGGAATTAGGATTGCAGTACTTGTATAGTTGTATTGA 1014
34D20_nubicola          ATGTTTATGCATTCTGCTGATTTAGGAATTAGGATTGCAGTACTTGTATAGTT--ATTGA 1009
                        **** **** ***************** * ********* ******* *****  *****


34D20_vesca             TCTGATATAACATAAATTTAATGAATCTAATAGACATTTTT-CCTAGTTAACAGAGGATA 1079
34D20_mandshurica       TCTGATATAACATAAATTTAATGAATCTAATAGACATTTTT-CCTAGTTAACAGAGGATA 1079
34D20_nilgerrensis      TCTGATATAACATAAATTTAATGAATCTAATATAAATTTTT-CCTAGTTAAG-------- 1059
34D20_iinumae           TCT------------------------------AAATTTTT-CCTAGTTAACAGAGGATA 1014
34D20_ananassa          TCT------------------------------AAATTTTT-CCTAGTTAACAGAGGATA 1043
34D20_viridis           TCTGATATAACTCCCATT-AATGAATCTAATATAAATTTTTTCCTAGTTAACAGAGGATA 1073
34D20_nubicola          TCTAATATAACATAAATTTAATGAATCTAATATAAATTTTT-CCTAGTTAAG-------- 1060
                        ***                              * ****** *********

34D20_vesca             GGTCTCCGGCTGACCTTATCCTACAAGGAAATAGAAACGTACAATTAACGCATTATACAC 1139
34D20_mandshurica       GGCCTCCGGCTGACCTTATCCTACAAGGAAATAGAAACGTACAATTAACGCATTATACAC 1139
34D20_nilgerrensis      -----CCAAAAGACCTTATCCTACAAGGAAACAGAAACGTACAATTAACGCATTATACAC 1114
34D20_iinumae           GGTCTCCGGCTGACGTTATCCTACAAGGAAACAGAAACGTACAATTAACGGATT---CAC 1071
34D20_ananassa          GGACTCCGGCTGACCTTATCCTACAAGGAAACAGAAACGTACAATTAACGGATT---CAC 1100
34D20_viridis           GGTCTCCGGCTGACCTTATCCAACAAGGAAACAGAAACATACAATTAACGCATTATACAC 1133
34D20_nubicola          -----TCAAAAGACCTTATCCTACAAGGAAACAGAAACGTACAATTAACGCATTATACAC 1115
                              *    *** ****** ********* ****** *********** ***   ***

34D20_vesca             AAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGTTTCATTGATCATATTGTCCT 1199
34D20_mandshurica       AAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGTTTCATTGATCATATTGTCCT 1199
34D20_nilgerrensis      AAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGT-------------------- 1154
34D20_iinumae           AAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGT-------------------- 1111
34D20_ananassa          AAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGT-------------------- 1140
34D20_viridis           AAGACTGGTCTATATAAGGCATCAAATTCTCTTTATCTGT-------------------- 1173
34D20_nubicola          AAGACCGGTCTATATAAGACATCAAAATCTCTTTATCTGT-------------------- 1155
                        ***** ************ ******* *************

34D20_vesca             CTTTATCTGTTTCATACTTTCAT--TGATCATATTGTCTAGTACTGGAAGAGCTATATTT 1257
34D20_mandshurica       CTTTATCTGTTTCATACTTTCAT--TGATCATATTGTCTAGTACTGGAAGAGCTATATTT 1257
34D20_nilgerrensis      ------------------TTCAT--TGATCATATTGTCTAGTACTGGAAGAGTTATATTT 1194
34D20_iinumae           ------------------TTCAT--TGATCATATTGTCTAGTACTGGAAGAGCTATATTT 1151
34D20_ananassa          ------------------TTCAT--TGATCATATTGTCCAGTACTGGAAGAGCTATATTT 1180
34D20_viridis           ------------------TTCAT--TGATCATATTGTCTAGTACTGGAAGAGCTATATTC 1213
34D20_nubicola          ------------------TTCATATTGATCATATTGTCTAGTACTGGAAGAGCTATATTT 1197
                                          *****  ************* ************* ******
```

```
34D20_vesca             ATCAGATAACAGAAAGTGCTTACTTGCTGGTTCATACTCAATATGGATCCGAAGGTCCTT 1317
34D20_mandshurica       ATCAGATAACAGAAAGTGCTTACTTGCTGGTTCATACTCAATATGGATCCGAAGGTCTTT 1317
34D20_nilgerrensis      ATCAGATAACAGAAAGTGCTTACTTGCTGGTTCATACTCAATATGGATCCGAAGGTCCTT 1254
34D20_iinumae           ATCTGATAACAGAAAGTGCTTACTTGCTGGTTCCTACTCATTATGGATCCGAAGGTCCTT 1211
34D20_ananassa          ATCTGATAACAGAAAGTGCGTACTTGCTGGTTCATACTCAATATGGATCCGAAGGTCCTT 1240
34D20_viridis           ATCTGATAACAGAAAGTGCTTACTTGCTGGTTCATACTCAATATGGATCCGAAGGTCCTT 1273
34D20_nubicola          ATCTGATAACAGAAAGTGCTTACTTGCTGGTTCATACTCAATATGGATCCGAAGGTCCTT 1257
                        *** *************** ************* ****** **************** **


34D20_vesca             AGTTACAATGGTGTTGACCTGAGCATGAGCGACTTGGATCTTCTTAGAGGCCCTTGTTAC 1377
34D20_mandshurica       AGTTACAATGGTGTTGACTTGAGCATGAGCGACTTGGATCTTCTTAGAGTCCCTTGTTAC 1377
34D20_nilgerrensis      AGTTACAATGGTGTTGACCTGAGCATGAGCGACTTGGATCTTCTTAGAGGCCCTTGTTAC 1314
34D20_iinumae           AGTTACAAAGGTGTTGACCTGAGCATGAGCGACTAGGATATTCTTAGAGGACCTTATTAC 1271
34D20_ananassa          AGTTCCCAAGGTGTTGACCTGAGAATGAGCGACTTGGATCTTCTTAGAGGCCCTTATTAC 1300
34D20_viridis           AGTTACAATGGTGTTGACCTGAGCATGAGCGACTTGGATCTTCTTAGAGGCCCTTATTAC 1333
34D20_nubicola          AGTTACAAAGGTGTAGACCTGAGCATGAGCGACTTGGATCTTCTTAGAGGCCCTTGTTAC 1317
                        **** * * ***** *** **** ********** **** *********  **** ****


34D20_vesca             TTAACCGATAGCATCATTTGATTCTATTTCACTCATCTTACTTCCCATTATGATGATGAT 1437
34D20_mandshurica       TTAACCGATAGCATCATTCGATTCTATTTCACTCATCTTACTTCCCATTATGATGATGAT 1437
34D20_nilgerrensis      TTAACCGATAGCATCATT-----------CACTTATCTTACTTCCCATTACGATGATGAT 1363
34D20_iinumae           TTAACCGATAGCATCATTCGATTCTATTTCACTTATCTTACTTCCCATTATGATGATGAT 1331
34D20_ananassa          TTAACCGATAGCATCATTCAATTCTATTTCACTTATCTTACTTCCCATTATGATGATGAT 1360
34D20_viridis           TTAACCGATAGCATCATTCGATTCTATTTCACTAATCTTAGTTCCCATTATGATGATGAT 1393
34D20_nubicola          TTAACCGATAGCATCATTCGATTCTATTTCACTCATCTTACTTCCCATTATGATGATGAT 1377
                        ******************           **** ****** ********* *********


34D20_vesca             ATCCTTCTGGTTTCCCCTAATATCTCTGATCTTCTGGTAAATTCTCCGGATCCCGAGGAT 1497
34D20_mandshurica       ATCCTTCTGGTTTCCCCTAATATCTCTGATCTTCTGGTAAATTCTCCGGATCCCGAGGAT 1497
34D20_nilgerrensis      ATCCTTCTGGTTTCCCCTAATATCTCTGATCTTCTGGTAAATTCTCTGGATCCTGAGGAT 1423
34D20_iinumae           ATCCTTCTGGTTTCCCCTAATATCTCTGATCTCCTGGTAAATTCTCCGGATCCCGAGGAT 1391
34D20_ananassa          ATCCTTCTGGTTTCCCCTAATATCTCTGATTTTCTGGTAAATTCTCCGGATCCCGAGGAT 1420
34D20_viridis           ATCCTTCTGGTTTCCCCTAATATCTCTGATCTTCTGGTAAATTCTCCGGATCCCGAGGAT 1453
34D20_nubicola          ATCCTTCTGGTTTCCCCTAATATCTCTGATCTTCTGGTAAATTCTCCGGATCCCGAGGAT 1437
                        ****************************** * ************* ****** ******

34D20_vesca             GAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATCTTCGCA 1557
34D20_mandshurica       GAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATCTTCGCA 1557
34D20_nilgerrensis      GAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATCTTCGCA 1483
34D20_iinumae           GAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATCTTCGCA 1451
34D20_ananassa          GAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATCTTCGCA 1480
34D20_viridis           GAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGGAAAGTTGTGATATTTGCA 1513
34D20_nubicola          GAGCTTAGAGCCTTTGCGGAGTCTGACCAACTTGGTAAAAGAAAAGTTGTGATCTTTGCA 1497
                        **************************************** ************ ** ***

34D20_vesca             GTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG 1617
34D20_mandshurica       GTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG 1617
34D20_nilgerrensis      GTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG 1543
34D20_iinumae           GTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG 1511
34D20_ananassa          GTGAATGATAACAAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG 1540
34D20_viridis           GTGAATGATAACGAAGATCCGAGTCGAAGCGACGGCGGAAACCATTGGAGCTTGCTGGTG 1573
34D20_nubicola          GTGAATGATAACGAAGATCCGAGTCGAAGCGACGGTGGAAACCATTGGAGCTTGCTGGTG 1557
                        ************ ********************** ************************

34D20_vesca             TATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTG 1677
34D20_mandshurica       TATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTG 1677
34D20_nilgerrensis      TATTTCAGAAAATCAAACGCATTCGTACACTACGACAGCTTGGGGGGTAACAATAGTTTG 1603
34D20_iinumae           TATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTG 1571
34D20_ananassa          TATTTCAGAAAATCAAACGCATTGGTACATTACGACAGCTTGGGGGGTAACAATAGTTTG 1600
34D20_viridis           TATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTG 1633
34D20_nubicola          TATTTCAGAAAATCAAACGCATTCGTACATTACGACAGCTTGGGGGGTAACAATAGTTTG 1617
                        *********************** ***** ******************************
```

```
34D20_vesca             GAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCA 1737
34D20_mandshurica       GAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCA 1737
34D20_nilgerrensis      GATGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCAGCTCCAGCAACACAAGCA 1663
34D20_iinumae           GAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCA 1631
34D20_ananassa          GATGCTAGGAAAATGTATACAGCATTCAAGAAACTTGTGGCAGTTCCAGCAACACAAGCA 1660
34D20_viridis           GAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCA 1693
34D20_nubicola          GAAGCTAGGAAAATGTATACAGTATTCAAGAAACTTGTGGCTGCTCCAGCAACACAAGCA 1677
                        ** ******************* ****************** * ****************

34D20_vesca             CCAATAACTCCAGCTGGGACTAGTAGTTTGGCTACCAACAACAGTTCTACAATGAGACAC 1797
34D20_mandshurica       CCAATAACTCCAGCTGGGACTAGTAGTTTGGCTACCAACAACAGTTCTACAATGAGACAC 1797
34D20_nilgerrensis      CCAA---CTCCATCTGGGACTAGTAGTTTGGTTACCAACAACAGTTCTACAATGCGACAC 1720
34D20_iinumae           CCAATAACTCCAGCTGGGACTAGTAGTTTGGCTACCAACAACAGTTCTACAATGAGACAC 1691
34D20_ananassa          CCAA---CTCCAGCTGGGAGTAGTAGTTTGGTTACCAACAACAGTTCTACAATGGGACAC 1717
34D20_viridis           CCAATAACTCCAGCTGGGACTAGTAGTTTGGTTACCAACAACAGTTCTACAATGAGACAC 1753
34D20_nubicola          CCAATAACTCCAGCTGGGACTAGTAGTTTGGTTACCGACAACAGTTCTACAATGAGACAC 1737
                        ***    ***** ****** *********** **** ***************** *****

34D20_vesca             GAGTGCCACTCTACGCAGTCGCGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTT 1857
34D20_mandshurica       GAGTGCCACTCTACGCAGTCGGGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTT 1857
34D20_nilgerrensis      GAGTGCTACTCTACGCGGTCGCGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTT 1780
34D20_iinumae           GAGTGCCACTCTACGCAGTCGCGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTT 1751
34D20_ananassa          GAGTGCTACTCTACGCAGTCGCGGCGGTTTATAGACCATACCAAGATAATGCTTCGGGTT 1777
34D20_viridis           GAGTGCCACTCTACGCAGTCGCGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTT 1813
34D20_nubicola          GAGTGCCACTCTACGCAGTCGCGGCGATTTATAGACTATACCAAGACAATGCTTGGGGTT 1797
                        ****** ********* **** **** ********* ********* ******* *****

34D20_vesca             TGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCAT 1917
34D20_mandshurica       TGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCAT 1917
34D20_nilgerrensis      TGGGGTTTTGTTGTCAATTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCAT 1840
34D20_iinumae           TGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCAT 1811
34D20_ananassa          GGGGGTTTTGTTGTCAAATACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCAT 1837
34D20_viridis           TGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCAT 1873
34D20_nubicola          TGGGGTTTTGTTGTCAACTACATTTTGTCAAAGTACTTGCGTCTGTTTGGAAATTATCAT 1857
                         **************** ******************************************

34D20_vesca             TATCATCCTTCGGAAGTGTGCTATCCCATGCAAAAAATCACCAATAGTA---ATCATGGA 1974
34D20_mandshurica       TATCATCCTTCGGAAGTGTGCTATCCCATGCAAAAAATCACCAATAGTA---ATCATGGA 1974
34D20_nilgerrensis      TATC---CTTCGAAAGTGTGTTATCCCATGCAAAAGATCACCAATAGTAGTAATCATGGA 1897
34D20_iinumae           TATCATCCTTCGGAAGTGTGCTATCCCATGCAAAAAATCACCAATAGTA---ATCATGGA 1868
34D20_ananassa          CATCATCCTTCGGAAGTGTGTTATCCCATGCAAAAGATCACCAATAGTA---ATCATGGA 1894
34D20_viridis           TATCATCCTTCGGAAGTGTGTTATCCCATGCAAAAAATCACGAATGGTA---ATCATGGA 1930
34D20_nubicola          TATCATCCTTCGGAAGTGTGTTATCCCATGCAAAAAATCACCAATAGTA---ATCATGGA 1914
                         ***   ***** ******* ************** ***** *** ***   ********

34D20_vesca             GATGATGA---------------TGATGTTAATGAACCTTGGTATAGAGAAG-AGACTCT 2018
34D20_mandshurica       GATGATGA---------------TGATGTTAATGAACCTTGGTATAGAGAAG-AGACTCT 2018
34D20_nilgerrensis      GATGATGAAGAGCGTAATGATGATGATGTTAATGAACCTTGGTATAGAGAAG-AGACCGT 1956
34D20_iinumae           GATGATGA---------------TGATGTTAATGAACCTTGGTATAGAGAAG-AGACTCT 1912
34D20_ananassa          GATGATGAAGAGCTTAATGATGATGATGTTAATGAACCTTGGTATAGAGAAG-AGACTCT 1953
34D20_viridis           GGTGATGA---------------TGATGTTAATGAACCTTGGTATAGAGAAG-AGACTCT 1974
34D20_nubicola          GATGATGA---------------TGATGTTAATGAACCTTGGTATAGAGAAGGAGACTCT 1959
                        * ******               ***************************** ****  *

34D20_vesca             TATGCCTCAGCAGACGAATTTTTACGACTGCG 2050
34D20_mandshurica       TATGCCTCAGCAGACGAATTTT-ACGAGTGCG 2049
34D20_nilgerrensis      TATGCCTCAGCAGACGAATTT--ACGACTNNN 1986
34D20_iinumae           T-TGCCTCAGCAGACGAATTT--ACGACTNNN 1941
34D20_ananassa          TATGCCTCGGCAGACGAATTTT-ACGACTGCG 1984
34D20_viridis           TATGCCTCAGCAGACGAATTT--ACGAC-GCG 2003
34D20_nubicola          TATGCCTCAGCAGACGAATT---ACGAC-GCA 1987
                        * ****** ***********   ****
```

40M11

```
40M11_vesca           CAACATTTTGGTGGCCTTCTTGACATTCCAGTTTCTGGCCCTCAGATGCCTTGCAATGGA 60
40M11_mandshurica     CAACATTTTGGTGGCCTTCTTGACATTCCAGTTTCTGGCCCTCAGATGCCTTGCAATGGA 60
                      ************************************************************

40M11_vesca           TGCATCAGAACAGTATGTGGACAGCTTCTCGGGTACTGCCTTTAACAATTTTCTCACCTC 120
40M11_mandshurica     TGCATCAGAACAGTATGTGGACAGCTTCTCGGGTACTGCCTTTAACAATTTTCTCACCTC 120
                      ************************************************************

40M11_vesca           ATTAATCTGCAAACAATAAGATTTTTTAGGCAAAGCAGAACTATGAGTTCCCCAAACTAA 180
40M11_mandshurica     ATTAATCTGCAAACAATAAGATTTTTTAGGCAAAGCGGAACTATGAGTTCCCCAAACTAA 180
                      ************************************ ***********************

40M11_vesca           TAGCTTTCAAACAAGTAGAGGAGCACATTTACTAAAGATACCTTTGCCTGCTGCTCTTCA 240
40M11_mandshurica     TAGCTTTCAAACAAGTAGAGGAGCACATTTACTAAAGATACCTTTGCCTGCTGCTCTTCA 240
                      ************************************************************

40M11_vesca           CTTGTTAAAATACTCTCAGAGCCATTTGAGGAAGATTTTTTTATTCCCGCACTCATAGTT 300
40M11_mandshurica     CTTGTTAAAATACTCTCAGAGCCATTTGAGGAAGATTTTTTTATTCCCGCACTCATAGTT 300
                      ************************************************************

40M11_vesca           TTGAGGGGAAACTCTGCAAATCAACAATGGAGATTTCAAAACTTATGTCCTAGTTTCACA 360
40M11_mandshurica     TTGAGGGGAAACTCTGCAAATCAACAATGGAGATTTCAAAACTTATGTCCTAGTTTCACA 360
                      ************************************************************

40M11_vesca           GTTCCCTTCGGTCTCCCATCACCATCAAATACAATAAATTTCAATATATTTAACAAAAAA 420
40M11_mandshurica     GTTCCCTTCGGTCTCCCATCACCATCAAATACAATAAATTTCAATATATTTAACAAAAAA 420
                      ************************************************************

40M11_vesca           ATTGCTCTTCATCCCACAAAACACAGAGTCCTCATCTTCATTGTTCAATATATCATTTGA 480
40M11_mandshurica     ATTGCTCTTCATCCCACAAAACACAGAGTCCTCATCTTCATTGTTCAATATATCATTTGA 480
                      ************************************************************

40M11_vesca           AATTAACAACTTTTATTCTTCTAGTCAACCACATTTCGCAGCTACTTGTTTAACTCATAA 540
40M11_mandshurica     AATTAACAACTTTTATTCTTCTAGTCAACCACATTTTGCAGCTACTTGTTTAACTCATAA 540
                      ************************************ ***********************

40M11_vesca           ACCCTTTCTTCCGATCCATAGCTATCAAATATCCAATCTAAACGAGACTACTACTTTGTT 600
40M11_mandshurica     ACCCTTTCTTCCGATCCATAGCTATCAAATATCCAATCTAAACGAGACTACTACTTTGTT 600
                      ************************************************************

40M11_vesca           CACAACGAATCCAACACAAAAGGATCAAAAAAACCATCCAAAACTCATGCACAACATAAT 660
40M11_mandshurica     CACAACGAATCCAACACAAAAGGATCAAAAAAACCATCCAAAACTCATGCACAACATAAT 660
                      ************************************************************

40M11_vesca           CAACCAAATATTTTAACCACAAAAACAAGCACAATTCTCCAAAGTACAAAAAGAAATGGG 720
40M11_mandshurica     CAACCAAATATTTTAACCACAAAAACAAGCACAATTCTCCAAAGTACAAAAAGAAATGGG 720
                      ************************************************************

40M11_vesca           CTTTAGACACCAGGAAGGCATATCAAACCGGCCCACACACGTTAAAGGGATACAAAGATC 780
40M11_mandshurica     CTTTAGACACCAGGAAGGCATATCAAACCGGCCCACACACGTTAAAGGGATACAAAGATC 780
                      ************************************************************

40M11_vesca           TCACCTGGACCAAAGACAGAACTGGGTGGTTGCTGACTGAGCAAAGCCAATATCTCGGAG 840
40M11_mandshurica     TCACCTGGACCAAAGACAGAACTGGGTGGTTGCTGACTGAGCAAAGCCAATATCTCGGAG 840
                      ************************************************************

40M11_vesca           CTCCTCAGATGTCGGAGAGACCCATCTGAACCCAAGTCAACTGCACTGTTACAGCAACTA 900
40M11_mandshurica     CTCCTCAGATGTCGGAGAGACCCATCTGAACCCAAGTCAACTGCNNNNNNNNNNNNNNNN 800
                      ********************************************

40M11_vesca           CAAAACGCAAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAGATGAG 960
40M11_mandshurica     NNNNNNNNNNNNNNNNNNNNNNNN------------------------------------ 924
```

```
40M11_vesca           TCACCCGACTGATCAAAGTCGTCCAAAGAATTCTCAAACTGTGGTTTCATAGGCCG 3116
40M11_mandshurica     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN ~3000
```

63F17

```
63F17_vesca           CGCTCTATGGAAGGGACAAGAGACACTGAAATAGCAATGGGG-TCCTACCAACCTCATCA 59
63F17_mandshurica     CGCTCTATGGAAGGGACAAGAGACACTGAAATAGCAATGGGGGTCCTACCAACCTCATCA 60
63F17_viridis         CGCTCTATGGAAGGGACAAGAGACACTGAAATAGCAATGGGGGTCCTACCAACCTCATCA 60
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 60


63F17_vesca           TACATGGGCAAGAAATCATTCTAGTCCTCTCGGACAGGTAATCACAGAATCCAGATTAGA 119
63F17_mandshurica     TACATGGGCAAGAAATCATTCTAGTCCTCTCGGACAGGTAATCACAGAATCCAGATTAGA 120
63F17_viridis         TACATGGGCAAGAAATCATTCTAGTCCTCTCGGACAGGTAATCACAGAATCCAGATTATA 120
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 120


63F17_vesca           TGCAGGTTTTGAATTATTAGAGTCTATAAAGGGACATAGTTACAACTGTTTGTATGCTTT 179
63F17_mandshurica     TGCAGGTTTTGAATTATTAGAGTCTATAAAGGGACATAGTTACAACTGTTTGTATGCTTT 180
63F17_viridis         TGCCGGTTTTGAATTATTAGAGTCTATAAAGGGAC-TAGTTACAACTGTTT------TTT 173
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180


63F17_vesca           TCCA--TTTTTTTTATTTTTTTATTTTTTGAGAATGTATGCTTTTTCACTTGTATGGCCT 237
63F17_mandshurica     TCCA--TTTTTTT-ATTTTTTTATTTTTTGAGAATGTATGCTTTTTCACTTATATGGCCT 237
63F17_viridis         TCCACTTTTTTTTTTTTTTTTTTTTTTTTGAGAATGTATGCTTTT-CACTTATATGGCCT 232
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240


63F17_vesca           GAAGTTGCGAATGTTTTGGTTGATAGATATTTGGATATAGAATGTCACTATGGGCAGAGC 297
63F17_mandshurica     GAAGTTGCGAATGTTTTGGTTGATAGATATTTGGATATAGAATGTCACTATGGGCAGAGC 297
63F17_viridis         GAAGTTGCGAATGTTTTGGTTGATAGATATTTGGATATAGAATGTCACTATGGGCAGAGC 292
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300


63F17_vesca           ACACAGGAACCGTTGAGGACTGTTTTAGAGAACCAGAGAGTCTTGAATGTGTTAGGAGAG 357
63F17_mandshurica     ACACAGGAACCGTTGAGGACTGTTTTAGAGAACCAGAGAGTCTTGAATGTGTTAGGAGAG 357
63F17_viridis         ACACAGGAACCGTTGAGGACTGTTTTAGAGAACCAGAGAGTCTTGAATGTGTTAGGAGAG 352
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360


63F17_vesca           TTAGAGCAATGGGTGAGATGAACTGGAAACAATTTGCTGCTGAGGAGGTTACAGAGATGA 417
63F17_mandshurica     TTAGAGCAATGGGTGAGATGAACTGGAAACAATTTGCTGCTGAGGAGGTTACAGAGATGA 417
63F17_viridis         TTAGAGCAATGGGTGAGATGAACTGGAGACAATTTGCTGCTGAGGAGGTTACAGAGATGA 412
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420


63F17_vesca           GGGGTCATCTATTGAAGTATCCAGTTGAAATTGATCGAAAAGGCAAAGTCACATCCCTTC 477
63F17_mandshurica     GGGGTCATCTATTGAAGTATCCAGTTGAAATTGATCGAAAAGGCAAAGTCACATCCCTTC 477
63F17_viridis         GGGGTCATCTATTGAAGTATCCAGTTGAAATTGATCGAAAAGGCAAAGTCACATCCCTTC 472
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480


63F17_vesca           CTGGATGTGAGAGTTTCCCCGATGCAGGAGGAAATATAACCGGTTCTTTCCTTGGCATTC 537
63F17_mandshurica     CTGGATGTGAGAGTTTCCCCGATGCAGGAGGAAGTATAACCGGTTCTTTCCTTGGCATTC 537
63F17_viridis         CTGGATGTGAGAGTTTCCCCGATGCAGGAGGAAATATAACCGGTTCTTTCCTTGGCATTC 532
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540


63F17_vesca           AAGAAAATTTGACAATTTGATCACCAGTTCAGTTTTATAGAAGAACTCAGTTAGTACAGT 597
63F17_mandshurica     AAGAAAATTTGACAATTTGATCACCAGTTCAATTTTATAGAAGAACTCAGTTAGTACAGT 597
63F17_viridis         AAGAAAATTTGACAATTTGATCACCAGTTCAGTTTTATTGAAGAACTCGGT----ACAGT 588
63F17Rrc_ananassa     NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600
```

```
63F17_vesca            TTTGAAACGTTTTTTTGTTGTATTTAGCAAACCCATAGGAGGATAGGGTTTTCTTTTATT 657
63F17_mandshurica      ATTGAAACGTTTTTTCGTTGTATTTAGCAAACCCATAGGAGGATAGGGTTTTCTTTTATT 657
63F17_viridis          TTTGAAACGTTTTTTTGTTGTATTTAGCAAACGCATAGGAGGATAGGGTTTTCTTTTATT 648
63F17Rrc_ananassa      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 660


63F17_vesca            CAACAGGGATATAGGCGCTTTTAGGGTTTCTTTTCCTATTCAATTTCGTTCTTTGGTAGA 717
63F17_mandshurica      CAACAGGGATATAGGCGCTTTTAGGGTTTCTTCTCCTATTCAATTTCATTCTTTGGTAGA 717
63F17_viridis          CAACAGGGATATAGGCGCTTTTAGGGTTTCTTTTCCTATTGAATTTCGTTCTTTGGTATA 708
63F17Rrc_ananassa      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTCAATTTCATTCTGGGATAGA 720
                                                             ** ****** ****  * ** *

63F17_vesca            CCAAGTCGCTTCTTTGGCATTCAAGGAAACCTGAGCATTTGATCTGCCTGTCATCACATC 777
63F17_mandshurica      CCAAGTCGCTTCTTTGGCATTCAAGGAAACCTGAGCATTTGATCTGCCTGTCATCACATC 777
63F17_viridis          CCAAGTCCCTTCTTTGGCATTCAAAGAAACCTAAGCATTTGATCTGCCTGTCATCACGTC 768
63F17Rrc_ananassa      CCAAGTCACTTTTTTGGCGTTCAAGGAAACCTGAGCATTTGATCTGCCGGTCATCACATC 780
                       ******* *** ****** ***** ******* *************** ******** **

63F17_vesca            CAGAGTTGCAGATTGTTTAGAGAAGAATTCCAATAAATTCCTTTTGTACAGTTTGGTTAA 837
63F17_mandshurica      CAGAGTTGCAGATTGTTTAGAGAAGAATTCCAAT----TCCTTTTGTACAGTTTGGTTAA 833
63F17_viridis          TAGAGTTGCAGATTGTTTAGAGAAGAATTCCAATAAATTCCTTTTGTACAGTTTGGTTAA 828
63F17Rrc_ananassa      CAGAGTTGCAGATTGTTTATAGAAGAATTCCAATAAATTCCTTTTGTACAGTTTGGTTAA 840
                        ****************** **************    **********************

63F17_vesca            CTTTTGGTATTCAACAACGCATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGAT 897
63F17_mandshurica      CTTTTGGTGTTCAACAACGCATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGAT 893
63F17_viridis          CTTCTGGTGTTCAACAACGCATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGAT 888
63F17Rrc_ananassa      CTTTTGGTGTTCAACAACGCATTGTACAACTCTGCCAATTTGGCACATTATAATGTTGAT 900
                       *** **** ***************************************************

63F17_vesca            ATGCAGGTAACATCTCTGACTATGCATCTTTGCTTTTTCTTCTTTTTTTGAGAACAAGGC 957
63F17_mandshurica      ATGCAGGTAACATCTCTGACTATGCATCTTTGCTTTTTCTTCTATTTTTGAGAACAAGGC 953
63F17_viridis          ATGCAGGTAACATCTCTGACTATGCATCTTTGCTTTTTCTTCTGTTTTTGAGAACAAGGC 948
63F17Rrc_ananassa      ATGCAGGTAACATCTCTGACTATGCATCTTTGCTTTTTCTTCTTTTTTTGAGAATAAGGC 960
                       ******************************************* ********** *****

63F17_vesca            ATCTTGTTTATGTGTAGCCAACTTGAAGCACTGTATTTAAATAATGCTAAAACAGTGTTA 1017
63F17_mandshurica      ATCTTGTTTATGTGTGGCCAACTTGAAGCACTGTATTTAAATAATGCTAAAACAGTGTTA 1013
63F17_viridis          ATCTTGTTTATTTGTGGCCAACTTGAAGCACTGTATTTAAATAATGCTAAGACCGTGTCA 1008
63F17Rrc_ananassa      ATCTTGTTTATGTGTAGCCAACTTGAAGCACTGTATTTAAATAATGCTAAAACAGTGTTA 1020
                       *********** *** ********************************** ** **** *

63F17_vesca            ATTTTGTTATAAAAGTGTAGGCAACAATGAACTTGAACTTGAACTTGAACTTGATAAGAA 1077
63F17_mandshurica      ATTTTGTTATAAAAGTGTAGGCAACAATGAACTTGAACTTGAAC------TTGATAAGAA 1067
63F17_viridis          ATTTTGTTACAAAAGTCTAGGCAACAATGAACTTGAACTT------------GATAAGAA 1056
63F17Rrc_ananassa      ATTTTGTTATAAAAGTGTAGGCAACAATGAACTTGAACTTGAACTTGAACTTGATAAGAA 1080
                       ********* ****** ***********************            ********

63F17_vesca            AATAGATCCCAGAGATGGTCTATCTACTACCCTTGACTACACAAGTTACTCATTCTTTAC 1137
63F17_mandshurica      AATAGATCCTAGAGATGGTCTATCTACTACTCTTGACTACACAAGTTACTCATTCTTTAC 1127
63F17_viridis          AATAGATCCTAGAGATGGTCTATCTACTACCCTTGACTACACAAGTTGCTCATTCTTTAC 1116
63F17Rrc_ananassa      AATAGATCCCAGAGATGGTCAATCTACTACCCTTGACTACACAAGTTACTCATTCTTTAC 1140
                       ********* ********** ********* **************** ************

63F17_vesca            ATGTGAAATGGCTATCCAGAGCAGTCGTA--TTTCATGAGATATTAACAAGCTTTGGACG 1195
63F17_mandshurica      ATGTGAAATGGCTATCCAGAGCAGTCGTAAATTTCATGAGATATTAACAAGCTTTGGACG 1187
63F17_viridis          ATGTGAAATGGCTATCCAGAGCAGTCGTAAATTTCATGAGATATTAACAAGCTTTGGACG 1176
63F17Rrc_ananassa      ATGTGAAATGGCTATCCAGAGCAGTCGTA--TTTCATGAGATATTAACAAGCTTTGGACG 1198
                       ****************************  ******************************

63F17_vesca            TCCAGTTGCCAGGTTCATCTCTACTCGGAGGCCAAGTCGAGCAAGGGCAGGCACATCAAC 1255
63F17_mandshurica      TCCAGTTGCCACGTTCATCTCTACTCGGAGGCCAAGTCGAGCAAGGGCAGGCACATCAAC 1247
63F17_viridis          TCCAGTTGCCAGGTTCATGTCTACTCGGAGACCAAGT-GAGCAAGGGCAGGCACATCAAC 1235
63F17Rrc_ananassa      TCCAGTTGCCAGGTTCATCTCTACTCGGAGGCCAAGTCGAGCAAGGGCAGGCACATCAAC 1258
                       *********** ****** *********** ****** **********************
```

```
63F17_vesca            AGACCCCTTAA 1266
63F17_mandshurica      AGACCC----- 1253
63F17_viridis          AGACCCCTAAA 1246
63F17Rrc_ananassa      AGACCCTCAA- 1268
                       ******
```

72E18

```
72E18_vesca             GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAAATCCGGCCTAAGCATTAATATC 60
72E18_mandshurica       GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAGACCCGGCCTAAACATTAACATC 60
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 60
72E18_viridis           NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTAATATC 60
72E18_iinumae           GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAAACCCGGCCTAAACATTAACATC 60
72E18_ananassa          GCTAGGGAAAACAGCTCGTGGAGCATCATCTCCAGCAAACCCGGCCTAAACATTAACATC 60


72E18_vesca             AAA-TCAGTCCTTGAGATTCGACATGCATAAAAAAGACAATAAAGGGTACAAAAACAACC 119
72E18_mandshurica       AAAATCAGTCCTTGAGATTCAACATGCATAACAAAGACAATAAAGGGTACAAAAACAACC 120
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 120
72E18_viridis           AAAATCAGTCCTTGAGATTCAACATGCATAACAAAGACAATAAACGGTACAAAAACAACC 120
72E18_iinumae           AAAATCAGTCCTTGAGATTCGACATGCATAAAAAAGACAATAAAGGGTACAAAAACAACC 120
72E18_ananassa          AAAATCAGTCCTTGAGATTCAACATGCATAAAAAAGACAATAACGGGTACAAAAACAACC 120


72E18_vesca             ACTCAAACAATCACAACATAATATCATTCAATACCTTGACCATTCCGGTTCCATTATCAC 179
72E18_mandshurica       ACTCAAACAATCACAACATAATATCATTCAATACCTTGACCATTCCGGTTCCATTATCAC 180
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
72E18_viridis           ACTCAAACAATCACAACATAATATCATTCAATACCTTGACCATTCCGGTTCCATTATCAC 180
72E18_iinumae           ACTCAAACAATCACAACATAATAGCATCCAATACCTTGACCATTCCGGTTCCATTATCAC 180
72E18_ananassa          ACTCAAACAATCACAACATAATATCATTCAATACCTTGACCATTCCGGTTCCATTATCAC 180


72E18_vesca             ACACAAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAACATACACCAC 239
72E18_mandshurica       ACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAACATACACCAC 240
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
72E18_viridis           ACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAACATACACCAC 240
72E18_iinumae           ACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCTACATACACCAC 240
72E18_ananassa          ACACGAGCGGCTGAATGTCCTCGGTTTCTGCCATCTTCTTCTACCTGCAACATACATCAC 240


72E18_vesca             AATCAATGACAACAATGCCTCATTCACACAACAAAGAAATAGACATTCAAAAACAAAACA 299
72E18_mandshurica       AATCAA--------ATGCTACATTCACACAACAAAGAAATAGACATTCAAAGACAAAACA 292
72E18_nilgerrensis      NNNNNN--------NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 292
72E18_viridis           AATCAATGACAACAATGCCTCATTCACACAACAAAGAAATAGACATTCAAAAACAAAACA 300
72E18_iinumae           AATCAATGACACCAATGCCTCATCGACACAACAAACAAATAGACATTCAAAAACAAAACA 300
72E18_ananassa          AATCAATGACAACAATGCCTCA-----------AACAAATAGACATTCAAAAACAAAACA 289


72E18_vesca             CAATACACACTACTAATGTG---------------------------------------- 319
72E18_mandshurica       CAA-ACACACTACTAACGTG---------------------------------------- 311
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNN---------------------------------------- 312
72E18_viridis           CAATACACACTACTAATGTG---------------------------------------- 320
72E18_iinumae           CAATACACGATGCTAACATTTCCCTAATCTCTCCTCCATCAACTAAAATCTCCATTCCAA 360
72E18_ananassa          CAATACACAATGCTAACATTTCCCTAATATCTCCTCCATCAACTAAAATCTCCATGCCAG 349


72E18_vesca             -------------GCACAGAAACCAAAGCATGATTCAAAACAAAACTAGAACATCTACAT 366
72E18_mandshurica       -------------GCACGGAAACCAAAGCATGATTCAAAACAAAACTAGAACATCTACAT 358
72E18_nilgerrensis      -------------NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 359
72E18_viridis           -------------GCACAGAAACCAAAGCATGATTCAAAACAAAACTAGAACATCTACAT 367
72E18_iinumae           ATCACACACTACTACACAGAAACCAAAGCATGATTCAAAACAAACCAAGAACATCTACAT 420
72E18_ananassa          ATCAC---------CACAGAAACCAAAGCATGATTCAAAACAAACCAAGAACATCTACAT 400
```

```
72E18_vesca             AGTTCTCTCACAATAGTAAAGAAACGATCTTTGACAATCAAAAGGCATCGAAAGCTAGTA 426
72E18_mandshurica       AGTTCTCTCACAATAGTAAAGAAACGATCGTTGACAATCAAAAGGCATCGAAAGCTAGTA 418
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 419
72E18_viridis           AGTTCTCTCACAATAGTAAAGAAACGATCTTTGACAATCAAAAGGCATCGAAAGCTAGTA 427
72E18_iinumae           AGTCCTCTCACAATAGTAAAGAAACGATCTTTGACAATCAAAAGGCATCGAAAGCTAGTA 480
72E18_ananassa          AGTTCTCTCACAATAGTAAAGAAACGATCTTTGACAATCAAAAGGCATCGAAAGCTAGTA 460


72E18_vesca             AAGAAACGATCTTTCAGATGGGAAATACCCAAATTTGATTGCTACATGCATAAAACCCTC 486
72E18_mandshurica       AAGAAACGATCTTTCAGATGGGAAATACCCAAATTTGATTGCTATATACATAAAACCCTC 478
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 479
72E18_viridis           AAGAAACGATCTTTCAGATGGGAAATACCCAAATTTGATTGCTACATGCATAAAACCCTC 487
72E18_iinumae           AAGAAACGATCTTTCAGATGGGAAATACCCAAATTTGATTGCTATATACATAAAACCCTC 540
72E18_ananassa          AAGAAACGATCTTTCAGATGGGAAATGCCCAAATTTGATTACTATATACATAAAACTCCC 520


72E18_vesca             AAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATAAAAAAAA-TTCAA 545
72E18_mandshurica       AAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATAAAAAAAAATTCAA 538
72E18_nilgerrensis      NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 539
72E18_viridis           AAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATAAAAAAAAAATTCA 547
72E18_iinumae           AAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACATTAAAAAAAAATCAA 600
72E18_ananassa          AAATTGATACGAAATCAAACAATGCAGCAATCAAATCATTCCACAG-AAAAAAAATTCAA 579


72E18_vesca             GAAAA-AAAGAGAGAGA--AAATTACAGATTTAAAGCGACGAACAA-TGAAAAGGAATGA 601
72E18_mandshurica       GAAAA-AAAGAGAGAGA--AAATTACAGATCTAAAGCGACGAACAG-TGAGAAGGAATGA 594
72E18_nilgerrensis      NNAAA-AAAGAGAGAGAGA---TTACAGATCTAN-GCGACGAACAA-TGAGAAGGAATGA 593
72E18_viridis           AGAAATAAAGAGAGAGA--AAATTACAGATCTAAAGTGACGAACAA-TGAGAAGGAATGA 604
72E18_iinumae           GAAAAAAAGAAGAGAGA--AAATTACAGATCTAAAGCGACGAACAAATGAGAAGGAATGA 658
72E18_ananassa          GAAAAAAAAAAGAGAGAGAAAATTACAGATCTAAAGCGACGAACAA-TGAGAAGGAATGA 638
                          *** **  ** ****   * ******** **  * ********  *** *********

72E18_vesca             GAGGCAAAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGGGAGAGA 661
72E18_mandshurica       GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGGGAGAGA 654
72E18_nilgerrensis      GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGT--------GA 645
72E18_viridis           GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGTGAGTGAGG--GAGA 662
72E18_iinumae           GAGACAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGAATGAGAGT--------GAGAGA 710
72E18_ananassa          GAGGCAGAGAGAAGAGATGAGGAAGTTGACCTTTGTGA----------GTGAGGGAGAGA 688
                        *** ** ******************************* ****   *** ***

72E18_vesca             GAGAGAGAGATCGACGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 721
72E18_mandshurica       GAGAGAGAGATCGACGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 714
72E18_nilgerrensis      GAGAGAGAGATCGAAGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 706
72E18_viridis           GAGAGAGAGATCGAAGACGAAGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 722
72E18_iinumae           GAGAGAGAGATCGAAGACGAGGCAGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 771
72E18_ananassa          GAGAGAGAGATCGAAGACGAAGCTGAGCGAAAGAGACGAGTGTGGTGTTTGTGAGTTGAG 749
                         ************* ***** ** ************************************

72E18_vesca             GCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATT 781
72E18_mandshurica       GCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATT 774
72E18_nilgerrensis      GCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATT 765
72E18_viridis           GCGAA-GAATTGNACCNNNATANAGGAGTGNGATTGACNAGTTATCTCNGCNGNTTGATT 781
72E18_iinumae           GCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATT 830
72E18_ananassa          GCGAAAGAATTGGAGCAAAATAAAGGAGTGGGATTGACGAGTAATCTCAGCCGTTTGATT 808
                        ***** ****** * *   *** ******* ******* *** ***** ** * ******

72E18_vesca             TATGGACCGCGTCTATTGAGCCCTTGTGGGG-CCATTACAGCTCCTTCCGCTGTTCCAGT 840
72E18_mandshurica       TATGGACCGCGTCTATTGCGCCCTTGTGGGG-CCATTACAGCTCCTTCCGCTGTTCCAGT 833
72E18_nilgerrensis      TATGGACCGCGTCCATTGCGCCCTTGTGGGG-CCATTACAGCTCCTTCCGCTGTTCCAGT 824
72E18_viridis           TATGGACCGCGTCCATTGTGCCNTTGTGGGG-CCATNACNGCTCCTNCCNCTGTNCCNGC 840
72E18_iinumae           TATGGACCGCGTCCATTGCGCCCTTGTGGGGGCCATTACAGCTCCTTCCGCTGTTCCAGT 890
72E18_ananassa          TATGGACCGCGTCCGTTGCGCCCTTGTGGGG-CCATTGCAGCTCCTTCCGCTGTTCCAGT 867
                        *************  *** *** ******** ****  * ****** ** **** ** *
```

```
72E18_vesca              CATTTTTTTCTCCACCTTCTTTACCTTTTTGCCCCTCAGTCCCTT-CCCTTTTCTCCCAA 899
72E18_mandshurica        CATTTTTTTCTCCACCTTCTTCACCTTTTTGCCCCTCAGTCCCTT-CCCTTTTCTCCCAA 892
72E18_nilgerrensis       CATTTTTTTCTCCACCTTCTG------------CCCCATTCCCTT-CCCTTTTCTCCCAA 871
72E18_viridis            CATNTTTATCTCCACCTTCTNCACNTTTNTGCCNCNCNTGNCCNT-CCCTTTTCTCCCNN 899
72E18_iinumae            CATTTTT--CTCCACCTTCTTCACCTTTCTGCCCCTCGTTCCCTTTCCCTCTTCTCCCAA 948
72E18_ananassa           CATTTTT--CTCCACCTTC----------------------------------------- 884
                         *** ***  **********


72E18_vesca              TTCTT---------------------------------------------TCTCAACTCT 914
72E18_mandshurica        TTCTT---------------------------------------------TCTCAACTCT 907
72E18_nilgerrensis       TTCTT---------------------------------------------TCTCAACTCT 886
72E18_viridis            TTCTN---------------------------------------------NCTCANATCT 914
72E18_iinumae            CTCTTCTTAGCCTAATTGCATTTTCATTTTAGTGCTTAGATCAATATGATTNNNNNNNNN 1008
72E18_ananassa           ------------------------------------------------------------


72E18_vesca              TCTTAAACCTAATTGCATTTCCCTAATTGCATTTTCATTTTAGTGCTTAGATCAATATGA 974
72E18_mandshurica        TCTTAAACCTAATTGCATTTCCCTAATTGCATTTTCATTTTAGTGCTGAGATCAATATGA 967
72E18_nilgerrensis       TCTTAAACCTAAT--------------TGCATTTTCATTTTANTGCTTANATCAATATGA 932
72E18_viridis            TCTNAANGCTNATNGCNTTTNCCNNANNGCATTTGNATTTNNGNGCTTCNATCAATNCGN 974
72E18_iinumae            NNNNNNNNNNNNNN------------NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1056
72E18_ananassa           ACCTTTTTGCCCCT------------CATTCCCTTTCCCTCTTCTCCGATCCCAACTTTC 932

72E18_vesca              TTAAGAAGCTTCATTTTGTCAACACAAGGCAACAAGGACA--CAGGGGAGCATTTCGATC 1032
72E18_mandshurica        TTAAGAAGCTTCATTTTGTCAACACAAGGCAACAAGGACA--CAAGGGAGCATGTCGATC 1025
72E18_nilgerrensis       TTAAGAAGCTCCATTTTGTCAACACAAGGCAACAAGGACN--TANGGGAGCATGTCGATC 990
72E18_viridis            NNNNAANGCTNCGTNTTGTCNTNNCANGGNTNCAAGGANCCTNNGGGANGCNNGTTGATC 1034
72E18_iinumae            NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1116
72E18_ananassa           TCAACTCTTCTTAACCCACCCAGTTGCATTTTCATTGTAGTGCTTAAATCAATATGGTTA 992

72E18_vesca              ATCGTTCCAGTCATTTTCGTATATAAT-TTGGGCTTGAAATGGTTGATC-GGTCGTAAAA 1090
72E18_mandshurica        ATCGTTCCAGTCATTTTCGTATATAAT-TTGGGCTTGAAATGGTTAATC-AATCGTAAAA 1083
72E18_nilgerrensis       ATCGTTCCGGTCACTTTCGTATATAAT-TTGGACTTAAAATGGTTGATC-GATCGTAAAA 1048
72E18_viridis            GTCAANTNCGGCCTGGTNATANANAATATNNGACNTNAAATGGTTGATTGNNTCGTTAAA 1094
72E18_iinumae            NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1176
72E18_ananassa           AGAAGCTTCATTTTGTCAACACAAGGCAACAAGGACATAGGGGAGCATGTGGATGATCGT 1052

72E18_vesca              TTTAAAATGACGTTTTGATATGATATCTATAAAGA-GGACATAATTTACTTTGTATGTCA 1149
72E18_mandshurica        TTTAAAATGACGTTTTGATATGATATCTATAAGGA-GGACATAATTTACTTTATATGTCA 1142
72E18_nilgerrensis       TTTGAAATGACGTTTTGGTATGATATTTGTAAAGA-GGACATAATTTACT---------- 1097
72E18_viridis            TTNGAAATAANNNNNNNNNNNNNNNNNNNNNNAAAAGGNCNTAATTTANTNNGTANNTCA 1154
72E18_iinumae            NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN----- 1231
72E18_ananassa           TTTTAAACTACGTTTTGGTATGATATCGTAAGGAGGACATAATTTACTTTGTATG----- 1107

72E18_vesca              GGTTTTAATATGGAA----CGAAGAGTATGGGTGAAA-GTGTC-ATCCCACACATTTTAA 1203
72E18_mandshurica        GGTTTTAATACGGAA----GGAAGAGTATGGGTGAAA-GTGTC-ATCCCACACATTTTAA 1196
72E18_nilgerrensis       ---------ATGGAA----CGAAGAGTATAGGTGAAA-GTGTC-ATCCCACACATTTTAA 1142
72E18_viridis            AGGTTTTAANTAGGGAAATGGAGGAGTGNGGGTGAAAAGTGTCCATCCCACACGTTTTAA 1214
72E18_iinumae            ---------NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1282
72E18_ananassa           ---------TCNGGTTTTTAATACGGAATGAAGAGTGTGGGAGAAAGTGTCATCCCACAC 1158

72E18_vesca              AAGAGCTGTAATGTAGGG-TAATGAGCA-CAACTGCAAGCTGCATCC-TATAA-TGGATC 1259
72E18_mandshurica        AAGAGCTGTAATGTAGGG-TAATGAGCA-CAACTGCAAGCTGCATCC-TATAA-TGGATC 1252
72E18_nilgerrensis       AAGAGCTTTAATGTAGGGGTAATGAGCA-CAACTACAAGCTGCATCC-TATAA-GGGATC 1199
72E18_viridis            ANGAGCNTTAANGNAGGG-NAANGNGCAGNAACNGCAAGCTGCATNCCTNTAATNGGATC 1273
72E18_iinumae            NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1342
72E18_ananassa           ATTCATTTTAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1218
```

```
72E18_vesca             AATCAGAACATTAA-ACAACGTAAAGAGGAAGGTATTTGCTTTACACAACCTTATAAAAT 1318
72E18_mandshurica       AATCAGAACAATAA-ACAACGTAAAGAGGAAGGTATTTGCTTTACACAACCTTATAAAAT 1311
72E18_nilgerrensis      AATCAGAACATTAA-ACAACGTAAAGAGGAAGGTATTTGCTTTACACAACCTTATAAAAT 1258
72E18_viridis           AATCAGAACATTAAGNCAACGTAAAGAGGAAGGTATTTGCTTTACACAACCTTATAAAAT 1333
72E18_iinumae           NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1402
72E18_ananassa          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1278


72E18_vesca             GATGAGGATCTACTC-AAAATCCAGACTACCATGGTTGGCAAAATTAGATCCTCACTGTA 1377
72E18_mandshurica       GATGAGGATCTACTC-AAAATCCAGACTACCATGGTTGGCAAAATTAGATCCTCACTGTA 1370
72E18_nilgerrensis      GATGAGGATCTACTC-AAAATCCAGACTACCATGGTTGGCAAAATTAGATCCTGCCTGTA 1317
72E18_viridis           GATGNGGATNTACTCCAAAGTGAGGACTACCATGGTCGGCAAAATTAGTTCCTGACTGTA 1393
72E18_iinumae           NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1462
72E18_ananassa          NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCAAAATCCAGACTACC 1338


72E18_vesca             ACCAGCTAGGCATTGGTAATGCATAATAGCTATAGCTAACTATAGGTGGGAGACTCATCA 1437
72E18_mandshurica       ACCAGCTAGGCATTGGTAATGCATAATAGCTATAGCTAACTATAGGTGGGAGACTCATCA 1430
72E18_nilgerrensis      ACCAGCTAGGCATTGGCAATGCATAATAGCTAGAGCTAACCATAGGTGGGAGACTCATCA 1377
72E18_viridis           ACCAGCTAGGCATNGGCAATGCATAATAGCTATAGCTGACTATAGGTGGGAGACTCATCA 1453
72E18_iinumae           NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1522
72E18_ananassa          GTGGTGGGCAAAATTAGATCCTCACCTATAGTAGGGTAACTATAGGTGGGAGACTNGTCA 1398


72E18_vesca             TTGAGATCATAGAAAAACAAAGATGAAAGAAAGAAATGAAGAAACAAGCAACAGCTATTC 1497
72E18_mandshurica       TTGAGATCATAGAAAAACAGAGATGAAAGAAAGAAATGATGAAACAAGCAACAGCTATTC 1490
72E18_nilgerrensis      TTGAGATCATAGGAAAA------------AAAATGATGA--AAACAAGCAACAGTTATTC 1423
72E18_viridis           TTGAGATCAGAGAAAAACNAAGATGAAAGAAAGAAATGATGAAACAAG-AATAGTTATTC 1512
72E18_iinumae           NNNNNNNNNNNNNNNNN---------------NNNNNNNNNNNNNNNNNNNNNNNNNNNN 1567
72E18_ananassa          TTGTGATCATAGAAAAA---------------GAAATGATGAAACAAGCAACAGTTACTC 1443


72E18_vesca             GAAAGCAAGTACAGAAGGGATTGTTCATGAAGTGTTCACCAAGTCACAGCTTAGGGCATT 1557
72E18_mandshurica       GAAAGCAAGTACAGAAGGGATTGTTCATGAAGTGTTCACCAAGTCACAGCTTAGGGCATT 1550
72E18_nilgerrensis      GAAAGCAAGTACAGAAGGGATTGTTCATTAAGTGTTCACCAAGTCACAGCTTAGGGCATT 1483
72E18_viridis           GAAAGCAAGTACAGAAGGGATTGTTCATGAAGTGTTCACCAAGTCACAGCTTAGGGCATT 1572
72E18_iinumae           NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1627
72E18_ananassa          GAAAGCAGGTACAGAAGGGATTGTTCATGAAGTGTTCACCAAGTCNCAGCTTAGGGCATT 1503


72E18_vesca             CTTAGAAGTAACAAGCTTACCAACTTCCATTTACTTGTTTCAAGTTCATGATGATATTAA 1617
72E18_mandshurica       CTTAGAAGCAACAAGCTTACCAACTTCCATTTACTTGTTTCAAGTTCATGATGATATTAA 1610
72E18_nilgerrensis      CTTAGAAGCAACAAGCTTACCAACTTCCATTTACTTGTTTCAAGTTCATGATGATATTTA 1543
72E18_viridis           CTTAGAAGCAACAAGCTTGCCAACTTCCATTTACTTGTTTCAAGTTCATGNTGATATTAN 1632
72E18_iinumae           NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1687
72E18_ananassa          CTTAGAAGCAACAAGCTTACCAACTTCCATTTACTTGTTTCAAGTTCAGGGTGACATTAC 1563


72E18_vesca             CCATCCAACGAAAACATCCAAAGGTACTGTGACAGAAAGCTCAAGGGGATATCTGTGTTT 1677
72E18_mandshurica       CCATCCAACGAAAACATCCAAAGGTACTGTGACAGAAAGCTCAAGGGGATATCTGTGTTT 1670
72E18_nilgerrensis      -----------------------GTCATGT-------------------------TGTTT 1555
72E18_viridis           CCATCCAACGAAAACATCCAAAGGTNCTGTGACTGAAAGCTCAAGGGGATATCNGTGTTT 1692
72E18_iinumae           NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCGAAAGCTCAATGGGATATCTGTGTTT 1747
72E18_ananassa          CCNTCCAACGAAAACATCCAAAGGTACTGTGACTGAAAGCCCAAGGGGATATCCGTGTTT 1623
                                                                               *****

72E18_vesca             AAAGCCACAACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTG 1737
72E18_mandshurica       AAAGCCACAACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTG 1730
72E18_nilgerrensis      AAAGCCACAACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTG 1615
72E18_viridis           AAAGCCACTACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTG 1752
72E18_iinumae           AAAGCCACAACATGACTAAATATAATTGCTCCCAATTTCTAAAGTTACATTCGTTTTGTA 1807
72E18_ananassa          AAAGCCACAACATGACTAAATATAATTGCTTCCAATTTCTAAAGTTACATTCGTTTTGTG 1683
                        ******** ********************* **************************** 
```

```
72E18_vesca             CAAATGACAAAACAGTTCAAATTGACTGCATAA-ATAGATTACTCTTGTATAGATCAACA 1796
72E18_mandshurica       CAAATGACAAAACAGTTCAAATTGACTGCATAA-ATAGAT-ACTCTTGTATAGATCAACA 1788
72E18_nilgerrensis      CAGATGACAAAACAGTTCAAATTGACTGCATAAGATAGATTACTCTTGTATAGATCAACA 1675
72E18_viridis           CAAATGACAAAACAGTTCTAATTGACTGCATAAGATAGATTACTCTTGTATAGATCAACA 1812
72E18_iinumae           CAAATGACAAAACAGTTCAAATTGACTGCATAA-ATAGATTACTCTTGTATAGATCAACA 1866
72E18_ananassa          CAAATGACAAAACAGTTCAAATTGACTGCATAAGATAGATTACTCTTGTATAGATCAACA 1743
                        ** *************** ************** ****** *******************


72E18_vesca             AGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAATACTAACATCAATATTGAAAT 1856
72E18_mandshurica       AGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAATACTAACATCAATATTGAAAT 1848
72E18_nilgerrensis      AGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAACACTAACATAAATATTGAAAT 1735
72E18_viridis           AGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAAAACTAACATCAATATTGAAAT 1872
72E18_iinumae           AGCAAATCTCCAAGTTCTTATTACAAAGTCTAAGCAGAATACTAACATCAATATTGAAAC 1926
72E18_ananassa          AGCAAATCTCCAAATTCTTATTACAAAGTCTAAGCAGAATACTAACATCAATATTGAAAT 1803
                        ************* ************************* ******** ********** 


72E18_vesca             TGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCGTAGGAAATGGAATTG 1916
72E18_mandshurica       TGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCGTAGGAAATGGAATTG 1908
72E18_nilgerrensis      TGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCATAGGAAATGGAATTG 1795
72E18_viridis           TGGATAAATATGCGATCTGAACTTCTTCACGTTGATGAGCTATCGTAGGAAATGGAATTG 1932
72E18_iinumae           TGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCGTAGGAAATGGAATTG 1986
72E18_ananassa          TGGATAAATATGCGATCTGAACTTCTTCACGTTGATGACCTATCGTAGGAAATGGAATTG 1863
                        ************************************** ***** ***************

72E18_vesca             AACACTTGACACCAAAGAGAACAATGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGG 1976
72E18_mandshurica       AACACTTGACACCAAAGAGAACAATGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGG 1968
72E18_nilgerrensis      AACACTTGACACCAAAGAGAACAACGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGG 1855
72E18_viridis           AACACTTGACACCAAAGAGAACAACGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGG 1992
72E18_iinumae           AACACTTGACACCAAAGAGAACAATGAAGGTAGCCTCGCCAATCACTTCTACAAGAATGG 2046
72E18_ananassa          AACACTCGACACCAAAGAGAACAACAAAGGTAGCCTCACCAATCACTTCTACAAGAATGG 1923
                        ****** *****************  *********** **********************

72E18_vesca             GGGTAGAATCACCCATCGACGTGGATACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTG 2036
72E18_mandshurica       GGGTAGAATCACCCATCGACGTGGATACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTG 2028
72E18_nilgerrensis      GGGTAGAATCACCCATCTACGTGGATACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTG 1915
72E18_viridis           AGGTAGAATCACCCATCGACGTGGATACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTG 2052
72E18_iinumae           GGGTAGAATCACCCATCGACGTGGATACTTGGGTCTTCCGTCCTTCCCATCAAATAGCTG 2106
72E18_ananassa          GGGTAGAATCACCCATCGACGTGGATACTTGGGTCGTCCGTCCTTCCCATCAAATAGCTG 1983
                         **************** ***************** ************************

72E18_vesca             GACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTACATACAAACT 2096
72E18_mandshurica       GACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTACATACAAACT 2088
72E18_nilgerrensis      GACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTACATACAAACT 1975
72E18_viridis           GACATGGCAGGGTGTCACAAAAGATCAATATTGCATGCAAAGAGCTTCTACATACAAACT 2112
72E18_iinumae           GACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTACATACAAACT 2166
72E18_ananassa          GACATGGCAGGGTGTCACAAAAGATCAATATTGCATGTAAAGAGCTTCTACATACAAACT 2043
                        ************************************* **********************

72E18_vesca             CAAATGGATATGTTCTGGCGCTTGTAGAATATAATTATGTATACAAATATGCATGTACAG 2156
72E18_mandshurica       CAAATGGATATGTTCTGGCGCTTGTAGAATATAATTATGTATACAAATATGCATGTACAG 2148
72E18_nilgerrensis      CATATGGATATGTTCTGGCGATTGCAGAATATAATTATGTATACAAATATGCATGTACAG 2035
72E18_viridis           CAAATGGAGATGTTCTGGCGCTTGTAGAATATAATTATGTATACAAATATGCATGTACAG 2172
72E18_iinumae           CAAATGGATATGTTCTGGCGCTTGTAGAATATAATTATGTATACAAATATGCATGTACAG 2226
72E18_ananassa          CAAATGGATATGTTCTGGCGCTCGTAGAATATAATTATGTATACAAATATGCATGTACAG 2103
                        ** ***** *********** * * ***********************************

72E18_vesca             AGCTTCCACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTCCAATAAA 2216
72E18_mandshurica       AGCTTCCACATACAAACTCATATGTATACTTGTAAATTTATGCAATTTAATTCCAATAAA 2208
72E18_nilgerrensis      AGCTTCTACATACAAACTCATACGAATACTTGTAAATTTAGGCAATTTAATTCCAATAAA 2095
72E18_viridis           AGCTTCCACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTCCAATAAA 2232
72E18_iinumae           AGCTTCCACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTCCAATAAA 2286
72E18_ananassa          AGCTTCCACATACAAACTCATATGAATACTTGTAAATTTATGCAATTTAATTCCAATAAA 2163
                        ****** *************** * *************** *******************
```

```
72E18_vesca             GGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAAAA---CAGACAAAACATTTAAGCA 2273
72E18_mandshurica       GGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAAAAA--CAGACAAAACATTTAAGCA 2266
72E18_nilgerrensis      GGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAAAA-----GACAAAACATTTAAGCA 2150
72E18_viridis           GGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAAAAAAAAAGACAAAACATTTAAGCA 2292
72E18_iinumae           GGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAAAAAC---AGACAAAACATTTAAGCA 2343
72E18_ananassa          GGTGAGTTTAAATAGACCAAGATGTTAGCTAAAAA--AC---AGACAAAACATTTAAGCA 2218
                        ***********************************  *     *****************

72E18_vesca             AAAGAAGAGCAGTAGAAGGTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAAAG 2333
72E18_mandshurica       AAAGAAGAGCAGTAGAAGGTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAAAG 2326
72E18_nilgerrensis      AAAGAAGAGCAGTAGAAGTTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAATG 2210
72E18_viridis           AAAGAAGAGCAGTAGAAGTTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAAAG 2352
72E18_iinumae           AAAGAAGAGCAGTAGAAGGTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAAAG 2403
72E18_ananassa          AAAGAAGAGCAGTAGAAGTTATTAAGATACCAAACAACATATTTGGGTTGGAGGACAAAG 2278
                        ***************** **************************************** *

72E18_vesca             TAGTATAGAGGAGTGTACCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATTGGTCAT 2393
72E18_mandshurica       TAGTATAGAGGAGTGTTCCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATTGGTCAT 2386
72E18_nilgerrensis      TAGCATAGAGGAGTGTACCTTCTTTAAACGGCGGTGCTTTCCTAGGGCCCAGTTGGTCAT 2270
72E18_viridis           TCGTATAGAGGAGTGTACCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATTGGTCAT 2412
72E18_iinumae           TAGTATAGAGGAGTGTACCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATTGGTCAT 2463
72E18_ananassa          TAGTATAGAGGAGTGTACCTTCTTTAAACGGCGGTGTTTTCCTAGGGCCCAATTGGTCAT 2338
                        * * ************ ******************* ************** ********

72E18_vesca             GATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAA 2453
72E18_mandshurica       GATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAA 2446
72E18_nilgerrensis      TATAGAAGCAGCAACTGCAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAA 2330
72E18_viridis           GATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACTGTCTGTGTTGCAATATTAAA 2472
72E18_iinumae           GATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAA 2523
72E18_ananassa          GATAGAAGCAGCAACTACAGCAAAAAGATAACCAGCAACCGTCTGTGTTGCAATATTAAA 2398
                         *************** ********************** ********************

72E18_vesca             ACCCAACCACTGATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAA 2513
72E18_mandshurica       ACCCAACCACTGATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAA 2506
72E18_nilgerrensis      ACCCAACCACTGATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAA 2390
72E18_viridis           ACCCAACCACTGATAAATCTCAGTTGTGTAATTTGCACATGTCACAATATTGAATAGAAA 2532
72E18_iinumae           ACCCAACCACTGATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAA 2583
72E18_ananassa          ACCCAACCACTGATAAATCTCAGTCGTGTAATTTGCACATGTCACAATATTGAATAGAAA 2458
                        ************************ ***********************************

72E18_vesca             ACCACGTGGTATTTGATAGCCTCCACTTCCATCAGGACTTCGCAGATTCCTCAGTAGAAT 2573
72E18_mandshurica       ACCACGTGGTATTTGATAGCCTCCACTTCCATCAAGACTTCTCAGATTCCTCAGTAGAAT 2566
72E18_nilgerrensis      ACCACGAGGTATTTGATAGCCTCCACTTCCATCAGGACTTCGCAGATTCCTCAGTAGAAT 2450
72E18_viridis           ACCACGTGGTATTTGATAGCCTCCACTTCCATCAGGACTTCGCAGATTCCTCAGTAGAAT 2592
72E18_iinumae           ACCACGTGGTATTTGATAGCCTCCACTTCCATCAGGACTTCGCAGATTCCTCAGTAGAAT 2643
72E18_ananassa          ACCACGTGGTATTTGATAGCCTCCGCTTCCATCAGGACTTCGCAGATTCCTCAGTAGAAT 2518
                        ****** ***************** ********* ****** ******************

72E18_vesca             ATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAACCCA 2622
72E18_mandshurica       ATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAACCCA 2615
72E18_nilgerrensis      ATGGCAATACAAGTTCGCGATTTGATTTATTATCCCAAAACCAAACCCA 2499
72E18_viridis           ATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAACCCA 2641
72E18_iinumae           ATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAACCCA 2692
72E18_ananassa          ATGGCAATAGAAGTTCGCAATTTGATTTATTATCCCAAAACCAAACCCA 2567
                        ********* ******** ******************************
```

LIST OF REFERENCES

Abdulova G, Ananiev E, Grozdanov P (2002) Isolation and purification of nuclear DNA from excised cotyledons of *Cucurbita pepo* L.(zucchini). Bulg. J. Plant Physiol. **28:** 3-11

Ahmadi H, Bringhurst RS, Voth V (1990) Modes of inheritance of photoperiodism in *Fragaria*. J. Amer. Soc. Hort. Sci. **115:** 146–152

Akiyama Y, Yamamoto Y, Ohmido N, Oshima M, Fukui K (2001) Estimation of the nuclear DNA content of strawberries (*Fragaria* spp.) compared with *Arabidopsis thaliana* by using dual-stem flow cytometry. Cytologia **66:** 431-436

Albani M, Battey NH, Wilkinson MJ (2004) The development of ISSR-derived SCAR markers around the Seasonal Flowering Locus (SFL) in *Fragaria vesca*. Theoretical & Applied Genetics **109:** 571-579

Aljanabi SM, Forget L, Dookun A (1999) An improved rapid protocol for the isolation of polysaccharide and polyphenol-free sugarcane DNA. Plant Mol. Biol. Rep. **17:** 1-8

Anonymous (1980) IEEE Transactions on Magnetics **16:** 387-490

Anonymous (1998) Montreal Protocol on Substances that Deplete the Ozone Layer. United Nations Environmental Program (UNEP).

Anonymous (2001) Journal of Magnetism and Magnetic Materials **225:** 1-314

Antonius K, Ahokas H (1996) Flow cytometric determination of polyploidy level in spontaneous clones of strawberries. Hereditas **124:** 285

Arnau G, Lallemand J, Bourgoin M (2003) Fast and reliable strawberry cultivar identificaion using inter simple sequence repeat (ISSR) amplification. Euphytica **129:** 69-79

Arulsekar S, Bringhurst RS, Voth V (1981) Inheritance of PGI phosphoglucoisomerase and LAP leucine aminopeptidase isozymes in octoploid cultivated strawberry. J Amer Soc Hort Sci **106:** 679-683

Ashley MV, Wilk JA, Styan SMN, Craft KJ, Jones KL, Feldheim KA, Lewers KS, Ashman TL (2003) High variability and disomic segregation of microsatellites in the octoploid *Fragaria virginiana* Mill. (Rosaceae). Theor Appl Genet. **107**

Barakat A, Matassi G, Bernardi G (1998) Distribution of genes in the genome of *Arabidopsis thaliana* and its implications for the genome organization of plants. Proc Natl Acad Sci U S A **95:** 10044–10049

Bedbrook J, Gerlach W, Thompson R, Flavell RB (1980) Emergent Techniques. University of Minnesota Press, Minneapolis

Bell JA, Simpson DW (1994) The use of isozyme polymorphisms as an aid for cultivar identification in strawberry. Euphytica **77:** 113-117

Benitez-Burraco A, Blanco-Portales R, Redondo-Nevado J, Bellido ML, Moyano E, Caballero JL, Munoz-Blanco J (2003) Cloning and characterization of two ripening-related strawberry (*Fragaria* × *ananassa* cv. Chandler) pectate lyase genes. J Exp Bot J Exp Bot **54:** 633-645

Bennett MD, Leitch IJ (2005) Nuclear DNA amounts in angiosperms: progress, problems and prospects. Annals of Botany **95:** 45-90

Bennett MD, Leitch IJ, Price HJ, Johnston S (2003) Comparisons with *Caenorhabditis* (~100 Mb) and *Drosophila* (~175 Mb) using flow cytometry show genome size in *Arabidopsis* to be ~ 157 Mb and thus ~25% larger than the *Arabidopsis* Genome Initiative estimate of ~125 Mb. Annals of Botany **91:** 547-557

Bennetzen JL, Kellogg EA (1997) Do plants have a one-way ticket to genomic obesity? Plant Cell **9:** 1509-1514

Bennetzen JL, Ma J, Devos KM (2005) Mechanisms of Recent Genome Size Variation in Flowering Plants. Annals of Botany **95:** 127-132

Bennetzen JL, SanMiguel P, Chen M, Tikhonov A, Francki M, Avramova Z (1998) Grass genomes. Proc Natl Acad Sci U S A **95:** 1975-1978

Besemer J, Borodovsky M (1999) Heuristic approach to deriving models for gene finding. Nucleic Acids Res. **27:** 3911-3920

Bies DH, Folta KM (2004) An effective substitute for triisopropylnaphthalenesulfonic acid in the preparation of plant RNA. Anal Biochem. **333:** 201-203

Birney E, Clamp M, Durbin R (2004) GeneWise and Genomewise. Genome Research **14:** 988-995

Bringhurst RS (1990) Cytogenetics and Evolution in American *Fragaria*. Hortscience **25:** 879-881

Bringhurst RS, Arulsekar S, Hancock JF, Voth V (1981) Electrophoretic characterization of strawberry cultivars. Journal of the American Society for Horticultural Science **106:** 684-687

Bringhurst RS, Gill T (1970) Origin of *Fragaria* polyploids. II. Unreduced and doubled-unreduced gametes. American Journal of Botany **57:** 969-976

Bullock WO, Fernandez JM, Short JM (1987) XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. Biotechniques **5:** 376-379

Burge C, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. J. Mol Biol **268:** 78-94

Chaudhry B, Yasmeen A, Husnain T, Riazuddin S (1999) Mini-scale genomic DNA extraction from cotton. Plant Molecular Biology Reporter **17:** 1-7

Chilton M-D, Drummond MH, Merlo DJ, Sciaky D, Montoya AL, Gordon MP, Nester EW (1977) Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. Cell **11:** 263-271

Chomczynski P, Mackey K, Drews R, Wilfinger W (1997) DNAzol: A reagent for the rapid isolation of genomic DNA. BioTechniques **22:** 550-553

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Analytical Biochemistry **162:** 156-159

Cohn EJ, Conant JB (1926) The molecular weight of proteins in phenol. PNAS **12:** 433- 438

Collins GG, Symons RH (1992) Extraction of nuclear DNA from grape vine leaves by a modified procedure. Plant Mol. Biol. Rep. **10:** 233-235

Cox RA (1968) The use of guanidine hydrochloride in the isolation of nucleic acid. *In* L Grossman, K Moldave, eds, Methods in Enzymology, Vol 12B. Academic Press, New York, pp 120-129

Craigie JS, McLachlan J (1964) Excretion of coloured ultraviolet absorbing substances by marine algae. Can J Bot **42:** 23-33

Crowley TM, Muralitharan MS, Stevenson TW (2003) Isolating conifer DNA: a superior polysaccharide elimination method. Plant Mol. Biol. Rep. **21:** 97a-97d

Dabo SM, Mitchell EDJ, Melcher U (1993) A method for the isolation of nuclear DNA from cotton (*Gossypium*) leaves. Anal Biochem. **210:** 34-38

Dale A, Sjulin TM (1990) Few cytoplasms contribute to North-American strawberry cultivars. Hortscience **25:** 1341–1342

Darrow GM (1966) The Strawberry. New York, Holt, Rinehart and Winston

Davis TM, DiMeglio LM (2004) Identification of putative diploid genome donors to the octoploid cultivated strawberry, *Fragaria* × *ananassa*. PAG-XII, San Diego, CA

Davis TM, Yu H (1997) A linkage map of the diploid strawberry, *Fragaria vesca*. Journal of Heredity **88:** 215-221

Davis TM, Yu H, Haigis KM, McGowan PJ (1995) Template mixing: a method of enhancing detection and interpretation of codominant RAPD markers. Theoretical and Applied Genetics **91:** 582 - 588

de la Cruz M, Ramirez F, Hernandez H (1997) DNA isolation and amplification from cacti. Plant Molecular Biology Reporter **15:** 319-325

Degani C, Rowland LJ, Levi A, A. HJ, Galletta GJ (1998) DNA fingerprinting of strawberry (*Fragaria* × *ananassa*) cultivars using randomly amplified polymorphic DNA (RAPD) markers. Euphytica **102:** 247-253

Degani C, Rowland LJ, Saunders JA, Hokanson SC, Ogden EL, Golan-Goldhirsh A, Galletta GJ (2001) A comparison of genetic relationship measures in strawberry (*Fragaria* × *ananassa* Duch.) based on AFLPs, RAPDs, and pedigree data. Euphytica **117:** 1-12

Deng C, Davis TM (2001) Molecular identification of the yellow fruit color (*c*) locus in diploid strawberry: a candidate gene approach. Theor Appl Genet. **103:** 316-322

Dirlewanger E, Cosson P, Tavaud M, Aranzana M, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theor Appl Genet. **105:** 127-138

Diwan N, Bouton JH, Kochert G, Cregan PB (2000) Mapping of simple sequence repeat (SSR) DNA markers in diploid and tetraploid alfalfa. Theor Appl Genet. **101:** 165-172

Doležel J, Bartoš J, Voglmayr H, Greilhuber J (2003) Nuclear DNA content and genome size of trout and human. Cytometry **51A:** 127-128

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin **19:** 11-15

Durbin M, Learn Jr G, Huttley G, Clegg M (1995) Evolution of the Chalcone Synthase Gene Family in the Genus *Ipomoea*. Proc Natl Acad Sci U S A **92:** 3338-3342

Fay EW (1903) Latin etymologies. The American Journal of Philology **24:** 67

Fedoroff N, Wessler S, Shure M (1983) Isolation of the transposable maize controlling elements Ac and Ds. Cell **35:** 235-242

Fedorova NJ (1946) Crossability and phylogenetic relationships in the main European species of *Fragaria*. Doklady Akademii Nauk SSSR **52:** 545-547

Feldmann KA (1991) T-DNA insertion mutagenesis in *Arabidopsis*. Plant J. **1:** 71–82

Feldmann KA, Marks MD (1987) Agrobacterium-mediated transformation of germinating seeds of *Arabidopsis thaliana*: A non-tissue culture approach. Mol Gen Genet **208:** 1-9

Flavell RB, Bennett MD, Smith JB, Smith DB (1974) Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochem Genet **12:** 257-269

Fleischmann R, Adams M, White O, Clayton R, Kirkness E, Kerlavage A, Bult C, Tomb J, Dougherty B, Merrick J, al. e (1995) Whole-genome random sequencing and assembly of *Haemophilus influenza* Rd. Science **269:** 496-512

Folta K, Dhingra A, Howard L, Stewart P, Chandler C (2006) Characterization of LF9, an octoploid strawberry genotype selected for rapid regeneration and transformation. Planta **224:** 1058-1067

Folta KM, Davis TM (2006) Strawberry genes and genomics. Critical Reviews in Plant Sciences **25:** 399-415

Folta KM, Kaufman LS (2000) Preparation of transcriptionally active nuclei from etiolated *Arabidopsis thaliana*. Plant Cell Reports **19:** 504-510

Folta KM, Staton M, Stewart PJ, Jung S, Bies DH, Jesdurai C, Main D (2005) Expressed sequence tags (ESTs) and simple sequence repeat (SSR) markers from octoploid strawberry (*Fragaria* × *ananassa*). BMC Plant Biology **5:** 12

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Molecular Biology Reporter **13:** 207-209

Gauch HG, Dugger Jr. WM (1953) The Role of Boron in the Translocation of Sucrose. Plant Physiol. **28:** 457-466

Gegenheimer P (1990) Preparation of extracts from plants. Meth. Enzymol. **182:** 174-193

Gidoni D, Rom M, Kunik T, Zur M, Izsak E, Izhar S, Firon N (1994) Strawberry cultivar identification using randomly amplified polymorphic DNA (RAPD) markers. Plant Breeding **113:** 339-342

Gilbert W (1978) Why genes in pieces? Nature **271:** 501-503

Goldberg RB (1978) DNA sequence organization in the soybean plant. Biochem Genet **16:** 45–68

Goldberg RB (2001) From cot curves to genomics. How gene cloning established new concepts in plant biology. Plant Physiol. **125:** 4-8

Graham J, McNicol RJ, McNicol JW (1996) A comparison of methods for the estimation of genetic diversity in strawberry cultivars. Theoretical and Applied Genetics **93:** 402 - 406

Greilhuber J, Doležel J, Lysak M, Bennett MD (2005) The origin, evolution and proposed stabilization of the terms 'genome size' and 'C-value' to describe nuclear DNA contents. Annals of Botany **95:** 255-260

Haas BJ, Wortman JR, Ronning CM, Hannick L, Jr RKS, Maiti R, Chan AP, Yu C, Farzad M, Wu D, White O, Town CD (2005) Complete reannotation of the *Arabidopsis* genome: methods, tools, protocols and the final release. BMC Biology **3**

Hadonou AM, Sargent DJ, Wilson F, James CM, Simpson DW (2004) Development of microsatellite markers in *Fragaria*, their use in genetic diversity analysis, and their potential for genetic linkage mapping. Genome **47:** 429-438

Hamilton RH, Künsch U, Temperli A (1972) Simple rapid procedures for isolation of tobacco leaf nuclei. Analytical Biochemistry **49:** 48-57

Hanahan D (1985) Techniques for transformation of E. coli. *In* DM Glover, ed, DNA Cloning: a Practical Approach. IRL Press, Oxford, pp 109-135

Hanania U, Velcheva M, Sahar N, Avihai P (2004) An improved method for isolating high-quality DNA from *Vitis vinifera* nuclei. Plant Mol Biol. Rep. **22:** 173-177

Hancock JF (1999) Strawberries. CABI Publ, Oxon

Hancock JF, Callow PA, Shaw DV (1994) Randomly amplified polymorphic DNAs in the cultivated strawberry, *Fragaria* × *ananassa*. J Amer Soc Hort Sci **119:** 862-864

Hancock JF, Serçe S, Portman CM, Callow PW, Luby JJ (2004) Taxonomic variation among North and South American subspecies of *Fragaria* virginiana Miller and *Fragaria chiloensis* (L.) Miller Can. J. Bot. **82:** 1632–1644

Harrison RE, Luby JJ, Furnier GR, Hancock JF (1997) Morphological and molecular variation among populations of octoploid *Fragaria virginiana* and *F. chiloensis* (*Rosaceae*) from North America. Am J Bot **84:** 612-620

Haymes KM, Henken B, Davis TM, van de Weg WE (1997) Identification of RAPD markers linked to a *Phytophthora fragariae* resistance gene (*Rpf1*) in the cultivated strawberry. Theor Appl Genet. **94:** 1097-1101

Haymes KM, Van de Weg WE, Arens P, Maas JL, Vosman B, Den Nijs APM (2000) Development of SCAR markers linked to a *Phytophthora fragariae* resistance gene and their assessment in European and North American strawberry genotypes. J Am Soc Hortic Sci **125:** 330–339

Helariutta Y, Kotilainen M, Elomaa P, Kalkkinen N, Bremer K, Teeri T, Albert V (1996) Duplication and functional divergence in the chalcone synthase gene family of Asteraceae: evolution with substrate change and catalytic simplification. Proc Natl Acad Sci U S A **93:** 9033-9038

Hofmeister F (1888) Zur Lehre von der Wirkung der Salze. Zweite Mittheilung. Arch. Exp. Pathol. Pharmakol. **24:** 247-260

Hummer KE, Sabitov A, Davis T (2005) Iturup and Sakhalin island strawberries. Hortscience **40:** 1127

Ichijima K (1926) Cytological and genetic studies on *Fragaria*. Genetics **11:** 590-604

James CM, Wilson F, Hadonou AM, Tobutt KR (2003) Isolation and characterisation of polymorphic microsatellites in diploid strawberry (*F. vesca* L.) for mapping, diversity studies and clone identification. Mol Ecol **3:** 171–173

Jeffreys A, Flavell R (1977) The rabbit square-globin gene contains a large insert in the coding sequence. Cell **12:** 1097–1108

Jin-Xia H, Li-Jia Q, Ji Y, Hao Y, Hong-Ya G (2004) A preliminary study on the origin and evolution of chalcone synthase (CHS) gene in angiosperms. Acta Botanica Sinica **46:** 10-19

Jones AS (1953) The isolation of bacterial nucleic acids using cetyltrimethylammonium bromide (cetavlon). Biochim Biophys Acta. **10:** 607-612

Kam-Morgan LNW, Gill BS, Muthukrishnan S (1989) DNA restriction fragment length polymorphisms: A strategy for genetic mapping of the D genome of wheat. Genome **23:** 724-732

Katterman FRH, Shattuck VL (1983) An effective method of DNA isolation from the mature leaves of *Gossypium* species that contain large amounts of phenolic terpenoids and tannins. Preparative Biochem **13:** 347-359

Kauzmann W (1954) Denaturation of proteins and enzymes. *In* WD McElroy, B Glass, eds, The Mechanism of Enzyme Action. Johns Hopkins Press, Baltimore.

Kay ER, Dounce AL (1953) The preparation of sodium ribonucleate with the use of sodium dodecyl sulfate. J. Am. Chem. Soc. **75:** 4041-4044

Keller B, Feuillet C (2000) Colinearity and gene density in grass genomes. Trends Plant Sci **5:** 246–251

Kim UJ, Shizuya H, Jong PJd, Birren B, Simon MI (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Research **20:** 1083-1085

King EE (1971) Extraction of cotton leaf enzymes with borate. Phytochemistry **10:** 2337-2341

Kirby KS (1956) A new method for the isolation of ribonucleic acids from mammalian tissues. Biochem. J. **64:** 405–410

Koch MA, Haubold B, Mitchell-Olds T (2000) Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, Arabis, and related genera (Brassicaceae). Molecular Biology and Evolution **17:** 1483-1498

Koes RE, Spelt CE, Mol JNM, Gerats AGM (1987) The chalcone synthase multigene family of *Petunia hybrida* (V30): sequence homology, chromosomal localization and evolutionary aspects. Plant Molecular Biology **10:** 159-169

Koller S (2001) Automated genomic DNA purification using the Wizard Magnetic 96 Plant DNA System. Promega Notes **79:** 25-28

Kunz W, Henle J, Ninham BW (2004) 'Zur Lehre von der Wirkung der Salze' (about the science of the effect of salts): Franz Hofmeister's historical papers. Current Opinion in Colloid & Interface Science **9:** 19-37

Landry BS, Rongqi L, Khanizadeh S, Dijkstra J (1997) Classification of 75 strawberry cultivars and breeding lines using RAPD markers. Acta Hort **439:** 101-105

Lerceteau-Köhler E, Guérin G, Laigret F, Denoyes-Rothan B (2003) Characterization of mixed disomic and polysomic inheritance in the octoploid strawberry (*Fragaria* × *ananassa*) using AFLP mapping. Theoretical and Applied Genetics **107**

Lerceteau-Köhler E, Roudeillac P, Markocic M, Guérin G, Praud K, Denoyes-Rothan B (2002) The use of molecular markers for durable resistance breeding in the cultivated strawberry (*Fragaria* × *ananassa*). *In* ISHS Acta Horticulturae 567: IV International Strawberry Symposium, Tampere, Finland

Leutwiler LS, Hough-Evans BR, Meyerowitz EM (1984) The DNA of *Arabidopsis thaliana*. Mol. Gen. Genet. **194:** 15-23

Levi A, Rowland LJ, Galletta GJ, Martelli G, Grego I (1994) Identification of strawberry genotypes and evaluation of their genetic relationships using random amplified polymorphic DNA (RAPD) analysis. Adv. Strawberry Res **13:** 36-39

Li H, Luo J, Hemphill JK, Wanf J-T, Gould JH (2001) A rapid and high yielding DNA miniprep for cotton (*Gossypium* spp.). Plant Mol. Biol. Rep. **19:** 1-5

Li J, Riehle MM, Zhang Y, Xu J, Oduol F, Gomez SM, Eiglmeier K, Ueberheide BM, Shabanowitz J, Hunt DF, Ribeiro JMC, Vernick KD (2006) *Anopheles gambiae* genome reannotation through synthesis of *ab initio* and comparative gene prediction algorithms. Genome Biology **7:** R24

Li W, Zhang P, Fellers JP, Friebe B, Gill BS (2004) Sequence composition, organization, and evolution of the core Triticeae genome. The Plant Journal **40:** 500-511

Liolios K, Tavernarakis N, Hugenholtz P, Kyrpides NC (2006) The Genomes On Line Database (GOLD) v.2: a monitor of genome projects worldwide. Nucleic Acids Res. **34:** D332–D334

Lipshitz R, Chargaff E (1956) Studies on nucleoproteins. IV. Preparation of the deoxyribonucleoprotein and fractionation of the deoxyribonucleic acid of wheat germ. Biochim. Biophys. Acta **19:** 256

Llop-Tous I, Dominguez-Puigjaner E, Palomer X, Vendrell M (1999) Characterization of two divergent endo-beta-1,4-glucanase cDNA clones highly expressed in the nonclimacteric strawberry fruit. Plant Physiol **119:** 1415-1422

Lodhi MA, Ye G-N, Weeden NF, Reisch BI (1994) A simple and efficient method for DNA extraction from grapevine cultivars,*Vitis* species and *Ampelopsis*. Plant Mol. Biol. Rep. **12:** 6-13

Logemann J, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant tissues. Anal. Biochem. **163:** 16-20

Loomis WD (1974) Overcoming problems of phenolics and quinones in the isolation of plant enzymes and organelles. Methods Enzymol **31:** 528-544

Mangelsdorf AJ, East EM (1927) Studies on the Genetics of *Fragaria*. Genetics **12:** 307-339

Manning K (1991) Isolation of nucleic acids from plants by differential solvent precipitation. Analytical Biochemistry **195:** 45-50

Milbourne D, Meyer RC, Collins AJ, Ramsay LD, Gebhardt C, Waugh R (1998) Isolation, characterisation and mapping of simple sequence repeat loci in potato. Mol. Gen. Genet. **259:** 233–245

Min Jou W, Haegeman G, Ysebaert M, Fiers W (1972) Nucleotide sequence of the gene coding for the bacteriophage MS2 coat protein. Nature **237:** 82-88

Monfort A, Vilanova S, Davis TM, Arús P (2005) A new set of polymorphic simple sequence repeat (SSR) markers from a wild strawberry (*Fragaria vesca*) are transferable to other diploid *Fragaria* species and to *Fragaria* × *ananassa*. Molecular Ecology Notes

Murai N, Sutton DW MM, Slightom JL Merlo DJ, Reichert NA, Sengupta-Gopalan C, Stock CA, Barker RF, Kemp JD, Hall TC (1983) Phaseolin gene from bean is expressed after transfer to sunflower via tumor-inducing plasmid vectors. Science **222:** 476–482

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. **8:** 4321-4325

National Plant Genomics Initiative (2002) Objectives for 2003–2008. Plant Physiol. **130:** 1741–1744

Nehra NS, Kartha KK, Stushnoff C (1991) Isozymes as markers for identification of tissue culture and greenhouse-grown strawberry cultivars. Can J Plant Sci **71:** 1195-1201

Nehra NS, Kartha KK, Stushnoff C (1991) Nuclear DNA content and isozyme variation in relation to morphogenic potential of strawberry (*Fragaria* × *ananassa*) callus cultures. Canadian Journal of Botany **69:** 239-244

Nourse SM, Fickus EW, Cregan PB, Hokanson SC (2002) Development of simple sequence repeat (SSR) molecular markers in strawberry. *In* SC Hokanson, AR Jamieson, eds, Strawberry research to 2001. ASHS Press, Alexandria, Virginia, pp 48-53

Parent JG, Page D (1995) Authentication of the 13 strawberry cultivars of Quebec's certification programme by random amplified polymorphic DNA analysis (RAPD). Can J Plant Sci **75:** 221-224

Peterson DG, Boehm KS, Stack SM (1997) Isolation of milligram quantities of nuclear DNA from tomato (*Lycopersicon esculentum*), a plant containing high levels of polyphenolic compounds. Plant Molecular Biology Reporter **15:** 148-153

Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH (2000) Construction of plant bacterial artificial chromosome (BAC) libraries: An illustrated guide. *In* Journal of Agricultural Genomics, Vol 5

Porebski S, Bailey LG, Braun BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol. Biol. Rep. **15:** 8-15

Potter D, Luby JJ, Harrison RE (2000) Phylogenetic relationships among species of *Fragaria* (Rosaceae) inferred from non-coding nuclear and chloroplast DNA sequences. Systematic Botany **25:** 337-348

Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc. of the IEEE **77:** 257-286

Rogstad SH (2003) Plant DNA extraction using silica. Plant Molecular Biology Reporter **21:** 463a-463g

Rosskopf EN, Chellemi DO, Kokalis-Burelle N, Church GT (2005) Alternatives to methyl bromide: a Florida perspective. *In* APS Feature Story,

Rout G, Samal S, Nayak S, Nanda R, Lenka P, Das P (2002) An alternative method of plant DNA extraction of cashew (*Anacardium occidentale* L.) for randomly amplified polymorphic DNA (RAPD) analysis. Gartenbauwissenschaft **67:** 114-118

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. *In* S Krawetz, S Misener, eds, Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386

Rychlik W, Spencer WJ, Rhoads RE (1990) Optimization of the annealing temperature for DNA amplification in vitro. Nucleic Acids Research **18:** 6409-6412

Salamov AA, Solovyev VV (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. **10:** 516-522

Sambrook J, Russell DW (2001) Molecular cloning: a laboratory manual, Ed 3rd. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press

Sargent DJ, Clarke J, Simpson DW, Tobutt KR, Arus P, Monfort A, Vilanova S, Denoyes-Rothan B, Rousseau M, Folta KM, Bassil NV, Battey NH (2006) An enhanced microssatellite map of diploid *Fragaria*. Theor Appl Genet. **112:** 1349-1359

Sargent DJ, Davis TM, Tobutt KR, Wilkinson MJ, Battey NH, Simpson DW (2004) A genetic linkage map of microsatellite, gene-specific and morphological markers in diploid *Fragaria*. Theoretical and Applied Genetics **109:** 1385 - 1391

Sargent DJ, Hadonou AM, Simpson DW (2003) Development and characterization of polymorphic microsatellite markers from *Fragaria viridis*, a wild diploid strawberry. Molecular Ecology Notes **3:** 550

Sargent DJ, Rys A, Nier S, Simpson DW, Tobutt KR (2007) The development and mapping of functional markers in *Fragaria* and their transferability and potential for mapping in other genera Theor Appl Genet. **114:** 373–384

Sawyer WH, Puckridge J (1973) The dissociation of proteins by chaotropic salts. The Journal of Biological Chemistry **248:** 8429-8433

Settles AM, Latshaw S, McCarty DR (2004) Molecular analysis of high-copy insertion sites in maize. Nucleic Acids Research **32:** e54

Sevag MG, Lackman DB, Smolens J (1938) The isolation of the components of streptococcal nucleoproteins in serologically active form. J. Biol. Chem. **124:** 425-436

Shapiro HS, Chargaff E (1960) Studies on the nucleotide arrangement in deoxyribonucleic acids. IV. Patterns of nucleotide sequence in the deoxyribonucleic acid of rye germ and its fractions. Biochim Biophys Acta. **39:** 68-82

Sjulin TM, Dale A (1987) Genetic diversity of North American strawberry cultivars. J. Am. Soc. Hort. Sci. **112:** 375–385

Soltis PS, Soltis DE (2000) The role of genetic and genomic attributes in the success of polyploids. Proc. Nat. Acad. Sci. **97:** 7051-7057

Staudt G (1973) *Fragaria iturupensis*, eine neue Erdbeerart aus Ostasien. Willenowia **7:** 101–104

Staudt G (2003) Notes on Asiatic species: III. *Fragaria orientalis* Losinsk. and *Fragaria mandshurica* spec. nov. Bot. Jahrb. Syst. **124:** 397–419

Staudt G (2005) Notes on the Asiatic *Fragaria* species: IV. *Fragaria iinumae*. . Bot. Jahrb. Syst. **126:** 163–175

Stein L (2001) Genome annotation: from sequence to biology. Nature Reviews Genetics **2:** 493-503

Sterck L, Rombauts S, Vandepoele K, Rouze P, Van de Peer Y (2007) How many genes are there in plants (... and why are they there)?

Sugimoto T, Tamaki K, Matsumoto J, Yamamoto Y, Shiwaku K, Watanabe K (2005) Detection of RAPD markers linked to the everbearing gene in Japanese cultivated strawberry. Plant Breeding **124:** 498–501

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Research **11:** 1441-1452

The American Heritage® (2006) Dictionary of the English Language, Ed Fourth Edition. Houghton Mifflin Company

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815

Thomas AJ, Sherratt HS (1956) The isolation of nucleic acid fractions from plant leaves and their purine and pyrimidine composition. Biochem J. **62:** 1-4

Travaglini EC, Meloni ML (1962) Extraction and separation of nucleic acids from cesium chloride homogenates of whole cells. Biochem Biophys Res Commun. **7:** 162-166

Tsai C-J, Harding SA, Tschaplinski TJ, Lindroth RL, Yuan Y (2006) Genome-wide analysis of the structural genes regulating defense phenylpropanoid metabolism in *Populus*. New Phytologist **172:** 47–62

Tuskan GA, Difazio S JS, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D. (2006) The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science **313:** 1596-1604

Van de Weg WE (1997) Resistance to *Phytophthora fragariae* var. *fragariae* in strawberry: the Rpf2 gene. Theoretical and Applied Genetics **94:** 1092 - 1096

Van Ooijen JW, Voorrips RE (2001) Joinmap 3.0: software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands

Venter JC AM, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. (2001) The sequence of the human genome. Science **291:** 1304-1351

Viruel MA, Sanchez D, Arus P (2002) An SSR and RFLP linkage map for the octoploid strawberry (*Fragaria* × *ananassa*). *In* Plant, animal and microbe genomes. Xth Conf, San Diego, California

Voet, Voet, Pratt (1998) Fundamentals of Biochemistry, Ed 1st edition. Wiley

Vogelstein B, Gillespie D (1979) Preparative and analytical purification of DNA from agarose. Proc Natl Acad Sci U S A **76:** 615-619

Voorrips RE (2002) Mapchart: software for the graphical presentation of linkage maps and QTLs. J Hered **93:** 77–78

Wang SY (2006) Fruits with high antioxidant activity as functional foods. *In* Functional Foods, pp 371-413

Watson JC, Thompson WF (1986) Purification and restriction endonuclease analysis of plant nuclear DNA. Methods Enzymol **118:** 57-75

Wein M, Lavid N, Lunkenbein S, Lewinsohn E, Schwab W, Kaldenhoff R (2002) Isolation, cloning and expression of a multifunctional O-methyltransferase capable of forming 2,5-dimethyl-4-methoxy-3(2H)-furanone, one of the key aroma compounds in strawberry fruits. Plant J. **31:** 755-765

Westphal O, Jann K (1965) Bacterial lipopolysaccharides: extraction with phenol-water and further applications of the procedure. *In* RL Whistler, ed, Methods in carbohydrate chemistry. Academic Press, Inc., New York, N.Y, pp 83–91

Westphal O, Luderitz O, F. B (1952) Über die extraktion von bakterien mit phenol/wasser. Z Naturforschung B. **7B:** 148–155

Wilcockson J (1973) The use of sodium perchlorate in deproteinization during the preparation of nucleic acids. Biochem J. **135:** 559-561

Wilhelm S, Sagen JE (1974) A history of the strawberry. From ancient gardens to modern markets. University of California, Berkeley. Division of Agricultural Sciences

Williamson R (1969) Purification of DNA by isopycnic banding in cesium chloride in a zonal rotor. Anal Biochem. **32:** 158-163

Williamson SC, Yu H, Davis TM (1995) Shikimate dehydrogenase allozymes: inheritance and close linkage to fruit color in the diploid strawberry. Journal of Heredity **86:** 75-76

Xu Q, Wen X, Deng X (2004) A simple protocol for isolating genomic DNA from chesnut rose (*Rosa roxburghii* Tratt) for RFLP and PCR analyses. Plant Mol. Biol. Rep. **22:** 301a-301g

Yandell MD, Majoros WH (2002) Genomics and natural language processing. Nature Reviews Genetics **3:** 601-610

Yu H, Davis TM (1995) Genetic linkage between runnering and phosphoglucoisomerase allozymes, and systematic distortion of monogenic segregation ratios in diploid strawberry. J Amer Soc Hort Sci **120:** 687-690

Yu J HS, Wang J, Wong GK-S, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Guangqiang Hu XH, Wenjie Li, Jian Li, Zhanwei Liu, Long Li, Jianping Liu, Qiuhui Qi, Jinsong Liu, Li Li, Tao Li, Xuegang Wang, Hong Lu, Tingting Wu, Miao Zhu, Peixiang Ni, Hua Han, Wei Dong, Xiaoyu Ren, Xiaoli Feng, Peng Cui, Xianran Li, Hao Wang, Xin Xu, Wenxue Zhai, Zhao Xu, Jinsong Zhang, Sijie He, Jianguo Zhang, Jichen Xu, Kunlin Zhang, Xianwu Zheng, Jianhai Dong, Wanyong Zeng, Lin Tao, Jia Ye, Jun Tan, Xide Ren, Xuewei Chen, Jun He, Daofeng Liu, Wei Tian, Chaoguang Tian, Hongai Xia, Qiyu Bao, Gang Li, Hui Gao, Ting Cao, Juan Wang, Wenming Zhao, Ping Li, Wei Chen, Xudong Wang, Yong Zhang, Jianfei Hu, Jing Wang, Song Liu, Jian Yang, Guangyu Zhang, Yuqing Xiong, Zhijie Li, Long Mao, Chengshu Zhou, Zhen Zhu, Runsheng Chen, Bailin Hao, Weimou Zheng, Shouyi Chen, Wei Guo, Guojie Li, Siqi Liu, Ming Tao, Jian Wang, Lihuang Zhu, Longping Yuan, Huanming Yang (2002) A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *indica*). Science **296:** 79-92

BIOGRAPHICAL SKETCH

Denise Cristina Manfrim Tombolato was born to Vadir and Marlene Tombolato on March 20, 1976, in Campinas, São Paulo, Brazil. She received her bachelor's degree in agronomic engineering from the Escola Superior de Agricultura "Luiz de Queiroz," at the University of São Paulo in 1998. While an undergraduate student in Brazil, Denise was granted the FAPESP and CNPq fellowships to investigate molecular markers for disease resistance in maize, working under the supervision of Dr. Luiz Eduardo Aranha Camargo at the Plant Pathology Department. During the last year of her undergraduate studies, she was introduced to University of Florida's professor Dr. Richard D. Berger, who was on sabbatical studies at ESALQ. Denise was invited to spend one semester in Dr. Berger's laboratory, carrying out investigation on plant disease epidemiology for the completion of her degree.

Dr. David Pete Weingartner granted her a research assistantship from 1999 to 2002, when she earned her master's degree from the Plant Pathology Department at the University of Florida. In 2002, she started her doctorate program at the Horticultural Sciences Department, where she attended the majority of the courses offered to the Plant Molecular and Cellular Biology program. Upon graduation, she will work as a genetic technologist for Ball Helix, the biotechnology research branch of Ball Horticultural Company in West Chicago. She is very excited about the upcoming changes in her life!