



Article **Crocus pallidus (Iridaceae)**—A Neglected Species for the **Bulgarian Flora and Critical Taxon in the Balkans**

Tsvetanka Raycheva¹, Kiril Stoyanov^{1,*}, Samir Naimov², Elena Apostolova-Kuzova² and Yulian Marinov³

- ¹ Department of Botany and Agrometeorology, Agricultural University—Plovdiv, 4000 Plovdiv, Bulgaria; raicheva.tzveta@gmail.com
- ² Department of Plant Physiology and Molecular Biology, University of Plovdiv, 4000 Plovdiv, Bulgaria; samir.naimov@gmail.com (S.N.); elena.apost@gmail.com (E.A.-K.)
- ³ Natural History Museum—Plovdiv, 4000 Plovdiv, Bulgaria; julianmarinov@abv.bg
- * Correspondence: nomtax@gmail.com

Abstract: For a long time, the Balkan endemic species *Crocus pallidus* has been unconfirmed and neglected for the flora of Bulgaria. It has remained an uncertain species from the Balkans, often listed as a synonym of *C. weldenii*. The morphological resemblance to the albinistic forms of *C. chrysanthus* has led to incorrect identification in the past, resulting in uncertainty regarding the distribution of this species in Bulgaria. In this regard, a detailed morphological and anatomical study of Bulgarian natural populations was carried out. A phylogenetic comparison in the ITS region of two Bulgarian populations of *C. pallidus* with other related species was conducted, revealing the distinction of *C. pallidus* populations from the closely related *C. weldenii*. The recently described new endemic species from Turkey, *C. thracicus*, shows very similar morphological and anatomical characteristics to *C. pallidus* and logically continues the distribution range in Eastern Thrace, along the Black Sea coast. We have a basis for suggesting that it should be treated as a synonym of *C. pallidus*.

Keywords: Iridaceae; Bulgaria; ITS region; *Crocus pallidus; Crocus thracicus; Nudiscapus;* anatomy; morphology

1. Introduction

The genus Crocus L., with a center of diversity on the Balkan Peninsula and Asia Minor [1], constitutes 235 species [2]. The number of species has increased due to the descriptions of the new species from this area in the last few years [3,4]. However, a taxonomic revision of the genus Crocus in Bulgarian flora has not been conducted in the last 50 years. After the last taxonomic treatment in Bulgaria, the genus included nine species [5]. Crocus pallidus Kitan. & Drenk. was reported as a new species, based on the revised specimens of C. chrysanthus Herb. var. albidus Maw, collected from the region of North Macedonia [6]. Later, this species persisted in the research of Kitanov et al. [7] and was included in the list of species for the flora of Dobrudja [8]. For more than 35 years, this species has been neglected in the Bulgarian floristic literature and identification keys [9]. However, C. pallidus was mentioned in a review for the Balkan endemics in Bulgaria [10]. On the basis of the publication of Kitanov et al. [7], the fourth edition of "Conspectus of the Bulgarian Vascular Flora" [11] included *C. pallidus*, with a distribution range along the Black Sea Coast (Northern) and Northeast Bulgaria. Kitanov and other authors collected the available deposited materials from Bulgarian sites during 1965–1972. The limited studies so far have caused a lack of morphological and anatomical characteristics for C. pallidus in the referent taxonomic sources. Various authors have perceived white-flowered crocuses in the Balkans with different taxonomic statuses. Pulević [12] considered that the specimens described by Kitanov & Drenkovsky represent C. chrysanthus var. citrinus Velen. Crocus *pallidus* was neglected by Mathew [13] and noted as a heterotypic synonym of *C. weldenii*



Citation: Raycheva, T.; Stoyanov, K.; Naimov, S.; Apostolova-Kuzova, E.; Marinov, Y. *Crocus pallidus* (Iridaceae)—A Neglected Species for the Bulgarian Flora and Critical Taxon in the Balkans. *Plants* **2022**, *11*, 686. https://doi.org/10.3390/ plants11050686

Academic Editor: Marco C. Simeone

Received: 29 January 2022 Accepted: 28 February 2022 Published: 2 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). Hoppe & Fürnrohr. This has probably been the reason for *C. pallidus* being listed as a synonym of *C. weldenii* in the system of WCSP [14].

Crocus weldenii is a taxon with controversial taxonomic status. It has been perceived as a subspecies of *C. biflorus* s.l. [13]. This taxon is Balkan-endemic and Illyrian–Adriatic-endemic, distributed from Trieste to the border of Albania, with general distribution in Slovenia, Croatia, Herzegovina, Montenegro, Serbia, and North Macedonia [15]. The northwestern border of the species' range reaches the vicinity of Trieste in NE Italy [16–18].

According to Randelović et al. [19], both species are found in Serbia. *Crocus weldenii* inhabits the hilly and southern regions, while *C. pallidus* is a species from the mountainous and more northern areas. The taxonomical status of *C. pallidus* remains very unclear [2].

All members of the genus *Crocus* have shown exclusively high diversity in their chromosome numbers [20]. The chromosome number of *C. biflorus* sensu Mathew (1982) varies between 2n = 8 and 2n = 36 [21]. The changes in the chromosome number in the related species play an important role in the speciation [22,23]. This could be an important feature if the close related taxa are different according to taxonomic numbers. Nevertheless, the representatives of the group discussed here have different morphology and the same chromosome number. The cytological models in *Crocus* are complex, with extensive dysploidy, and not satisfactorily explained [13].

The white-flowering crocuses on the Balkan Peninsula and Adriatic coast [19] have not been covered by the recent molecular and phylogenetic studies [24,25]. Furthermore, no annotated ITS sequences of *C. pallidus* were found in the NCBI GenBank Nucleotide database [26].

In the protologue of the older taxonomic literature, the morphology of *C. pallidus* has been laconic and incompletely reflected [5,17], which also leads to uncertainty of its status. A species named *C. thracicus* with a similar but more detailed morphology was described in northwest Turkey [27].

The information about the species in the Bulgarian floristic literature is based on sporadic reports, without precise localities and current diagnostic data. The various assumptions about the taxon have led to confusion and neglect in the Bulgarian literature. Moreover, the contradictory taxonomic treatments of white-flowered Crocuses from the Balkans need further study of a more thorough sample.

In this regard, our study aimed to provide detailed information on natural Bulgarian populations of *C. pallidus*, using the traditional morphological method, leaf anatomy, and an analysis of the literature data. To clarify the phylogenetic position among the related taxa, we used ITS internal and external spacers as taxonomically useful phylogenetic markers in *Crocus* [28].

2. Results

2.1. Description Based on Bulgarian Materials

A spring synanthous geophyte is shown in Figure 1 with the following features: corm oval-suborbicular, 13.5–25 mm in diameter, with membranous-coriaceous tunic, yellowish-brown, with separating basal rings; Basal rings 1.2–2.5 mm thick (Figure 1G) with unclear, slightly visible irregular dentation; plant height 100–210 mm; leaves 4–6(–7), 0.8–2.3 mm wide; flowers 2–3(–4); perigone segments niveous-white or outer ones with sprayed pale-blue coloration (Figure 1A,C); rounded to slightly acuminate; 19–33 × 6–11 mm long and 15–26 × 5–12 mm wide; perigone tube 6–10 cm long, white or bluish-lilac. (Figure 2); throat glabrous, white; filaments 3.8–10 mm; anthers 8–12(16) mm long with blackish basal lobes, or exceptionally with blackish stripes on the internal side, along the thecas (Figure 1B), on both sides of the white connective; stigma orange-reddish, trilobate (stylodia 3–5 mm), usually shorter than or equal to anther length; Capsule ellipsoid 15–21 × 4.3–7.4 mm (Figure 1D); seeds oval, dark brown 2.3–3.4 × 1.5–2.1 mm.





Figure 1. Morphological features of *Crocus pallidus* (specimen SOA 062791, grid 1 mm): (**A**) whole plant; (**B**) anthers and stigma; (**C**) flower section; (**D**) mature capsules; (**E**) seeds; (**F**) corm; (**G**) basal rings and bottom of the corm.





Figure 2. Variation in the color of the outer perigone segments of *Crocus pallidus*—white and speckled violet (population with voucher 062791).

A comparison with the known morphological features of the closely related species is provided in Table 1.

Table 1. A comparison of morphological features of *Crocus pallidus* (current data), *C. weldenii* [13,17],and *C. thracicus* [27].

	C. pallidus	C. weldenii	C thracicus
Corm	13.5–25 mm	13–19.7 mm	10–12 mm
Plant height	(8)10–20 cm	8–17 cm	7.35–12 mm
Teeth of the basal rings	Irregular	Regular/Irregular	Irregular
Count of leaves	3–5	3–5	3-4(-5)
Leaf width	0.84–2.3 mm	max 1 mm	0.75–1.3 mm
Ribs on the abaxial leaf surface	Missing or vaguely visible	1–2, clearly visible	Missing
Color of the perigone tube	White, rarely violet	White, often violet	White, rarely violet near the apex
Color of outer perigone segments	White or sprayed in violet	White	White or sprayed in violet
Size of perigone segments	$19-33 \times 6-11/$	10.27 × 4.8.10.8	$17-24 \times 6-9/$
(out/in)	$15-26 \times 5-12 \text{ mm}$	$19-37 \times 4.8-12.8$	$14.6-23 \times 5.7-8.7 \text{ mm}$
Perigone throat	Glabrous, white	Yellow	Glabrous, white
Filaments	3.8–13.7 mm	8–16 mm	10–13.5 mm
Anthers	8–16 mm, yellow, with blackish basal lobes, or entirely black edge	yellow, without blackish lobes	7.3–11.8 mm, yellow, with blackish basal lobes.
Capsule	Ellipsoid; 15–21 mm long; 4.3–7.4 mm wide.		Ellipsoid, about 14 mm long
Seeds	2.2–3.4 $ imes$ 1.5–2.1 mm, with convex caruncle		2.5 mm in diameter, with convex caruncle
Distribution	Bulgaria (Eastern parts), North Macedonia, Serbia.	Italy, Albania, Serbia.	Turkey in Europe (Thrakia)
Elevation	30–190 m	100–750 m	45–170 m
Flowering period	February–March		February–March

2.2. Distribution

The species is Balkan-endemic. The distribution area extends over the territory from Serbia (eastern and southeastern), through Kosovo, North Macedonia, Bulgaria, and prob-



ably the European part of Turkey. The Black Sea coast represents the eastern border (Figure 3).

Figure 3. Localities of *Crocus pallidus* (**A**) in Bulgaria and *C. thracicus* (**B**) in Turkey. Map created at GPSVisualizer.com (accessed on 16 March 2021). Leaflet | NGS maps from ESRI/ArcGIS.

2.3. Phenology and Habitats

The flowering period lasts from the end of January to the end of March. The population of *C. pallidus* near the village of Kamen Bryag (SOA 062791) inhabits the open grass communities in the margin of an oak grove, accompanied by individuals of *Paliurus spinachristi* Mill., *Ficaria verna* Huds., *Crocus chrysanthus* (Herb.) Herb., *Geranium mole* L., *Fragaria vesca* L., *Viola* sp. div., *Trifolium* sp. div., Poaceae sp. div., and tufts of mosses. The whole habitat is surrounded by agricultural land. Approximately 150–200 individuals represent the population. A dense sympatric population of *C. chrysanthus* was found inside the grove. On the margin of the grove, populations of both species are mixed, whereas hybrids may occur (Figure 4). The locality is under anthropogenic pressure due to agricultural work, the main road laying on the western border of the population, and proximity to Kamen Bryag.



Figure 4. *Crocus pallidus* (A), *C. chrysanthus* (C), and their hybrid (B) in the population with voucher 062791.

The population in Pomorie Narrow (SOA 062791) exists in sparse oak forest with the same accompanying species and *Anemone blanda* Schott & Kotchy. The few observed individuals of *C. chrysanhus* were flowering earlier than *C. pallidus* flowering time. The anthropogenic pressure is also significant (the habitat is an illegal landfill for household waste).

Under high anthropogenic pressure is the locus classicus of *C. p.* f. *bulgaricus* Kitan. & Drenk. It is an oak grove, subjected to unregulated logging and pollution, situated on the main road near the city of Dobrich (SO 21305).

The population near the Rudnik suburb of Burgas (SOA 063064) is under relatively low anthropogenic pressure. It is located in an open habitat, protected by the shrubs of *Paliurus*, and accompanied by *C. chrysanthus*.

2.4. Leaf Anatomy

The anatomical features of *C. pallidus* from the evaluated populations are presented in Table 2.

Table 2. Anatomical parameters of the examined populations of *Crocus pallidus*. Values are given as a range (minimum–maximum), mean \pm standard deviation.

Population *	062791	062797	063064
	Kamen Bryag	Pomorie Narrow	Roudnik
Section width, µm	1701-2224	955-1722	910-2017
	1927.6 ± 133	1210.9 ± 193	1432.6 ± 327.9
Section height, µm	709-807	356-664	480-633
	761.1 ± 39.8	481.8 ± 89.2	562.8 ± 54.8
Arm width, μm	657-973 830.1 ± 96.3	$385-696 \\ 547.2 \pm 91.1$	455-1071 703.4 ± 146.6
White stripe to leaf width ratio, %	$15.1-21 \\ 17.4 \pm 2$	$\begin{array}{c} 12.7 – 28.8 \\ 20 \pm 4.2 \end{array}$	$\begin{array}{c} 14.9 {-} 19.3 \\ 16.64 \pm {1.5} \end{array}$
Vascular bundles, count	$\begin{array}{c} 2325\\ 23.8\pm1\end{array}$	$\begin{array}{c} 1315\\ 13.8\pm1\end{array}$	$15-17 \\ 15.5 \pm 0.3$

Population *	062791 Kamen Bryag	062797 Pomorie Narrow	063064 Roudnik
Vascular bundles, height	25.9–177	27.5–187	33.2–190
um	81.7 ± 45.3	95.6 ± 45.8	104.6 ± 50.1
Vascular bundles, width,	17.9–103	22.9–141	23.9–127
um	59.2 ± 24.6	73.2 ± 36.5	74 ± 29.8
Vascular tissues in the	22.2-50.7	7–61	4.1-30.8
bundles, %	37.3 ± 8.2	19.9 ± 9.3	19.4 ± 7.7
Palisade tissue:	13.7-77	38-80	25.1-75
thickness, μm	44 ± 14.5	56.8 ± 8.8	57.9 ± 11.2
Spongy tissue-thickness,	25-91	43-104	33.8-65
μm	48 ± 16.6	61.6 ± 13.8	46.3 ± 9.7
	287-407	150-295	247-305
White stripe, µm	341 ± 49.2	241.1 ± 44.5	276.9 ± 22.8
Adaxial epidermal cell:	173-290	98-431	155-473
length, μm	236 ± 47.3	237.2 ± 61.7	286.2 ± 66.8
Adaxial epidermal cell:	15-23.8	11–22.6	6.7–21
height, μm	18.5 ± 1.96	17.83 ± 2.56	15.3 ± 3
Adaxial epidermal cell:	10.4-23.1	11.5-20.8	9.8-23.5
width, μm	17.5 ± 2.93	16.45 ± 2.4	16.8 ± 3
Deline de celle le ciele (com	13.8–36.4	21.3-40.7	17.2-41.3
Palisade cell: neight, µm	26.4 ± 4.25	32.1 ± 6.49	30.3 ± 5.3
Dalica da calle reidth um	8.8-21.2	9.6–18.6	8.9–17.5
Fallsade cell: width, µm	13.9 ± 2.47	13.53 ± 2.46	12.7 ± 2
Cooper cally beight um	14.4-29.6	19.1–25.1	13.9–28.6
Spongy cell: height, µm	21.4 ± 3.14	22.36 ± 2.07	20.7 ± 4.1
Spangy cally width	8.3-23.1	13.7-17.9	7.4–21.5
spongy cell: width, µm	14.7 ± 3.39	15.74 ± 1.57	14.2 ± 2.6
Abaxial epidermal cell:	11.2-24.3	15–28	7.7-18.1
height, μm	19.6 ± 2.92	18.2 ± 3.33	13.5 ± 2.7

Table 2. Cont.

* The populations are represented by the number of the voucher specimens in SOA.

The leaves in genus *Crocus* have a unique bifacial profile of the cross-section, with central rectangular or square keel, and two lateral arms.

In the investigated populations, the keel is square to rectangular, with a wide base, and with a wide white stripe (15–30% of the leaf width). The lateral zones are convex due to the developed mesophyll around the bigger vascular bundles (Figure 5).



Figure 5. *Crocus pallidus* leaf cross-sections $(4 \times)$. Abbreviations: ad—adaxial side, ab—abaxial side, la—lacuna area. Voucher numbers are shown near the cross-sections.

100 µm 200 µm

The arms' length and their curving degree vary in the investigated populations. The ribs are present on the abaxial side of the arms (Figure 6B,E). Papillae are located in the apical regions of the arms and along the ribs (Figure 6A).

Figure 6. Leaf anatomy of Crocus pallidus. (A) Papilles on the ribs-adaxial surface; (B) section of the shoulder (SOA 062791); (C) detail of the leaf with visible adaxial stomata (SOA 062797); (D) crystals; (E) subterminal vascular bundle in the shoulder; (F) adaxial epidermis with stomata; (G) abaxial epidermis with stomata (SOA 062791). Abbreviations: ad-adaxial side, ab-abaxial side, la-lacuna area, e—epidermis, pp—palisade parenchyma, sp—spongy parenchyma, sc—sclerenchyma cap, ph-phloem, xy-xylem.

The adaxial epidermal cells are rectangular, with a thicker cuticle. The abaxial epidermal cells are elliptic in shape, while the cuticle layer is thinner. The stomata are located on the abaxial side of the leaves, (Figure 6E,G) in the zones of the arms and keel. The stomatal type is anomocytic. Single stomata are also sporadically present on the abaxial epidermis (Figure 6C,F).

The parenchyma in the central zone of the keel is represented by big rectangular cells with thin walls and without chloroplasts, forming a lacuna zone, visible like a white stripe along the leaf blade. The assimilating mesophyll in the zone of the arms consists of palisade and spongy layers. The palisade cells are oblong, organized in two rows, without any space between them. The spongy cells form zones of 2–4 rows near the abaxial epidermis, between the ribs. They have an elliptic to irregular shape, with slightly visible intercellular spaces between them (Figure 6C,E). Crystalline formations as oxalate sand or crystals with an approximately cubic shape are present in all three studied populations of C. pallidus (Figure 6D).

The vascular bundles are collateral, located in a row along the arms and the keel zone. The vascular tissues take 7-40(-61)% while the remaining bundles (about 2/3) consist of sclerenchymatous "caps". The subterminal bundles in the folding area of the arms are bigger than the other (81–187 μ m \times 56–141 μ m). The terminal bundles in the keel are the biggest in that area (89–190 μ m \times 70–136 μ m). The number and size of the other bundles



vary depending on the width of the section. The population from Kamen Bryag (SOA 062791) shows a bigger number of vascular bundles. Despite the quantitative variation in their number, in all of the sections, the biggest vascular bundle is subterminal, followed by 1–3 narrower terminal bundles in the shoulder.

2.5. ITS Sequences

On the basis of the ITS sequences, the phylogenetic tree (Figure 7) shows the known grouping of the species in section *Nudiscapus* B. Mathew. Two separated branches contain the series *Biflori* and the series *Reticulati*. The branch of series *Biflori* contains the Bulgarian samples of *C. pallidus* (in bold blue in the tree) together with *C. minutus*, *C. leucostylosus*, *C. ranjeloviciorum*, *C. punctatus*, *C. alexandrii*, *C. weldenii*, *C. pulchricolor*, and *C. adamioides*. A highly divergent subgroup contains the species from the Aegean islands and Italy (*C. biflorus*, *C. babadagnensis*, and *C. rhodensis*). In a divergent position remains a group of specimens determined as *C. chrysanthus*.



Figure 7. Placement of *Crocus pallidus* in the phylogenetic tree of sect. *Nudiscapus* obtained by Bayesian phylogenetic inference of the nuclear rDNA ITS regions using the methodology of Harpke et al. [28]. Posterior probabilities are designated by numbers. See Table A1 for details.

3. Discussion

Crocus pallidus belongs to subgenus *Crocus*, section *Nudiscapus*, series *Biflori* [13]. Due to the taxonomic changes and the growing number of the new species described, especially in the critical and volumetric section *Nudiscapus*, more detailed morphological descriptions are necessary for identification [29]. A disadvantage in the descriptions of the crocuses by earlier authors is the lack of detailed and comprehensive diagnostic descriptions. In the diagnosis by Kitanov and Drenkovski [6], as well as in the description by Ranđelović et al. [17], we did not find the data about the color of the anther's basal lobes, or about the color variation in the outer perigone segments in *C. pallidus*. On the other hand, in the herbarium materials, deposited by Kitanov (SO 69556, 32784, 32795), the blackish lobes of the anthers are distinctly visible, although some of the samples are damaged and unsuitable for analysis (SO 13220; 103010). We suspect that this feature was overlooked by the authors because no other white-flowering *Crocus* species are known for the flora of Bulgaria.

Kitanov and Drenkovski [6] reported the two localities of C. pallidus in Bulgaria. According to a revision of the herbarium material from northeast Bulgaria of C. chrysanthus (SO 21305), the authors published a new form C. pallidus f. bulgaricus. The described variability is based on the relation between the leaf and the flower size. Crocus p. f. pallidus from North Macedonia (locus classicus, SO 30393) and Serbia is characterized by the leaves shorter than flowers, while f. *bulgaricus* can be recognized by the longer leaves compared to flowers. This form can be found in the region of Balchik (SO 13220) and Dobrich (SO 21305). So far, a reasonable position has been expressed in the literature for *C. pallidus* f. bulgaricus. Since the species and all its infraspecific taxa are equated as synonyms of *C. weldenii*, we consider that it is necessary to specify the infraspecific variability on the basis of the collected material and the data from the relevant literature sources. The variation in the meristic features of the evaluated populations is related to specific ecological conditions of the habitats. In our opinion, this variability is not a sufficient basis for taxonomic ranking. The studied populations also reveal polymorphism in the domain of leaf anatomy (Figure 5). The leaves of the plants near Kamen Bryag have two indistinct ribs along the arms, as mentioned by Rukšāns [2]. This feature may be related to the hybridization process and high variability in this population. The leaves of the other two populations have no ribs in the arms as reported for *C. thracicus* [27].

Our study reveals the presence of nonfunctioning stomata on the abaxial epidermis in the two populations from the Black Sea coast (Figure 6F). This detail was not registered in the population from Roudnik. Data about the abaxial stomata in *Crocus* were not found in the relevant anatomical literature. Moreover, the available information about the leaf epidermal structures is pretty limited [28,30].

Crocus chrysanthus and *C. pallidus* are spring synanthous geophytes, with an overlapping flowering period. They are often found in sympatric populations, occasionally resulting in hybrids between them (Figure 4). Our morphological findings correspond to the descriptions of *C. pallidus* [2] and *C. thracicus* [27], with minor quantitative differences. For example, the corms in the population from Kamen Bryag are much bigger than the other examined populations and the data from the literature. This feature could be a consequence of the differences in the biotope (soil type, humidity). Very variable color of the outer perigon segments, from pale blue to white with a slightly visible yellow stripe, was noted in the population near Kamen Bryag. A probable reason for the variation is the overlapping flowering period with *C. chrysanthus*. We noticed the hybrid with an intermediate coloration of the perigone segments (white and yellow) in the same population (Figure 4B).

The existence of clear and mixed populations demonstrates that *C. pallidus* is a distinctly differentiated species. The spontaneous hybrid forms and the unclear morphological differentiation of the features between *C. pallidus* and *C. chrysanthus* s. 1. are evidence of the close relationship and probably recent divergence, assuming the distribution area of both species. We consider that introgressive hybridization is a reason for the higher variability in the mixed population. Sympatric populations of *C. pallidus* with *C. crysanthus* with overlapping flowering time have already been mentioned [17]. A similar feature was stated for C. thracicus \times C. chrysanthus [27]. Our investigation shows a morphological similarity between C. pallidus and C. thracicus. Since an ITS sequence of C. thracicus has not been deposited in the NCBI Nucleotide database, its phylogenetic relationship with Bulgarian populations of C. pallidus remains unknown. Geographically, populations of *C. thracicus* are located very close to the Bulgarian localities of *C. pallidus*, especially those in the south (Figure 3). Furthermore, the authors compare *C. thracicus* with the closely related C. alexandri Ničič & Velen. and C. weldenii, while C. pallidus is mentioned but neglected as a morphological description at the same time. According to Yüzbaşioğlu et al. [27], C. pallidus has a yellow throat. This qualitative feature is not correctly cited. In our study on four Bulgarian populations, as well as in the deposited herbarium specimens (see Section 4), the perigone throat is white. We found the same data about *C. pallidus* in the literature (Mathew, 1982; Rukšāns 2017) and in the evaluated herbarium specimens in SO. Kitanov and Drenkovsky [6] and Randelović et al. [17] did not describe the key features such as the throat color, as well as the dark basal lobes of the anthers, in C. pallidus. The examined specimens, including the deposited material by Kitanov and Drenkovsky, stably show the white throat and blackish basal lobes of the anthers. The same morphology has been noted in the description of C. thracicus. The reason for this assumption is also given by the close morphological and anatomical parameters of the two species, the habitat conditions, and the elevations (Table 1). In the white-flowering crocuses group of the north of Balkan Peninsula and southeastern Europe, the only member with a yellow throat is *Crocus malyi* Vis. (section Crocus, series Versicolores) with a distribution area on Velebit Mts. in Croatia, Bosna, and Herzegovina [2,31,32].

Molecular and phylogenetic methods and analyses are reliable tools for explaining the complex relationships between taxa in the nonmonophyletic section *Nudiscapus*, particularly the white-flowered species of the *Biflori* series distributed in the Balkans. The sequences of both populations of *C. pallidus* with Bulgarian origin are identical. The generated phylogenetic tree with the *C. adamii* s. str. group as the outgroup complements the data from Harpke et al. [28] without major changes in the topology. The samples of *C. pallidus* are situated in the group of the polymorphic series *Biflori*, together with *C. minutus* Kerndorff & Pasche [33] from Turkey (Lycia), *C. punctatus* Kerndorff, Pasche & Harpke [24] from southern Turkey (Isparta and Burdur), and *C. leucostylosus* Kerndorff, Pasche & Harpke in [24], with a native range Western Turkey (Denizli), *C. alexandri*, *C. weldenii*, and *C. randjeloviciorum*. The big clade in ser. *Biflori* contains species with discrete anatomical and morphological differences; however, according to the ITS similarity, they are genetically closely related and remain in an unresolved phylogenetic position.

4. Materials and Methods

4.1. Examined Specimens

The samples were collected during terrain work in 2019–2021. We used the materials deposited in the national herbariums to compare some qualitative and metric features. The differences between this taxon with the related *C. thracicus* and *C. weldenii* were taken from protologues and photos of herbarium specimens and were reviewed in a table.

The examined specimens are listed below. The vouchers for the ITS sequences are signed with their NCBI numbers. New and unpublished chorological data are signed with an asterisk (*). The specimens are listed by country, floristic region, MGRS square, toponym, elevation, exact geographic coordinates (if present), date, and collector, followed by herbarium acronym and entry number.

Crocus pallidus:

North Macedonia: 34TDM71. Near Mavrovo. 19 March 1968 (coll. Drenkovski & Kitanov) SO 32794; 34TEM11. "Selo Zrkle, Porec–dolina na Treska", near the village of Zrkle, Treska Valley, 2 March 1972 (coll. B. Drenkovski sub *C. chrysanthus* Herb var. *albidus* G.Maw, rev. B. Kitanov, Type) SO 30393; 34TEM36. Skopska Crna Gora, 600 m. 19 February 1972 (coll. Drenkovski sub *C. crysanthus* Herb. var. *albidus* G.Maw; rev. Kitan. et

Drenk, Isotype) SO 30392, 32795; BULGARIA: * Black Sea Coast (Southern): 35TNH52. Pomorie Narrow–Pandarlak locality, N42.71911 E27.72632, 183 m, 27 February 2020 (coll. Raycheva & Stoyanov) SOA 062797 (NCBI record MW775330). Black Sea Coast (Northern): 35TNH89. "Zlatni Pyasatsi" Nature park, N43.303 E28.034214, 250 m, 13 March 2015 (coll. Y. Marinov) SOA 063074; 35TNJ90. Near the town of Balchik, April 1965 (coll. Minchev & Kitanov sub *C. biflorus* Mill. var. *albus* Herb.; rev. *C. pallidus* Kitan. 20 Mar. 1973) SO 13220; 35TPJ21. Kamen Bryag, 7 March 1973 (coll. Georgieva & Kitanov) SO 69557, 8 March 1973 (coll. Dimova & Kitanov) SO 69556; N43.45199 E28.54402, 41 m, 28 February 2020 (coll. Raycheva & Stoyanov) SOA 062791 (NCBI record MW775331). Northeast Bulgaria: 35TNJ62. Kabaklaka locality, near the city of Dobrich, May 1971 (coll. Dencheva; Type: C. p. f. bulgaricus Kitan. & Drenk.) SO 21305; 35TNJ70. The village of Batovo, 1975 (coll.?) SO 103010; 35TNJ80. Dezpelin locality, near the village of Topola, 59 m, N43.403085 E28.069111, 7 April 2020 (obs. Zh. Barzov) photo set; 35TPJ14. Bezhanovo, 36-69 m, N43.71868 E28.41449, 16 February 2018 (obs. Zh. Barzov) photo set. * Toundja Hilly Plain: 35TNH42. Noth of the Roudnik neighbourhood of Bourgas, 150 m, N42.638119 E27.489508, 27 February 2021 (coll. T. Raycheva & K. Stoyanov) SOA 063064.

Crocus thracicus (photo scans):

Turkey: A1(E), Kırklareli: Vize, Saray-Vize yolu, c. 10 km, Quercus sp. & Paliurus spina-christi Mill. açıklıkları, 170 m, 8 February 2014 (coll. S. Yüzbaşioğlu, S. Aslan, İ. Sözen & F. Canız, **Holotype**–ISTE 102922; **Isotype**–DUOF 5630.

Crocus weldenii (photo scans [34-38]):

Croatia: Istrien Kvarner/Carnaro/Quarner Bucht, okot Krk/isola di Veglia/Insel Vögls, NW Draga Bašćanska/Bescavalle, E-exponierter Hang zum Bach Suha Ričina/ Fiumera, W über dem Wasserfall; 14°41′12,40″ E 45°01′09,61″ N, 174 m, 16 February 2019 (coll. Rottensteiner, Walter K. & Rottensteiner, Marica) GJO 0099926; Gemeinde Bakar; 30–300 m Seehöhe; 14°32′00″ E 45°18′25″ N, 4 February 1886 GJO 004853; mt. Marian, nan loin de Spalato, May 1830 (coll. Steudel) K 000802457; **ITALY**: Istria. In rupestribus ad Gabrovizza prope Prosecco (coll. Solla) O V 2187017; **cult**. L 1476872.

4.2. Anatomical Investigations

Quantitative measurements and observations of qualitative features of fresh samples for the morphological analysis were conducted. To examine the leaf anatomy, 10 individuals per population, from three populations in total, were collected during the flowering time and conserved in 75% ethanol. Transverse cross-sections and epidermal areas of leaves from each individual were manually constructed from the middle part. The epidermis was peeled using a surgical scalpel. The microscopic slides were fixed with glycerin. The snapshots of the microscope slides were taken using a Leica 750 digital microscope, while the measurements of the characters were completed using Micam 2.4 software [39]. The total of 24 anatomical features included section width and height, arm length, white stripe width, the number and size of vascular bundles, palisade tissue height, height and width of palisade and spongy cells, and the height and width of the adaxial and abaxial epidermal cells. Abaxial epidermal cells, including the size of the stomata, were measured in the area of the stomatal rows (between the ribs). The values were expressed through the basic statistic parameters (mean, standard deviation, maximum and minimum).

4.3. Molecular Methods

Three populations were evaluated by anatomical parameters. DNA, isolated from two of the populations, was used for sequencing of the small subunit ribosomal RNA gene, partial sequence (ITS 1, 5.8S ribosomal RNA gene, and ITS 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence). The genomic DNA was isolated using a Dneasy Plant Mini Kit (QIAGEN). Briefly, 50 μ g of plant material was ground in liquid nitrogen and processed according to the manufacturer's requirements. DNA concentration and quality were determined spectrophotometrically at a wavelength of 260 nm using an

Epoch microtiter plate reader and T3 plate protocol. The ITS region (ITS1, 5.8S rDNA, ITS2) was amplified using the following primers: ITS-A (5'–GGAAGGAGAAGTCGTAACAAGG-3') and ITS-B (5'–CTTTTCCTCCGCTTATTGATATG–3'), as described in Tirel et al. [40]. The 50 μ L reaction mixture contained 1× reaction buffer, 200 μ M of dNTPs, 0.2 μ M of each primer, 100 ng of genomic DNA, and one unit of Q5 High Fidelity DNA polymerase (New England Biolabs). The PCR thermal cycler steps were as follows: initial melting of the reaction mixtures at 94 °C for 45 s, followed by 30 cycles at 94 °C for 10 s for denaturation, 10 s at 62 °C for primer annealing, 30 s at 72 °C for primer extension, and a final elongation step of 2 min at 72 °C. Amplified PCR products were separated by 0.8% agarose gel electrophoresis, excided from the gel, and purified using a QIAquick Gel Extraction Kit (QIAGEN). The purified DNA fragments were subsequently sequenced with Microsynth Company (Balgach, Switzerland) technology. Chromatograms were corrected manually with DNAStar software (Lasergene, Madison, WI, USA). The sequences were uploaded to the NCBI Gene database under accession numbers MW775331 (voucher SOA 062791), MW775330 (voucher SOA 062797), and OM368594 (voucher SOA 062596).

4.4. Phylogenetic Analyses

The obtained nucleotide sequences were blasted against nucleotide sequences from the NCBI Nucleotide database [41]. The best hits were downloaded and used for phylogenetic analysis. The alignment of the sequences (Figure A1) was achieved using the ClustalW Multiple alignments [42]. The phylogenetic analysis was conducted using Bayesian phylogenetic inference with MrBayes 3.2 [43]. The parameters of the analysis were the same as described by Harpke et al. [28]: 2×4 chains for two million generations, nuclear data set $\Gamma TP + G + I$, sampling tree per 1000 generations, two independent runs. The result was visualized as a tree using TreeGraph 2 [44]. The analysis included 35 nucleotide sequences, cited as numbers of NCBI entries in the phylogenetic tree (Figure 7, Table A1). We accepted the outgroup as the *C. adamii* s. str. group [28].

5. Conclusions

The results of the taxonomical, morphological, anatomical, and molecular studies of four Bulgarian populations of *C. pallidus* confirmed the presence of this species in the flora of Bulgaria. In this study, we provided detailed information about the leaf anatomy of *C. pallidus* from Bulgaria for the first time.

Our studies were specifically focused on obtaining ITS data for *C. pallidus* that have not been studied so far. The investigated ITS sequences of the two populations of *C. pallidus* were identical but different when compared to *C. weldenii*. The morphological differences between these two species were also noticeable. Consequently, *C. pallidus* should be treated as an autonomous species. On the other hand, *C. thracicus* shares similar morphological and anatomical features with *C. pallidus*. It is distributed close to the localities of *C. pallidus*. Both species have hybrids with *C. chrysanthus*. Summarizing the previous facts, there is a possibility that *C. thracicus* represents a taxon conspecific with *C. pallidus*.

According to recent data, the distribution area of *C. pallidus* consists of disjunct populations in Serbia–Macedonia and East Bulgaria–European Turkey.

We believe that this information will be helpful for future studies related to the taxonomy and phylogeny of the white-flowering *Crocus* species from the Balkans and neighbouring regions.

Author Contributions: Conceptualization, T.R. and K.S.; formal analysis, T.R., K.S., S.N., and E.A.-K.; Investigation, T.R., K.S., S.N., and E.A.-K.; resources, Y.M., T.R., and K.S.; data curation, T.R, K.S., S.N., and E.A.-K.; visualization: T.R. and K.S.; writing—original draft preparation, T.R., K.S., S.N., E.A.-K., and Y.M.; writing—review and editing: T.R., K.S., S.N., and E.A.-K.; supervision, T.R.; project administration, T.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the Bulgarian Science Fund, grant number KP-06-N31/5, and the Agricultural University of Plovdiv, grant number 17–12.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors express their heartfelt gratitude to Zhivko Barzov for the kindly provided data and photos of *C. pallidus*, from the localities in Varna District. The authors also thank Ress. Asst. Onur Tuncay (Istanbul University, Faculty of Pharmacy, Department of Pharmaceutical Botany) and Serdar Aslan (Düzce University, Faculty of Forestry, Department of Forest Botany) for the high-quality images of the types of *C. thracicus*.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

		20		40 I		60		80		100		120
1 T222361	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	COTOTTATAT	TTGCCAAA	CAAAACGA.C	TTOGGTTOGT	22222222222	222222222	ADDDATADDA	GACGGAGGAA	ACGAAACCCC
MN955426	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	TTGCCAAC	CAAAACGA-C	TTCGGTTCGT	000000000000000000000000000000000000000	CTCGCGGGGGC	GGCGAGGCGA	GACGGAGGAA	ACGAAACCCC
HE664018	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	TTGCCAAC	CAAAACGA-C	TICGGTICGT	00000000000	CTCGCGGGGGC	GGCGAGGCGA	GACGGAGGAA	ACGAAACCCC
MF766260	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CAAAACGA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
HE663973	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CAAAACGA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGG-C	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
HE663991	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CAAAACGA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
HE801149	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CGAAACGA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
LT222362	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CGAAACGA-C	TTCGGTTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
HE664005	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CAAAACGA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
MW775330	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CAAAACGA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
MW775331	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	- TTGCCAAC	CAAAACGA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
LT222368	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGGTTCGT	CGCGCCGCCC	CTCGCGGGGC.	GGCGCGGCGA	GACGGAGGAA	ACGAACCCCC
LT222363	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	GAAAACAT - C	TTCGGTTTGT	CGCGCCGCCC	CTCGCGGGAC	GACGCGGCGA	GACGGAGGAA	ACGAACCCCC
LT222365	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAAAC	TTCGGTTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGCGGCGA	GACGGAGGAA	ACGAACCCCC
OM368594	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGGTTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
HE663982	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGGTTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGCGGCGA	GACGGAGGAA	ACGAAACCCC
HE664009	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGGCTCGT	CCCGCCGCCC	CTCGCGGGGC	GGCGTGGCGA	GACGGAGGAA	ACGAACCCCC
HE663992	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGTTCCGT	CCCGCCGCCC	CTCGCGGCGC	GGCGTGGCGA	GACGGAGGAA	ACGAACCCCC
LM993453	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGTCTCGT	CCCGCCGCCC	CTCGCGGCGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
LM993448	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC.	CAAAACAA-C	TTCGTCTCGT	CCTGCCGCCC	CTCGCGGCGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
LM993466	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTTGTCTCGT	CCCGCCGCCC	CTCGCGGCGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
LM993461	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGACTCGT	CCCGTCGCCC	CTCGTGGCGC	GACGTGGCGA	GCCGGAGGAA	ACGAAACCCC
LM993438	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAAACAA-C	TTCGGTTCGT	CCCGTCGCCC	CTCGTGGCGC	GACGTGGCGA	GCCGGAGGAA	ACGAAACCCC
HE801159	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTAAC	CAAACCAA-C	TCCGGTTCGT	CCCGCCGCCC	CTCGCGGTGC	GGCGTGGCGA	GACGGAGGAA	ACGAAACCCC
HE663983	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTCATAG	ATTTGCTTAC	CAACAA-C	TCCGG - TCGC	CACGCCGCTC	CTCTCGGCAC	GGTGCGGCGA	GACGGAGGAA	ACGAACCCCC
HE801141	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTCATAT	ATTTGCTTAC	CAACGA-C	TCCGGCTCGT	CATGCCGCTC	CTCGCGGTGC	GGCGCGGCGA	GACGGAGGAA	ACGAACCCCC
HE801175	TCGAGACCCG	AACGAACAGA	CGATCGCGAA	CGTGTCATAT	ATTTGCTCAG	TAACAA-C	TTCGGCTCGC	TACGTTGCTC	CTCACGGCGC	GACGCGACGA	GACGGAGGAA	ACGAACCCC-
HE801063	TCGAGACC-G	AACGAACAGA	CGATCGCGAA	CGTGTCATAT	ATTTGCTCAG	TAACAA-C	TTCGGCTCGC	CACGCTGCTC	CTTGCGGCGC	GACGCGGCGA	GACGGAGGAA	ACGAACCCCC
HE801127	TCGAGACCCG	AACGAACGGA	CGATTGCGAA	CGAGTTATAT	ATTTGCTTAC	TAACCA-C	TCCGCTTCGC	CACGTCGCAC	CTTGCGGCGC	GGCGTGACGA	GACGGAGGAA	ACGAAACCCC
HE801123	TCGAGACCCG	AACGAACGGA	CGATTGCGAA	CGAGTTATAT	ATTTGCTTAC	TAACCA-C	TCCGCTTCGC	CACGTCGCAC	CTTGCGGCGC	GGCGTGACGA	GACGGAGGAA	ACGAAACCCC
HE801166	TCGAGACCCG	AACGGA	CGATCGCGAA	CGTGTTATAT	ATTTGCTTAC	- AA CAA- C	TCCGGTTCGC	CCCGCCGCTC	CTCGCGGAAC	CGCGTGACAA	GACGGAGGAA	ACGAAACCCC
HE663958	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTACAT	ATCTGCTCCC	TAACGA-C	TCCGGTTCGT	CCCGCCGCTC	CTCGCGGCGC	GGCGTGACGA	GACGGAGGAA	ACGAAACCCC
HE663964	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTACAT	ATCTGCTCCC	TAATGA-C	TCCGGTTCGT	CCCGCCGCTC	CTCGCGGCGC	GGCGTGACGA	GACGGAGGAA	ACGAAACCCC
HE664017	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTACAT	ATCTGCTCCC	TAACGA-A	TCCGGTTCGT	CCCGTCGCTC	CTTGCGGCAC	GGCGTGACGA	GACGGAGGAA	ACGAAACCCC
HE663969	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTACAT	ATCTGCTCCC	TAACGA-C	TCCGGTTCGT	CCCGCCGCTC	CTCGCGGCGC	GGCGTGACGA	GACGGAGGAA	ACGAAACCCC
HE663995	TCGAGACCCG	AACGAACGGA	CGATCGCGAA	CGTGTTACAT	ATCTGCTCCC	TAACGA-C	TCCG-TTCGT	CCCGCCGCTC	CTCGCGGCGC	GGCGTGACGA	GACGGAGGAA	ACGAAACCCC
100%								V				
Conservation												
		140		160		180		200		220		240
		140 1		160 1		180 I	-	200 1		220 1		240 I
LT222361	GGCGCGGTGG	140 I GCGCCAAGGA	ACACT - TCTT	AGAAGCGCCG	TCGCGGCCTC	180 I CTCC		200 1		220 1		240 1
LT222361 MN955426	GGCGCGGTGG GGCGCGGTGG	GCGCCAAGGA GCGCCAAGGA	ACACT - TCTT ACACT - TTTT	AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC	180 CTCC CTCC	·····	200 I		220 1		240 1
LT222361 MN955426 HE664018	GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG	GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TCTT ACACT - TTTT ACACT - TTTT	AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	180 CTCC CTCC CTCC		200 1		220 1		240 I GG GG
LT222361 MN955426 HE664018 MF766260	GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG	I40 I GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TCTT ACACT - TTTT ACACT - TTTT ACACT - TTTT	AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	180 I CTCC CTCC CTCC		200 1		220 1		240
LT222361 MN955426 HE664018 MF766260 HE663973	GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG	140 J GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TCTT ACACT - TTTT ACACT - TTTT ACACT - TTTT ACACT - TTTT	AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	180 CTCC CTCC CTCC CTCC CTCC		200 T		220 1 		240 I GG GG GG GG GG GG GG
LT222361 MN955426 HE664018 MF766260 HE663973 HE663991	GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG	Ido I GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TCTT ACACT - TTTT ACACT - TTTT ACACT - TTTT ACACT - TTTT ACACT - TTTT	AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	180 CTCC CTCC CTCC CTCC CTCC CTCC		200 1		220 1		240
LT222361 MN955426 HE664018 MF766260 HE663973 HE663991 HE801149	GGCGCGGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG	140 GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT-TCTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT	AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	180 CTCC CTCC CTCC CTCC CTCC CTCC		200 1		220 1		240 1
LT222361 MN955426 HE664018 MF766260 HE663973 HE663991 HE801149 LT222362	GCCGCGGTGG GCCGCGGTGG GCCGCGGTGG GCCGCGGTGG GCCGCGGTGG GCCCGGTGG GCCCCGGTGG	140 I GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TCTT ACACT - TTTT ACACT - TTTT ACACT - TTTT ACACT - TTTT ACACT - TTTT ACACT - TTTT ACACT - TTTT	100 I AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	180 CTCC CTCC CTCC CTCC CTCC CTCC CTCC CTCC CTCC CTCC		200 1		220 1		240 I GG GG GG GG GG GG GG GG GG GG GG GG G
LT222361 MN955426 HE664018 MF766260 HE663973 HE663991 HE801149 LT222362 HE664005	GGCCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCCGGGTGG GGCCGGGTGG	140 CCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TCIT ACACT - TITT ACACT - TITT	160 I AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	TB0 CTCC - -		200 1		220 1 		240 I G G G G G G G G G G G G G G G G G G
LT222361 MN955426 HE664018 MF766260 HE663973 HE603991 HE801149 LT222362 HE664005 MW775330	GCCCCGGTGG GCCCCGGTGG GCCCCGGTGG GCCCCGGTGG GCCCCGGTGG GCCCCGGTGG GCCCCGGTGG GCCCCGTGG GCCCCGTGG	140 ECGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT-TCTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT	100 I AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	TEO 100 CTCC - -		200 T		220 1		240 I G G G G G G G G G G G G G G G G G G
LT222361 MN955426 HE664018 MF766260 HE663973 HE801149 LT222362 HE664005 MW775330 MW775331	GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCCGGTGG GGCCGGTGG GGCCGGTGG	140 CCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TCTT ACACT - TTTT ACACT - TTTT	160 AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	180 CTCC		200		220		240 I G
LT222361 MN955426 HE664018 MF766260 HE663973 HE663991 HE801149 LT222362 HE664005 MW775330 MW775331 LT222362	GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCGCGGTGG GGCCGGGTGG GGCCGGTGG GGCCGGTGG GGCCGGTGG GGCCGGTGG	140 GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TOTT ACACT - TITT ACACT - TITT	AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC	THE CTCC - -			C114GG			240 1 6 6 6 6 6 6 6 6 6 6 6 6 6
LT222361 MN955426 HE664018 MF766260 HE663973 HE801149 LT222362 HE664005 MW775330 MW775331 LT222368 LT222368 LT222367	GCCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG GCCCCGTGG	140 CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA CCCCCAAGGA	ACACT-TCTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT ACACT-TTTT	100 AGAAGCGCCG AGAACGCCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC		CCTACTACCT CCTACTACCT CCTACTACCT		CTTAG G	220 GAGGGAGGGA GAGGGAGGGA GAGGGAGGGA	GGGAGG- GG GGGAGG- GG GGGAGG- GG	
LT222361 MN955426 HE664018 MF766260 HE663973 HE663973 HE861919 LT222362 HE664005 MW775330 MW775331 LT222363 LT222363 LT222363 LT222363	GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC	140 GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA GCGCCAAGGA	ACACT - TOTT ACACT - TITT ACACT - TITT	AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAACGCCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG	TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC		CCTACTACCT CCTACTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC	CITAG- C CITAG- S CITAG- S		GGGAGG - GG GGGAGG - GG GGGAGG - GG	240 4 4 4 4 4 4 4 4 4 4 4 4 4
LT222361 MN955426 HE664018 MF766260 HE663973 HE863991 HE801149 LT222362 HE664005 MW775331 LT222368 LT222368 LT222365 OM385594 HE863969	00000000000000000000000000000000000000	140 CCCCCAAGA CCCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA	ACACT - TOTT ACACT - TITT ACACT - TITT	100 1 AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAACCG AGAACCG A	C6000000000000000000000000000000000000		CTACTACCT CCTACTACCT CCTACTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC	CTTAG G CTTAG G		GGGAGG - GG GGGAGG - GG GGGAGG - GG	240 1 1 3 4 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
LT222361 MN955426 HE664018 MF766260 HE663973 HE663973 HE801194 LT222362 HE664005 MW775330 MW775331 LT222363 LT222363 LT222363 LT222363 HE663982 HE663982	00000000000000000000000000000000000000	140 CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA	ACACT - TGTT ACACT - TTTTT ACACT - TTTTT	100 C C C C C C C C C C C C C	TC8C666CTC TC6C666CTC TC8C666CTC TC8C666CTC TC8C666CTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C68CCTC		CTACTACO CTACTACO CTACTACO CTACTACO		CITAG G CITAG G CITAG G		GGGAGG- GG GGGAGG- GG GGGAGG- GG	240
LT222361 MN955426 HE664018 MF766280 HE663973 HE863973 HE863991 HE801149 LT222362 MW775331 LT222368 LT222365 LT222365 LT222365 M38554 HE663982 HE664009 HE863900	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	140 CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA	ACACT - TO IT ACACT - TITT ACACT - TITT	100 1 AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAAGCGCCG AGAACCGCCG AGAAGCGCCG AGAACCGCCG	C6000000000000000000000000000000000000		CTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CITAG G CITAG G CITAG G G G G G G G G G G G G G G G G G G G	GAAGGAAGGA AGGAAGGA GAGGAAGGA AGGAAGGA	000406-000 000406-000 000406-000 000406-000 000406-000	240 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2
LT222361 MN955426 HE663973 HE663971 HE801149 LT222362 HE664005 MW775330 MW775330 LT222363 LT22363 LT2256 LT2256 LT2256 LT2256 LT256 LT256 LT256 LT256 LT256 LT256	GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC	140 CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA	ACACT - TGTT ACACT - TTTTT ACACT - TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	100 C C C C C C C C C C C C C C C C C C C	TC8C666CCTC TC8C666CCTC TC8C666CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C66CCTC TC8C68CCTC TC8C68CCTC TC8C68CCTC TC8C68CCTC TC8C68CCTC TC8C68CCTC TC8C68CCTC TC8C68CCTC		CTACTACT CTACTACT CTACTACT CTACTACT CTACTACT CTACTACT CTACTACT	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTIACCTAC	CITAG - G CITAG - G AGGGCCGGAG	GAGGAACGA GAGGAACGA GAGGAACGA GAGGAACGA GAAGGAACGA	GGGAGG- GG GGGAGG- GG GGGAGG- GG GGGAGG- GG	240
LT222361 MN955426 HE664018 MF766260 HE6633073 HE6633073 HE66340149 LT222362 LT222362 LT222368 LT222368 LT222368 LT222365 LT22355 LT23555 LT2355 LT2355 LT2355 LT2355 LT23555 LT23555 LT23555 LT23555 LT23555 LT23555 LT23555 LT23555 LT235555 LT235555 LT2355555555 LT23555555555555555555555555555555555555	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	140 CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA	ACACT - TO TT ACACT - TITT ACACT - TITT	100 1 AGAAGCCCCC AGAACCCCCCC AGAACCCCCC AGAACCCCCCC AGAACCCCCCC AGAACCCCCCC AGAACCCCCCC AGAACCCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCCC AGAACCCCCCC AGAACCCCCCC AGAACCCCCCC AGAACCCCCCCC AGAACCCCCC	C6000000000000000000000000000000000000		CTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CITAG G CITAG G CITAG G AGGCCGGGAG		666466 66 666466 66 666466 66 666466 66 666466 66	240 1 2 2 2 2 2 2 2 2 2 2 2 2 2
LT222361 MIN95426 HE664018 MI776280 HE6633973 HE8633973 HE864005 MIV775331 LT222365 UT222363 UT222363 UT222363 UT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222364 LT222365 LT22365 LT2365 LT25	00000000000000000000000000000000000000	140 CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA	ACACT - TGTT ACACT - TTTTT ACACT - TTTTT	100 C C C C C C C C C C C C C	TCSCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC		CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT		CITAG C CITAG C AGGGOGGGAG	GAGGAGGA GAGGAGGA GAGGAGGA AGGAGGA GAAGGA GAAGGA GAAGGA CGA C		240
LT222361 MIN955426 HE664018 HE766200 HE863071 HE801149 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222364 LT22364 LT222364 LT22564 LT22564 LT22564 LT22564 LT22564 LT22564 LT22564 LT22564 LT2564 L	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	140 CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA	ACACT - TO TT ACACT - TITT ACACT - TITTT ACACT - TITTT ACACT - TITTT ACACT - TITTT ACACT - TITTT	100 10 10 10 10 10 10 10 10 10	C6000000000000000000000000000000000000		CTACTACT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCCTACT CCCTACCT CCCTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CITAG G CITAG G CITAG G AGGCCGGGAG		666466 - 66 666466 - 66	240 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2
LT222361 MI965426 HE5663973 HE6663973 HE6633973 HE6633973 HE76264005 MW775331 LT222365 LT222365 C03388594 HE6664009 HE6663982 LT222385 LM993465 LM993446 LM993448	GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC	140 CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA CCCCCAAGA	ACACT - TGTT ACACT - TTTTT ACACT - TTTTT	100 C C C C C C C C C C C C C	TCSCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC TCGCGGCCTC				CITAG C CITAG C AGGGCGGGAG			240
LT222361 MIN955426 HE664018 HE663091 HE683091 HE801109 HE801109 HE801109 HE801109 HE801109 HE803102 HE664009 HE663982 HE864009 HE663982 LM993448 LM993453 LM993453 LM993453	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	140 CCCCCAAGA CCCCCAAGA CCCCAAGA	ACACT - TGTT ACACT - TITT ACACT - TITT	100 1 AGAAGCCCCC AGAAGCCCCC AGAAGCCCCC AGAAGCCCCC AGAAGCCCCC AGAAGCCCCC AGAAGCCCCC AGAAGCCCCC AGAACCCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCC	TC60060000 TC60000000 TC60000000 TC60000000 TC60000000 TC60000000 TC60000000 TC60000000 TC60000000 TC60000000 TC60000000 TC600000000 TC6000000000 TC60000000000			ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CTTAG S CTTAG S AGGGCGGGGAG			240 1 2 2 2 2 2 2 2 2 2 2 2 2 2
LT222361 MIN955426 HE664018 HE6633991 HE6633991 LT222362 HE664005 MIV775330 MIV775331 LT222368 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT22365 LT	GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC		ACACT - TGTT ACACT - TTTTT ACACT - TTTTTT ACACT - TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	100 C C C C C C C C C C C C C	TCSCGGCCTC TCGCGGCTC				CITAG C CITAG C AGGGCGGGAG	200 CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGA CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGA CAGGA CAGGAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAG CAG		240
LT222361 MIN955426 HE664018 HE664018 HE663091 HE803119 HE801149 HE801149 HE801149 HE72362 HE664009 HE66308594 HE664009 HE663082 LM993448 LM993463 L	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	140 CCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITT	100 C C C C C C C C C C C C C C C C C C C	TC60060000 TC6000000 TC6000000 TC6000000 TC6000000 TC60000000 TC6000000 TC6000000 TC6000000 TC6000000 TC60000000000			ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CTTAG S CTTAG S AGGGCGGGGAG	220 1 2 2 2 2 2 2 2 2 2 2 2 2 2		240
LT222361 MIN955426 HE664018 HE6663091 HE663091 LT222362 HE664005 UN7775310 MIX775310 HE763002 LT222368 LT222365 LT222365 LT222365 LT222365 LT222365 LT222365 HE664009 HE663092 HE664093 HE664093 HE664093 HE664093 HE66309 HE663	GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCCGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC GCCCGGTGC		ACACT - TGTT ACACT - TTTT ACACT - TTTTT ACACT - TTTTT ACACT - TTTTT ACACT - TTTTT	100 C C C C C C C C C C C C C	TCSCGGCCTC TCGCGGCTC		CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CITAG G CITAG G AGGGCCGGAG	200 CAGGAGGA CAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGAGGA CAGGA CAGGAGGA CAGCA CAG CAG		240
LT222361 MIN95426 HE664018 HE766200 HE663391 HE663391 HE863400 LT222362 HE664009 HE72330 MW775331 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 HE664009 HE663983 HE663983 HE663983 HE6801165 HE801116 HE8011175 HE8011175	GCCCCGTGC GCCCCGTGC	140 CCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITTT ACACT - TITTT ACACT - TITTT ACACT - TITTT ACACT - TITTT	100 C C C C C C C C C C C C C C C C C C C	TC60060000 TC60000000 TC60000000 TC60000000 TC60000000000			ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CTTAG S CTTAG S AGGGCGGGGAG	220 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2		240
LT222361 MIN955426 HE664018 HE6633991 HE6633991 HE6633991 LT222362 HE763391 LT222362 LT222362 LT222362 LT222365 LT222365 LT222365 LT222365 LM933458 LM933461 LM93348 HE664009 HE663392 HE664099 HE663392 HE664099 HE663392 HE663392 HE664091 HE664099 HE663392 HE664091 HE664000 HE664000 HE664000 HE664000 HE664000 HE664000 HE664000 HE664000 HE664000 HE664000 HE664000 HE664000	GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCCGTGC GCCCCGTGC		ACACT - TUTT ACACT - TITT ACACT - TITT		TCSCGGCCTC TCGCGGCTC		CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CITAG G CITAG G AGGGCCGGGAG	200 CAGGAAGGA GAGGAAGGA GAGGAAGGA - GAGGAGGA - GAGGAGGA - GAGGAGGA - TAGGA - TAG		240
LT222361 MIN955426 HE664018 HE766200 HE663391 HE663391 HE763391 HE78310 HE775331 LT222362 LT222362 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 HE764009 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE763393 HE773 HE763393 HE773 HE7734 H	GCCCCGTGG GCCCCGTGG	140 CCCCCAAGA CCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITT	100 C C C C C C C C C C C C C C C C C C C	TCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC			ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CTTAG G CTTAG G AGGGCCGGGAG	220 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2	CGGACGACGACGACGACGACGACGACGACGACGACGACGA	240
LT222361 MIN955426 HE664018 HE6633991 HE6633991 HE6633991 LT222362 HE763309 HE763309 HE763309 LT2222362 LT222368 LT222368 LT222368 LT222368 LT222368 LM933469 LM933469 LM93348 HE664009 HE663392 HE722369 HE72330 HE72330 HE72330 HE733300 HE73330 HE73330 HE73330 HE733300 HE733300 HE733300 HE733300 HE733300 HE733300 HE733300 HE733300 HE733300 HE733300 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE733000 HE7330000 HE73300000000000000000000000000000000000	Control of the second s	140 CCCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITT		TCSCGGCCTC TCGCGGCCTC		CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CITAG G CITAG G AGGGCCGGGAG	200 CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGA CAGGA CAGGA CAGGAGGA CAG CAG		240
LT222361 MIN956428 HE66407 HE663971 HE663971 HE663971 HE66407 HE76397 HE76397 HE76407 HE775330 MIV775330 MIV775330 HE764009 HE664009 HE663982 HE722368 LM93348 LM93348 LM93348 LM93348 LM93348 LM93348 LM93348 HE663983 HE661175 HE601163 HE801163 HE801163 HE801165	Control of the second state of the second stat	140 CCCCCAAGA CCCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITT	100 C C C C C C C C C C C C C C C C C C C	TCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC			ACCIACCIAC ACCIACCIAC ACCIACCIAC ACCIACCI	CTTAG G CTTAG G AGGGCCGGGAG	220 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2	CGAACGA CGAACGA CGAACGA CGAACGA CGACGA CGACG	240
LT222361 MIN955426 HE664018 HE6633991 HE6633991 HE6633991 LT222362 HE763391 LT222362 LT222362 LT222362 LT222363 LT222365 LT222365 LT222365 LT222365 LT222365 HE664009 HE663392 HE663392 HE663392 HE663392 HE663392 HE663393 HE801159 HE663394 HE801159 HE801159 HE801151 HE801161 HE801161 HE801123	00000000000000000000000000000000000000	140 CCCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA CCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITT		TCSCGGCCTC TCGCGGCCTC			ACCIACCIAC ACCIACCIAC ACCIACCIAC ACCIACCI	CITAG G CITAG G AGGGCCGGGAG	220 CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGAGGA CAGGA CAGGA CAGGA CAGGA CAGGAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAGGA CAG CAG		240
LT222361 MIN956426 HE660391 HE663971 HE663971 HE663971 HE66400 HE75330 MIV775330 MIV775330 HE722368 LT222366 OM368594 HE664009 HE663382 HE663382 HE663383 HE663383 HE6801165 HE801163 HE801117 HE80112 HE801117 HE80112 HE80112 HE801117 HE80112 HE8012 HE80112	Control of the second state of the second stat	140 CCCCCAAGA CCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITT	100 C C C C C C C C C C C C C C C C C C C	Construction of the constr			ACCIACCIAC ACCIACCIAC ACCIACCIAC ACCIACCI	CTTAG G CTTAG G AGGGCCGGGAG	220 1 1 1 1 1 1 1 1 1 1 1 1 1	CGAACGA CGAACGA CGAACGA CGAACGA CG	240
LT222361 MIN955426 HE664018 HE6633991 HE6633991 LT222362 HE664005 UN7775330 MIX775330 MIX775330 LT2222368 LT2222368 LT2222368 LM933469 LE664009 HE663092 LM933469 LM93348 HE801159 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE663092 HE801159 HE801159 HE801151 HE801161 HE801161 HE801121 HE80121 HE80	Control of the second s		ACACT - TUTT ACACT - TITT ACACT - TITT		TCSCGGCCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGGCTC TCGCGCCTC			ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CTTAG C CTTAG C AGGGCCGGGAG	220 T C C C C C C C C C C C C C		240
LT222361 MIN956426 HE660307 HE6633991 HE6633991 HE6633991 MIV775330 MIV775330 LT222382 LT222382 LT222382 LT222382 LT222382 LT222382 LT222382 LT222382 LT222382 LM933448 LM933453 LM934553 LM9345553 LM9345552 LM93455552 LM9345552 LM9345552 LM9345552 LM9345552 LM9345552 LM9345552 LM93455552 LM9355552 LM9355552 LM9355552 LM9355552 LM9355552 LM935555555552 LM93555555555555555555555555555555555555	2000 Control C	140 CCCCCAAGA CCCCAAGA	ACACT - TUTT ACACT - TITT ACACT - TITT	100 CARACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACAACCOCCC CACACCCCCC CACCCCCC CACACCCCCC CACACCCCCC CACCCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCC CACCCCCCC CACCCCCC CACCCCCC CACCCCCCCC	Construction of the second sec			ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CITAG G CITAG G AGGGCCGGGAS	220 1 2 2 2 2 2 2 2 2 2 2 2 2 2	CGAACGAC C CGAACGAC C CGAACGAC C CGAACGAC C CGAACGAC C C CGAACGAC C C C	240 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2
LT222361 MIN955426 HE6604018 HE6633931 HE6633931 HE663391 LT222362 HE664005 UN7775330 MIX775330 LT222368 LT222368 LT222368 LM93463 LM93468 LM933468 LM93348 HE663092 HE663092 HE663092 HE663098 HE601123 HE801169 HE663098 HE801169 HE663098 HE801169 HE801173 HE801169 HE801173 HE801169 HE801173 HE801173 HE801174 HE801175 HE801169 HE80175 HE801169 HE80175 HE801169 HE80175 HE801169 HE80175 HE801169 HE80175 HE801173 HE801175 HE801169 HE80175 HE801169 HE80175	Control of the contro	140 1 1 1 1 1 1 1 1 1 1 1 1 1	ACACT - TUTT ACACT - TITT ACACT - TITT	100 1 AGAACCCCCCC AGAACCCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCCCC AGAACCCC	Construction of the second sec		CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCTACCT CCCCTACCT CCCCTACCT CCCCTACCT CCCCTCCT CCCCATTCC CCCCATTCC CCCCATTCC CCCCATTCC	ACCTACCTAC ACCTACCTAC ACCTACCTAC ACCTACCT	CTIAG C CTIAG C AGGCCCGGAG	CAGGAAGGA GAAGGAGGAGGA GAAGGAGGAGGA AGGAGG		240 1 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2

Figure A1. Cont.

		280		280		300		320		340		360	
1 7000264	000000000	CONTROLATAT	ACATATOOTA	TTTTCTACCA	CTCTCCCCAA	CCCATATOTA	CONTRACCO	TOCATCAACA	ACCTACCCAA	ATCCCATACT	TOCTOTOAAT	TOCACAATCO	
MN955426	GGGAACGCGG	GCGTGGATAT	AGATATOGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE664018	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
MF766260	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE663973	GGGAACGCGG	GCGTGGACAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE663991	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE801149	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
LT222362	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE664005	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
MW775330	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
MW775331	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
LT222368	GGGAACGCGG	GCGTGGATAT	GGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
LT222363	GGGAATGCGG	GCGTGGATAT	GGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
LT222365	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
OM368594	GGGGACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE663982	GGGGACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE664009	GGGAACGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE663992	IGGA CGCGG	GCGTGGATAT	AGATAICGIA	TTTTGTACGA	CICICGGCAA	CGGATAICIA	GGCTCTCGCA	ICGAIGAAGA	ACGIAGCGAA	AIGCGATACT	IGGIGIGAAI	IGCAGAAICC	
LM993453	GGGAACGCGG	GCGIGGATAT	AGATATCGTA	TTTTGTACGA	CICICGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
LM993448	GGGAACGCGG	GCGIGGATAT	AGATATCGTA	TTTTGTACGA	CICICGGCAA	CGGATAICIA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
LM993466	GGGAALGCGG	GCGTGGATAT	AGATATCGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TOGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
LM993461	GGGAACGCGG	GCGTGGATAT	AGATATOGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TOGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TOCAGAATCC	
LW993436	GGGAACGCGG	GCGTGGATAT	AGATATOGTA	TTTTGTACGA	CTCTCGGCAA	COGATATOTA	GGCTCTCGCA	TOCATCAACA	ACGTAGCGAA	ATGOGATACT	TOOTOTOAAT	TOCAGAATCC	
HE601109	CCCAACGCGG	CCCTCCACAT	AGATATCGTA	TTTTCTACCA	CTCTCCCCCAA	CCCATATCTA	COCTOTOGOA	TCCATCAACA	ACGTAGUGAA	ATGCGATACT	TCCTCTCAAT	TECAGAATEC	
HE003963	CCCAACCCCC	COCTOCACAT	ACATATCOTA	TTTTCTACCA	CTCTCCCCCAA	CCCATATCTA	COCTOTOGOA	TOCATCAACA	ACCTACCCAA	ATCCCATACT	TCCTCTCAAT	TCCACAATCC	
HE801175	GGGAACGCGG	CCGTCGACAT	AGATATOGTA	TTTTGTACGA	CTCTCGGCAA	CGGATATOTA	GOCTOTOGOA	TCGATGAAGA	ACGTAGOGAA	ATGCGATACT	TGGTGTGTGAAT	TGCAGAATCC	
HE901063	CCC ACCCCC	CCGTGAACAT	AGATATCOTA	TTTTGTACGA	CTCTCGGCAA	CCCATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TCCTCTCAAT	TGTAGAATCC	
HE801127	GGGAACGCGG	GCGTGGATAT	GATATOGTA	TTETETACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE801123	GGGAACGCGG	GCGTGGATAT	GATATOGTA	TTETETACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE801166	GGGAACGCGG	GCGTGGATAT	GATATOGTA	TTOTOTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE663958	GCGAACGCGG	GCGTGGATAT	AGATATCGTA	TTCTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE663964	GCGAACGCGG	GCGTGGATAT	AGATATCGTA	TTCTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE664017	GCGAACGCGG	GCGTGGATAT	AGATATCGTA	TTCTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
HE663969	GCGAACGCGG	GCGTGGATAT	AGATATCGTA	TTCTGTACGA	CTCTCGGCAA	CGGATATCTA	GGCTCTCGCA	TCGATGAAGA	ACGTAGCGAA	ATGCGATACT	TGGTGTGAAT	TGCAGAATCC	
115000005				TTOTOTIONI									

HE664017 HE663969	GCGAACGCGG GCGAACGCGG	GCGTGGATAT GCGTGGATAT	AGATATCGTA AGATATCGTA	TTCTGTACGA	CTCTCGGCAA CTCTCGGCAA	CGGATATCTA CGGATATCTA	GGCTCTCGCA GGCTCTCGCA	TCGATGAAGA TCGATGAAGA	ACGTAGCGAA ACGTAGCGAA	ATGCGATACT ATGCGATACT	TGGTGTGAAT TGGTGTGAAT	TGCAGAATCC TGCAGAATCC
Conservation	o on no o o o o o	ood toomini	NOATATOOTA		01010000///	ocontratorn	0001010000	TOGNTONNON	NeerNeeenN	Areconther	Toororonan	100/10/07/100
67%.		380	1	400		420		440		460		48
		ĩ		Ť		ĩ		ĩ		ĩ		ĩ
L1222361 MN955426	CGTGAACCAT	CGAGICITIG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
HE664018	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
MF766260	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
HE663991	CGTGAACCAT	CGAGICITIG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
HE801149	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCCAC	CCGTGTGCGG	ATGCGGAGAT
LT222362	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
HE664005	CGTGAACCAT	CGAGICITIG	AACGCAAGTT	GCGCCCGAGG	CCATCOGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCCAC	COTOTOCOG	ATGCGGAGAT
MW775331	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCCAC	CCGTGTGCGG	ATGCGGAGAT
LT222368	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
L1222363	CGIGAACCAT	CGAGICITIG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GIGIGCGG	ATGCGGAGAT
OM368594	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
HE663982	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	AGGGAACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
HE664009	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
LM993453	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCCAAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
LM993448	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
LM993466	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
LM993461	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGTC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCC	GTGTGCGG	ATGCGGAGAT
HE801159	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACACCCC	GCGTGTGCGG	ATGCGGAGAT
HE663983	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE801141 HE801175	CGTGAACCAT	CGAGICITIG	AACGCAAGTT	GCGCCCGAGG	CCATCOGGIC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	COGTOGOTOC	CCACAGCC	GIGCGG	ATGCGGAGAT
HE801063	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGTC	GTGCGG	ATGCGGAGAT
HE801127	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE801123	CGTGAACCAT	CGAGICITIG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE663958	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE663964	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE664017	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE003909	COTGAACCAT	CGAGICITIG	AAGGGAAGTT	GUGUUUGAGG	CCATCCGGCC	GAGGGGGGGGGGG	CIGCUIGGGC	GICACGCCIC	CCGTCGCTCC	CCACAGCC	GIGCGG	ALGOGGAGAL
HE663995	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GICACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE663995 conservation	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE663995 scone Conservation	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC	CTGCCTGGGC	GTCACGCCTC	CCGTCGCTCC	CCACAGCC	GTGCGG	ATGCGGAGAT
HE663995 toon Conservation	CGTGAACCAT	CGAGTCTTTG	AACGCAAGTT	GCGCCCGAGG	CCATCCGGCC	GAGGGCACGC 540	CTGCCTGGGC	GTCACGCCTC 500	CCGTCGCTCC	CCACAGCC		ATGCGGAGAT
LT222361 MN955426	TGGCCCACCG	CGAGTCTTTG 500 TGCCTCGTGC TGCCTCGTGC	AACGCAAGTT GCGGCGGGGTC	GCGCCCGAGG 520 GAAGTGCGGG GAAGTGCGGG	CCATCCGGCC	GAGGGCACGC 54(GCCGGGCGCG GCCGGGCGCGCG	GCGAATGGTG	GACGAATACT		CCACAGCC	CCGTGCCCCG	ATGCGGAGAT
HE663995 conservation LT222361 MN955426 HE664018	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG 500 I TGCCTCGIGC TGCCTCGCGC TGCCTCGCGC	AACGCAAGTT GCGGCGGGGTC GCGGCGGGTC GCGGCGGGTC	GCGCCCGAGG 520 I GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG	CCATCCGGCC CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	GAGGGCACGC 540 I GCCGGGCGCGC GCCGGGCGCGC GCCGGGCGCGC	CTGCCTGGGC GCGAATGGTG GCGAATGGTG GCGAATGGTG	GTCACGCCTC 500 GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTTG	CCACAGCC 500 I TGCCTACTCT TGCCTACTCT TGCCTACTCT	CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT
HE663995 conservation m LT222361 MN955426 HE664018 MF766260	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG 500 I TGCCTCGTGC TGCCTCGCGC TGCCTCGCGC TGCCCCGTGC	AACGCAAGTT GCGGCGGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC	GCGCCCGAGG 520 I GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG	CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	GAGGGCACGC 54(1 GCCGGGCGCG GCCGGGCGCG GCCGGGCGCG GCCGGGCGCG	CTGCCTGGGC GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG	GICACGCCTC 500 I GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG	CCACAGCC 580 I TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT	CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT
LT222361 MN955426 HE664018 MF766260 HE663991	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTITG 500 I TGCCTCGTGC TGCCTCGGC TGCCTCGGC TGCCCCGTGC TGCCCCGTGC	AACGCAAGTT GCGGCGGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC	GCGCCCGAGG 520 I GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG	CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	GAGGGCACGC 5% 1 GCCGGGCCGG GCCGGGCGCG GCCGGGCGCG GCCGGCGCGCG GCCGGCGCGCG GCCGGCGCGCG	CTGCCTGGGC GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG	GICACGCCTC I GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG	CCACAGCC 5500 I TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT	CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT
HE663995 Conservation MN955426 HE664018 MF766260 HE663973 HE663973 HE6639149	TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAAGTT GCGGCGGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC	GCGCCCGAGG 520 1 GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG	CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	6A66620266 140 15020202020 15020202020 15020202020 15020202020 15020202020 15020202020 15020202020	CTGCCTGGGC GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG	GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG	CCACAGCC 500 TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT	CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT
HE6639955 Conservation MN955426 HE664018 MF766260 HE663973 HE663991 HE801149 LT222362	TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTITG 500 1 TGCCTCGTGC TGCCTCGCGC TGCCCCGTGC TGCCCCGTGC TGCCCCGTGC TGCCCCGTGC TGCCCCGTGC	AACGCAAGTT GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC	GCGCCCGAGG 520 J GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG	CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	GAGGGCACGC GAGGGCACGC I GCCGGCGCGC GCCGGCGCGC GCCGGCGCGC GCCGGCGCGC GCCGGCCGCG GCCGGCCGCG GCCGGCCGCG	CTGCCTGGGC CCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG	GICACGCCTC GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG	CCACAGCC 500 1 TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT	CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT 6 GCCC CAACA GCCC CAACA GCCC CAACA GCCC CAACA GCCC CAACA GCCC CAACA GCCC CAACA GCCC CAACA
HE663995 Conservation LT222361 MN955426 HE664018 MF766260 HE663973 HE663973 HE663991 HE801149 LT222362 HE664005	TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAAGTT GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC	CCGCCCGAGG 5000 I GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG	CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	GAGGGCACGC 544 GCCGGGCGCG GCCGGCGCG GCCGGCCGC GCCGGCCGC GCCGGCCGC GCCGGCCGC GCCGGCCCCG GCCGGCCCCG GCCGGCCCCG	CTGCCTGGGC GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG	GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG	5500 TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT TGCCTACTCT	CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT
HE663995 Conservation LT222361 MN955426 HE664018 MF766260 HE663973 HE6639973 HE801149 LT222362 HE664005 MW775331	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAAGTT GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC GCGGCGGGTC	GCGCCCGAGG 520 I GAAGTCCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG GAAGTGCGGG	CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	*** *** *** **************************	CTGCCTGGGC GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG GCGAATGGTG	GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG	Solo Solo	CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT
HE663995 Conservation 0 LT222361 MN955426 HE664018 MF766260 HE663973 HE663973 HE663973 HE663973 HE663973 HE661973 HE664005 MW775331 LT222368	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTIG	AACGCAAGTT GCGGCGGGGTC GCGCCGGGTC GCGCCGGGTC GCGCCGGGTC GCGCCGGGTC GCGCCGGGTC GCGCCGGGTC GCGCCGGGTC	202 202 202 202 202 202 202 202	CCGTCGTCGGCC CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG CCGTCGTCGG	** ** 0502692020 8050260205 8050260205 8050260205 8050260205 8050260205 8050260205 8050260205 8050260205 8050260205 8050260205	CTGCCTGGGC CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG	GICACGCCIC GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTG TCGTTGTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG TCGTTGTTTG	CCACAGCC CONTRACTOR I GCCTACTOR I GCCT		ATGCGGAGAT 00 1 0 0 0 0 0 0 0 0 0 0 0 0 0
HE663995 Conservation LT222361 MN955426 HE664018 MF766260 HE663971 HE801149 LT222362 MV775330 MV775331 LT222368 LT222368 LT222368	TGGCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAAGTT GCGCGCGGTC GCGCCGGTC GCGCCGGTC GCGCCGGTC GCGCCGGTC GCGCCGGTC GCGCCGGTC GCGCCGGTC GCGCCGGTC GCGCCGGTC	4445 4455 4455 44577 44577 44577 44577 44577 44577 445777 4457777 4457777777777	CCGTCCGCCC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCGC CCGTCGTCCG CCGTCGTCCG CCGTCGTCCG	202042929429 20204292020 202092020 20209200 2020920 20209 2020920 2020900 20209 2020920 202000 20209 20209	CTECCTEGEC CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG	GICACGCCIC GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI	CCGTCGCTCC CCGTCGTTG CCGTTGTTG CCGTTGTTG CCGTCGTTGTTG CCGTCGTTGTTG CCGTCGTTGTTG CCGTCGTTGTTG CCGTCGTTGTTG CCGTCGTTG CCGTCGTTG CCGTCGTTG CCGTCGTTG CCGTCGTTG	CCACAGCC	CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6 CC6160CCC6	ATGCGGAGAT () GCCC CAACA GCCC CAACA
HE663995 Conservation MN955426 HE664018 MF765260 HE663973 HE663973 HE663973 HE663973 HE663973 HE663973 HE801149 LT222362 HE64005 MW775330 MW775331 LT222363 LT222363 LT222363 LT222363	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG		AACGCAGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC CCGCCCGGTC	202462/2020 202029126A6 20202126A6 20202126A6 20202126A6 20202126A6 2020226A6 2020226A6 2020226A6 2020226A6 2020000000000	CCGTCCGCCC CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG CCGTCCTCCG	2222429298292 ** 2322992222 232299222 232299222 232299222 232299222 232299222 232299222 232299222 232299222 2322922 23229222 23229222 23229222 23229222 23229222 23229222 23229222 23229222 23229222 23229222 2322922 2322922 2322922 2322922 2322922 2322922 2322922 2322922 2322922 2322922 2322922 2322922 2322922 2322922 232292	CTACCTORE CARACON CARA	GICACGCCIC GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC CCGTCGTTCC CCGTCGTTGTTC CCGTCGTTGTTC CCGTCGTTGTTC CCGTCGTCGTC CCGTCGTCGTC CCGTCGTTC CCGTCGTTCTTC CCGTCGTTCTTC CCGTCGTTCGT	CCACAGCC TGCCTACTCT GCCTACTCT		ATGCGGAGAT 6 6 6 6 6 6 6 6 6 6 6 6 6
HE663995 Conservation NN955426 HE664018 MF766260 HE663973 HE663973 HE663991 HE801149 LT222363 MW775331 LT222363 LT222363 LT222363 LT222363 LT222363	TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAAGTT CCCCCCCCCCCCCCCCCCCCCCCCCCCCC	000 000 000 000 000 000 000 000	CCATCCGCCC CCATCCGCCC CCATCCGCCC CCATCATCCC CCATCATCCC CCATCATCCC CCATCATCCC CCATCATCCC CCATCATCCC CCATCATCCC CCATCATCCC CCATCCTCCC CCATCCCCCC CCATCCCCCC CCATCCCCCCC CCATCCCCCCCC	44 44 45 45 45 45 45 45 45 45	CTECCTOGEC CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE CCEANTEGTE	GICACGCCIC GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGEC	CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCCC CGTTCCCCCC CGTTCCCCCC	ATGCCGAGAT
HE663995 Conservation MN955426 HE664018 MF766200 MM776330 MW775310 MW775330 MW775330 MW775330 MW775330 LT222368 LT22368 LT2368 LT22368 LT22368 LT22368 LT22368 LT22368 LT	TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAAGTT GCGCCGCGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC GCGCCGGCC		CCATCCGCCC CCATCCGCCC CCATCGTCGC CCATCGTCC CCATCGTCC CCATCGTCC CCATCGTCC CCATCCCC CCATCCC CCATCCC CCATCCC CCATCCC CCATCCCCCC CCATCCCC CCATCCC CCATCCCCCCC CCATCCCCCCCC	202042992429 20204292020 20202920202 2020292020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 202092020 20209200 20200000 202090000 202090000000000	CCCATCGCC CCCATCGCC CCCACTC CCCAC	GICACGCCIC C CACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI GACGAATACI	CCGTCGCTCC CCGTCGTTG TCGTTGTTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGTTG TCGTTGT TCGTTGT TCGTTGTTG TCGTTGTTG TCGTTGT	CCACAGECC SUP TGCCTACTOT	CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000 CC61600000	ATGCGGAGAT () GCCC CAACA GCCC CAACA GC
HE663995 Conservation MN955426 HE664018 MF765200 HE663973 HE80391 HE801149 LT222362 HE664005 MV775331 LT222365 UT222365 UT222365 UT222365 UT222365 UT222365 UT222365 UT222365 UT222365 UT222365 HE664009 HE6630992 HE6630992 HE663092	CGTGAACCAT TGGCCCACCG		AACGCAGTC CCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	20046272620 200201756400 20020175640 200201756400 200201756400 200201756400 200201756400 20020000000000000000000000000000000	CCGTCCGCCC CCGTCCGCC CCGTCCGCC CCGTCCGCC CCGTCCGCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCC CCGTCGCCCCCCCC	2222429829842 24 25 25 25 25 25 25 25 25 25 25	CTACCTORES CALL	GICACGGCTC GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC CCGTCGCTCC CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTCGTCG CCGTCGTCGTCG CCGTCGTCGTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG	CCACAGCC See CCACAGCC CCCACCC CCCCACCC CCCACCC CCCACCC CCCACCC CCCACCC CCCACCC	CCGTGCCCCG CCGTGCCCCG	ATGCGGAGAT
HE663995 Conservation LT222361 MN955426 HE664018 HE664018 HE664019 HE801149 LT222362 HE664009 HE663991 LT222363 LT222362 LT222363 LT22363 LT222363 LT2363 LT22363 LT22363 LT22363 LT22363 LT363 LT2363 LT	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAGAGTT CCCCCCCCCCCCCCCCCCCCCCCCCCCC	000 000 000 000 000 000 000 000	CCATCCGGCC CCATCCGGCC CCATCATCGG CCATCATCATCATCG CCATCATCATCGG CCATCATCATCGG CCATCATCATCATCATCGG CCATCATCATCATCATCATCATCATCATCATCATCATCAT	** ** ** ** ** ** ** ** ** ** ** ** **	CTECCTOGGC CCAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG CCGAATGGTG	GICACGCCIC GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGECC TGCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT GCCTACTOT	CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCC CGTTCCCCCC CGTTCCCCCC CGTTCCCCCC CGTTCCCCCC	ATGCCGAGAT
HE663995 Conservation HE7222361 MN955426 HE664018 HE663973 HE663973 HE663997 HE663974 HE663974 HE663975 HE663975 HE663992 HE664009 HE7222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 LT222363 HE663995 HE663995 HE663995 LT222363 HE663995 HE663995 HE663995 HE663995 HE663995 HE663995 HE663995 HE663995 HE663995 HE663995 HE66395 HE66	CGTGAACCAT TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG TGGCCCACCG	CGAGTCTTTG	AACGCAAGTT GCGCCGCGCTC GCGCCGCGCC GCGCCGCGCC	\$	CCATCCGGCC CCATCCGCGC CCATCGTCGC	202042992429 2020429202 20202920202 20202920202 2020292020 2020200 202000 202000 202000	CCGAATGGTG GCGAATGGTG	GICACGCCIC	CCGTCGCTCC CCGTCGTTG CCGTTGTTTG	CCACAGECC Signific Content of the second		ATGCGGAGAT
HE663995 Conservation HE664018 HE664018 HE664018 HE664018 HE664018 HE663991 HE663991 HE801149 LT222362 HE664005 MW775330 MW775330 MW775330 HW77530 HW775530 HW775530 HW775530 HW775	CGTGAACCAT TGGCCCACCG		AACGCAGT CCGCCCGGTC	200462726252 200727564 200727564 200727564 200727564 200727564 200727564 200727564 200727564 200727565 200725555 200725555 200725555 200725555 2007255555 2007255555 20075555555555	CCGTCCCCCC CCGTCCCCCC CCGTCCCCCC CCGTCCCCCCC CCGTCCCCCCC CCGTCCCCCCCC	*** *** *** **************************	CCGAATGGTG CCGAATGGTG	GICACGGCTC M GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT GACGAATACT	CCGTCGCTCC CCGTCGTTCC CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTCGT	CCACAGCC 500 CCCACAGCC CCCACACCC GCCTACTCT GCC	CCGTGCCCCG CCGTGCCCCG	
HE663995 Conservation LT222361 MN955426 HE664018 MF766280 HE663991 HE80149 LT222362 HE664099 LT222368 LT22368 LT2368 LT22368 LT22368 LT22368 LT22368 LT22368 LT22368 LT2368 L	CGTGAACCAT TGGCCCACCG TGGCCACCG TGGCCCACCG TGGCCACCG TGG		AACGCAGAGTC CCCCCCCCCCCCCCCCCCCCCCCCCCCC	000 000 000 000 000 000 000 000	CCATCCGGCC CCATCCGGCC CCGTCGTCGG CCGTCGTCG CCGTCG CCGTCGTCG CCG C	** ** ** ** ** ** ** ** ** ** ** ** **	CTECCTOGGC CCAATGGTG CCGAATGGTG	GICACGCCIC GACGAATACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGECC T GCCTACTOT		ATGCCGAGAT
HE663995 Conservation L1222361 MM955426 HE664018 HE663991 HE603918 HE603991 HE603992 HE664005 MW775331 L1222363 L1222363 L12222863 L12222863 L12222863 L12222863 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L12363 L1236	CGTGAACCAT TGGCCCACCG TGGCCACCG TGG		AACGCAGATC CGCCCCCCCCCCCCCCCCCCCCCCCCCCC	200 200 200 200 200 200 200 200	CCATCCGCCC CCATCCGCCC CCATCCGCCCC CCATCCGCCCCCCCCCC	232A2362382A2 4 % 5 232382323 5 2323823 5 23238232 5 2323823 5 232382 5 23282 5 23282 5 23282 5 25	CCGAATGGTG GCGAATGGTG	CACGAATACT CACGAATACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGEC		ATGCGGAGAT
HE663995 Conservation LT222361 MM955426 HE664018 M7766280 HE663991 HE663991 HE801149 LT222382 HE801199 HE663992 LT222382 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT22385 LT222385 LT222385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT22385 LT2	CGTGAACCAT TGGCCCACCG GGCCCACCG		AACGCA64GTT CCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		CCGTCCCCCC CCGTCCCCCC CCGTCCCCCC CCGTCCCCCCCC	**************************************	CCCATCOLOGIC CCCATCOLOGIC CCCATCOLOGIC CCCATCOLO CCCATCO	GICACGGCTC M GACGAATACT	CCGTCGCTCC CCGTCGTTC CCGTCGTTGTTC CCGTCGTTGTTC CCGTCGTTGTTC CCGTCGTTGTTC CCGTCGTTGTTC CCGTCGTC CCGTC CCGTCGTC CCGTCGTC CCGTCGTC CCGTC		CCGTGCCCCC CCGTGCCCCCC	
HE663995 Conservation LT222361 MN955426 HE664018 MF766280 HE663991 HE663991 HE80140 LT222368 LT22368 LT2368 LT2368 LT2368 LT22368 LT2568 LT2568 LT	CGTGAACCAT TGGCCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCGC TGGCCACCGC TGGCC		AACGCA84 CCACCONCENT CCACCONC	000 000 000 000 000 000 000 000	CCATCCGCCC CCATCCGCCC CCATCCGCCC CCATCCGCCC CCATCCGTCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCCC CCATCCTCCCC CCATCCTCCCCC CCATCCTCCCCC CCATCCTCCCCC CCATCCTCCCCCCCC	2323A238288A8 2323A2382323 23232882323 23232882323 23232882328 2323288238 2323288238 2323288238 2323288238 2323288238 2323288238 232328238 2323282328 2323282328 2323282328 2323282328 2323282328 2323282328 2323282328 23232828 232388 232388 232388 232388 232388 232388 232388 232388 232388 232388 232388 232388 232388 232388 232388 232388 23238 232388 23238 23238 2328	CTACCTORECT CALL CALL CALL CALL	GICACGCCIC GACGAATACT	CCGTCGTCC CCGTCGTCC CCGTCGTCC CCGTCGTCC CCGTCGTTCC CCGTCGTCC CCGTC	CCACAGECC T GCCTACTOT GCCTAC		ATGCGGAGAT
HE663995 Conservation MP56426 HE664018 HE664018 HE663991 HE601149 L1222361 HE663991 HE601149 L1222363 L122363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L12363 L1363 L1236	CGTGAACCAT TGGCCCACCG TGGCCACCG TGGCCA		AACGCAGAGTT CCCCCCCCCCCCCCCCCCCCCCCCCCCC	200 200 200 200 200 200 200 200	CCATCCGGCC CCATCCGGCC CCATCCGGCC CCATCGTCGG CCATCGTCGCC	232A298298A8 44 5022989202 502298202 5029820 50298 5029820 502980 5029800 5029800 5029800 5029800 5029800 5029800	CCGAATGGTG GCGAATGGTG	CACGAATACT CACGAATACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGEC TGCCTACTCT TGCCTA		ATGCGGAGAT
HE663995 Conservation HE664018 HE664018 HE664018 HE664018 HE664018 HE664018 HE664018 HE664018 HE664092 HE664092 HE664092 HE664099 HE663992 HE664099 HE663992 HE664099 HE663992 HE664099 HE663992 HE664093 HE664003 HE664003 HE664003 HE664003 HE664003 HE664000	CGTGAACCAT TGGCCCACCG GGCCCCCCCC		AACGCA64GTT CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		CCGTCCCCCC CCGTCCCCCC CCGTCCCCCCCCCCCCC	*** *** *** *** *** *** *** *** *** **	CCCATCOLO CCCATC	GICACGGCIC GACGAATACT	CCGTCGCCCC CCGTCGCTCC CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTCG CCGTCGTCG CCGTCGTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTCGTCG CCGTCGTCG CCGTCGTCG CCGT		CCGTGCCCCG CCGTGCCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG CCGTGCCCCG	
HE663995 Conservation LT222381 MN955426 HE684018 HE684018 HE68491 HE68491 HE80149 LT222382 HE64409 HE752388 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 HE684099 HE683993 HE683993 HE683993 HE683993 HE801187 HE801125 HE8	CGTGAACCAT CGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCGC TGGCCAC		AACGCA84 TTAAACGCA84 CTA82020202 CTA82020202 CTA8202020 CTA82020	000 000 000	CCATCCGGCC CCATCCGCCC CCATCCGCCC CCATCCGCCC CCATCCGTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCCC CCATCCTCCCC CCATCCTCCCCC CCATCCTCCCCC CCATCCTCCCCCCCC	2323A238288A8 2323A2382323 23232882323 2323288232 232328232 23232823 232328232 232328232 232328232 232328232 23232823 23232823 23232823 23232823 23232823 23232823 23232823 2323282 2323282 2323282 2323282 2323282 2323282 2323282 232382	CTACCTORECT CONTRACTOR CONTRACTOR CANAROSTO CONTRACTO CO	CACGAATACT CACGAATACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGECC T GCCTACTOT GCCTAC		ATECCEGAGAT 0 0 0 0 0 0 0 0 0 0 0 0 0
HE663995 Conservation MP56426 HE664018 HE664018 HE664091 HE664091 HE664095 HE663991 HE663991 HE663992 HE664005 MV775331 L1222362 L1222363 L12222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L1222363 L122363 L122363 L122363 HE663993 HE663993 HE663993 HE663993 HE601123 HE801163 HE801123 HE80123 HE	CGTGAACCAT TGGCCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCACCG TGGCCA		AACGCA64CTT CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	000 000 000	CCATCCGGCC CCATCCGGCC CCATCCGGCC CCATCGTCGGC	232A2362382A2 2022A2382323 20228823 2022882323 2022882323 2022882323 2022882323 2022882323 2022882323 2022882323 2022882323 2022882323 2022882323 2022882323 2022882323 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 2022882 20288 202882 202882 202882 202882 202882 20288 202882 20288	CCCATCGTC CCCCATCGTC CCCAT	GICACGCCIC CACGAATACT GACGAATACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGEC CCACAGEC TGCCTACTOT TGCCTA		ATGCGGAGAT
HE663995 Conservation LT222361 MM955426 HE664018 M7766280 HE664018 HE80149 LT222382 HE80408 M7775330 M7775330 M7775330 M7775330 M7775330 M7775330 M7775330 HE78298 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT222386 LT2386 LT2386	CGTGAACCAT TGGCCCACCG GGCCCCCCCC		AACGCA84 CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		CCGTCCGCCC CCGTCCGCC CCGTCCGCC CCGTCCGCC CCGTCCGCCC CCGTCCGCCC CCGTCCGCCC CCGTCCGCCC CCGTCCGCCC CCGTCGTCCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCC CCGTCGTCCCC CCGTCGTCCCCCCCC	2023A2020A2020A2 2022A202022 202208202 20220820 20220820 20220820 20220820 20220820 202 202	CCGAATGGTG CCGGAATGGTG CCGAATGGTG	CACGAGATACT GACGAATACT	CCGTCGTCGTCC CCGTCGTCGTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTTGTTCG CCGTCGTCG CCGTCGTCG CCGTCGTCG CCGTCGTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCGTTCG CCGTCG CCGTCG C			
HE663995 Conservation LT222381 MN955426 HE684018 HE684018 HE684018 HE801408 LT222382 HE644009 HE722386 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 LT222385 HE644009 HE663992 HE663993 HE663993 HE801103 HE801105	CGTGAACCAT CGCCACCG TGGCCACCGC TGGCCACCGC TGGCCACCGC TGGCCACCGC TGGCCACC		AACGCA84 CT284202222 CT2842022 CT2842022 CT2842022 CT284202	000 000 000	CCATCCGCCC CCATCCGCCC CCATCCGCCC CCATCCGCCC CCATCCGTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCC CCATCCTCCCCC CCATCCTCCCCC CCATCCTCCCCC CCATCCTCCCCC CCATCCTCCCCCCCC	232AA28028A2 232AA280232 2332382332 2332388233 2332388233 2332388233 2332388233 2332388233 2332388233 2332388233 233238232 233238232 233238232 233238232 233238232 233238232 233238232 233238232 233238232 233238232 233238232 233238232 233238232 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 23323823 2332382 233282 23	CTACCTORES CONTRACTOR CONTACTOR CONTRACTON CONTRACTON CONTRACTON CONTRACTON CONTRACTON C	CACGACTACT CACGACTACT	CCGTCGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCACAGECC See CCACAGEC Sec CCCACTON CCCTACTON CCCCTACTON CCCCTACTON CCCTACTON CCCTACTON CCCTACTO		ATECCEGAGAT
HE663995 Conservation L1222361 MN955426 HE664018 HE664019 HE601919 HE601919 HE601919 HE601919 HE601919 HE601919 HE601919 HE601919 HE603992 HE664009 HE603992 HE663983 HE60197 HE60192	CGTGAACCAT TGGCCCACCG TGGCCCCCCCCCCCG TGGCCCCCCCC TGGCCCCCCCCCC		AACGCA04GTT CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		CCATCCGGCC CCATCCGGCC CCATCCGGCC CCATCGTCGCC CCATCGTCGCC CCATCGTCGCC CCATCGTCGTCGC CCATCGTCGCC CCATCGTCGC CCAT	232A29829848 44 44 5302989203 5302988203 53029820 53029820 53029820 530298200 530298200 53029820 53029820 53029820 53	CTECCTEGEC CCCATECT CCCCATECT CCCATECT	GICACGGCIC GACGAATACT	CCGTCGCCCC CCGTCGTTCC CCGTCGTCC CCGTCGTCC CCGTCGTCC CCGTCGTCC CCGTCGTCC CCGTCGTTCC CCGTCGTCC CCGTCCC CCGTCC CCGTCC CCGTCCC CCGTCC CCGTCC	CCACAGEC CCACAGEC TGCCTACTOT TGCCTA		ATECCEGAGAT

Figure A1. Cont.

		620		640		660		680		700	
LT222361	ACGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AAGGAATGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 607
MN955426	ACGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AAGGAATGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 607
HE664018	ACGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AAGGAATGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 607
MF766260	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAATGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 607
HE663973	AYGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAYGAT	TCCGTCTYGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 606
HE663991	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAACGAT	TCCGTCTGGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 607
HE801149	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAACGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 611
LT222362	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAACGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 607
HE664005	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAACGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 607
MW775330	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAACGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 611
MW775331	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAACGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 611
LT222368	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 668
LT222363	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 668
LT222365	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CATCC 657
OM368594	AYGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAT	TTCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGACC	CGTCC 610
HE663982	ACGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAT	TTCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGAGC	CGTCC 610
HE664009	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAT	TCCGTCTCGC	TTTC		GA	CAAGGAAGG-	GAGAAGGAGC	CGTCC 652
HE663992	ATGCGACGTG	CCGTCGTCGG	ACCCCTCCCC	AAGGAAGGAT	TTCGTCTCGC	TTTC		GA	TAAGAAAGG-	GAGAAGGAGC	CGTTC 639
LM993453	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAT	TTCGTCTCGC	TCTC		G <mark>G</mark>	GAAGAGAGGG	AGAAGGACC	CGTCC 652
LM993448	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGA <mark>G</mark>	TTCGTCTCGC	TCTC		G <mark>G</mark>	GAAGAGAGGG	AGAAGGACC	CGTCC 652
LM993466	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAT	TTGTCTCGC	TCTC		G <mark>G</mark>	GAAGAGAGGG	AGAAGGACC	CGTCC 649
LM993461	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGA <mark>G</mark>	TCGTCTCGC	TCTC		GA	TAAGAGAGG-	GAGAAGGATC	CGTCC 652
LM993438	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGA <mark>G</mark>	TTCGTCTCGC	TCTC		GA	TAAGAGAGG-	GAGAAGGATC	CGTCC 656
HE801159	AAGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGA-	CCCCTCAC	Τ		· · · · · · · · · · · · · · · · · · ·		AGAAGGATC	CGTCC 631
HE663983	ATGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAC	CCTCTCTTGA	TTGCTTGCTT	GCTTTTC	TAAGAAGA	GAAAGAAAGA	AGGAAGGATC	CGTCC 665
HE801141	ATGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAC	CCTCTCTTGC	TTGC		GA	GAAAGAAAGA	AGGAAGGATC	CGTCC 647
HE801175	ATGCGACGTG	CCATCGTCGG	ACCCCTCACC	AAGGAAGGAC	CCTCTCTTGC	CTGCTTGCCT	TTC	TAAGAGAA	GAAAGTAAGG	AGGAAGGATT	CGTCC 663
HE801063	ATGCGACGTG	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAC	CCTCTCTTGC	CCGCTTGCCT	TTC	TAAGAGAA	GAAAGTACGG	AGCAAGGATT	CATCC 664
HE801127	GTGCGACGCG	CCGTCGTCGG	ACCCCTCACC	GAGGAAGGAC	CCTCTTCTCC	TTCT	TCCT	AGAAG	AAGGAGAAGG	ATGGATACGT	- GTCC 639
HE801123	GTGCGACGCG	CCGTCGTCGG	ACCCCTCACC	GAGGAAGGAC	CCTCTTCTCC	TT CT	TCCT	CGAAG	AAGGAGAAGG	ATGGATACGT	- GTCC 639
HE801166	GTACGACGCA	CCGTCGTCGG	ACCCCTCACC	AAGGAAGGAC	CCTCTTCTGC	TTCT	TCCAGA	AGAAG	AAGAAGGAGG	AGGGATCCGT	- GTCC 636
HE663958	ATGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGTC	CCTCTCTTGC	TT-CTTCTTC	TCCTAAT-AG	CTAGGAGGAG	GAGGAGGAAG	AGGAAGGAAT	CGTCC 671
HE663964	ATGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGTC	CCTCTCTTGC	TT-CTTCTTC	TCCTAAT-AG	CTGGGAGGAG	GAGGAGGAAG	AGGAAGGAAT	CGTCC 671
HE664017	A T G C G A <mark>T</mark> G <mark>C</mark> G	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGTC	CCTCTCTTGC	TTCTTC	TCCTAAT-AG	CT AGGAG	GAGGAGGAAG	AGGAAGGAAT	CGTCC 666
HE663969	ATGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGTC	CCTCCCTCTC	TTGCTTCTTC	GCCTAGTTAT	CTAGGAGGAG	GAGGAGGAAG	AGGAAGGAAT	CGTCC 678
HE663995	ATGCGACGCG	CCGTCGTCGG	ACCCCTCACC	AGGGAAGGTC	CCTCTCTCTC	TTGCTTCTTC	TCCTAGTTAG	CTAGGAGGAG	GAGGAGGAGG	AGGAAGGAAT	CGTCC 673
Concervation											
GUIISEIVAUUII											

Figure A1. Alignment of the evaluated ITS sequences: visualization [45].

Table A1. Data for the ITS sequences from the NCBI GenBank database used for the phylogenetic analysis (Figure 7). Author names are abbreviated as follows: DH—D. Harpke, SM—S. Meng, TR—T. Rutten, HK—K. Kerndorff, FB—F. Blattner, EP—E. Pasche, LP—L. Peruzzi, SN—S. Naimov, EA—E. Apostolova, KS—K. Stoyanov, TR—T. Raycheva, TK—T. Karamplianis, TC—T. Constantinidis, VR—V. Randjelovic, MJ—M. Jusckovic, NR—N. Randjelovic, and IR—I. Raca. Evaluated herbarium specimens are signed with "!". Voucher specimens reviewed via GBIF [34,35].

Species	NCBI Genbank Accession Number	Authors	Country	Voucher Specimen
C. adamii	HE663958	DH, SM, TR, HK, FB	Armenia	GAT 7161 -!
C. adamii	HE663964	DH, SM, TR, HK, FB	Armenia	B 10 0355320
C. adamioides	LT222361	DH, HK, EP, LP	Turkey	GAT 7136 -!
C. adamioides	MN955426	SN, EA, KS, TR	Bulgaria	SOA 062625 -!
C. aerius	HE663995	DH, SM, TR, HK, FB	Turkey	GAT 7178 -!
C. albocoronatus	HE664017	DH, SM, TR, HK, FB	-	GAT 7477 -!
C. aleppicus	HE801175	DH, SM, TR, HK, FB	Jordan	IABH 18357
C. alexandri	LT222362	DH, HK, EP, LP	Serbia	GAT 23024 -!
C. ancyrensis	LM993461	DH, LP, HK, TK, TC, VR, NR, MJ, FB	Turkey	GAT 29387 -!
C. babadagnensis	LT222363	DH, HK, EP, LP		GAT 7185
C. biflorus	LT222365	DH, HK, EP, LP	Italy	GAT25812 -!
C. boryi	HE801127	DH, SM, TR, HK, FB	cult.	GAT 7209 -!
C. chrysanthus	OM368594	SN, EA, KS, TR	Bulgaria	SOA 062596 -!
C. chrysanthus	HE663982	DH, SM, TR, HK, FB	Turkey	GAT 7138 -!
C. cf. chrysanthus	HE664009	DH, SM, TR, HK, FB	Turkey	GAT 7449 -!
C. danfordiae	HE663992	DH, SM, TR, HK, FB	Turkey	GAT 7123 -!
C. danubensis	LM993453	DH, LP, HK, TK, TC, VR, NR, MJ, FB	Serbia	GAT 23019 -!
C. fleischeri	HE663983	DH, SM, TR, HK, FB	Turkey	GAT 7139 -!
C. kerndorffiorum	HE801159	DH, SM, TR, HK, FB	Turkey	K 00802496 -!
C. laevigatus	HE801166	DH, SM, TR, HK, FB	cult.	GAT 7224 -!
C. leucostylosus	HE663973	DH, SM, TR, HK, FB	Turkey	GAT 7471 -!
C. micranthus	LM993438	DH, LP, HK, TK, TC, VR, NR, MJ, FB	Turkey	GAT 25831 -!
C. minutus	HE664005	DH, SM, TR, HK, FB	Turkey	GAT 7470 -!
C. pallidus	MW775330	SN, EA, KS, TR	Bulgaria	SOA 062797 -!
C. pallidus	MW775331	SN, EA, KS, TR	Bulgaria	SOA 062791 -!
C. pestalozzae	HE801141	DH, SM, TR, HK, FB	Turkey	GAT 7261 -!
C. pseudonubigena	HE663969	DH, SM, TR, HK, FB	Turkey	GAT 7452 -!
C. pulchricolor	HE664018	DH, SM, TR, HK, FB	Turkey	GAT 7188 -!

Species	NCBI Genbank Accession Number	Authors	Country	Voucher Specimen
C. punctatus	HE663991	DH, SM, TR, HK, FB	Turkey	GAT 7155 -!
C. randjeloviciorum	MF766260	DH, HK, IR, EP	Serbia	GAT 23042 -!
C. reticulatus	LM993448	DH, LP, HK, TK, TC, VR, NR, J, FB	Russia	GAT 23082 -!
C. rhodensis	LT222368	DH, HK, EP, LP	Greece	GAT s/n
C. tournefortii	HE801123	DH, SM, TR, HK, FB	Greece	GAT 7202 -!
C. variegatus	LM993466	DH, LP, HK, TK, TC, VR, NR, MJ, FB		GAT 7264
C. veneris	HE801063	DH, SM, TR, HK, FB	cult.	GAT 7201 -!
C. weldenii	HE801149	DH, SM, TR, FB	cult.	GAT 7392 -!

Table A1. Cont.

References

- Harpke, D.; Meng, S.; Rutten, T.; Kerndorff, H.; Blattner, F.R. Phylogeny of *Crocus* (Iridaceae) based on one chloroplast and two nuclear loci: Ancient hybridization and chromosome number evolution. *Mol. Phylogenetics Evol.* 2013, 66, 617–627. [CrossRef] [PubMed]
- 2. Rukšāns, J. The World of Crocuses; Latvian Academy of Sciences: Riga, Latvia, 2017; p. 568.
- 3. Yüzbaşıoğlu, S. *Crocus keltepensis* (Iridaceae): A new *Crocus* species from the highest part of the Samanlı Mountains, NW Turkey. *Phytotaxa* **2019**, *418*, 230–236. [CrossRef]
- 4. Raca, I.; Harpke, D.; Shuka, L.; Ranđelović, V. A new species of *Crocus* ser. *Verni* (Iridaceae) with 2n = 12 chromosomes from the Balkans. *Plant Biosyst.—Int. J. Deal. All Asp. Plant Biol.* **2020**, 1–7. [CrossRef]
- Velchev, V. Crocus L. In Flora Reipublicae Popularis Bulgaricae, 2nd ed.; Jordanov, D., Ed.; Academia Scientarum Bulgaricae: Sofia, Bulgaria, 1964; pp. 328–338.
- Kitanov, B.; Drenkovski, R. Über die Taxonomie einiger Sippen der Gattung Crocus L. auf der Balkanhalbinsel. In Problems of Balkan Flora and Vegetation, Proceedings of the First International Symposium on Balkan Flora and Vegetation, Varna, Bulgaria, 7–14 June 1973; Bulgarian Academy of Science: Sofia, Bulgaria, 1975; pp. 212–215.
- Kitanov, B.; Penev, I.; Krachunov, I. Materials on the flora of Dobroudja. Annu. Sofia Univ. St. Kliment Ohridski Fac. Biol. 1977, 70, 85–87.
- 8. Kitanov, B.; Penev, I. Flora of Dobroudja; Naouka & Izkoustvo: Sofia, Bulgaria, 1980; p. 630.
- 9. Delipavlov, D. *Crocus* L. In *Key to the Plants in Bulgaria;* Delipavlov, D., Cheschmedzhiev, I., Eds.; Agrarian University Academic Pres: Plovdiv, Bulgaria, 2003; pp. 456–457.
- 10. Petrova, A.; Vladimirov, V. Balkan endemics in the Bulgarian flora. *Phytol. Balc.* 2010, *16*, 293–311.
- 11. Assyov, B.; Petrova, A. *Conspectus of the Bulgarian Vascular Flora*, 4th ed.; Distribution Maps and Floristic Elements; Bulgarian Biodiversity Foundation: Sofia, Bulgaria, 2012; pp. 156–157.
- 12. Pulević, V. Two new species of the genus *Crocus* L. from Yugoslavia. *Glas. Republičkog Zavoda Zaštitu Prir. Prir. Zbirke Titogr.* **1976**, *9*, 39–43.
- 13. Mathew, B. The Crocus, a Revision of the Genus Crocus; BT Batsford Ltd.: London, UK, 1982; p. 127.
- 14. WCSP. World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. 2021. Available online: http://wcsp.science.kew.org/ (accessed on 4 January 2021).
- 15. Pavletić, Z.; Trinajstić, I. Prilog poznavanju horologije taksona *Crocus weldenii* Hoppe et Fürnrohr u srednjoj Dalmaciji. *Acta Bot. Croat.* **1979**, *38*, 163–166.
- Pulević, V. Prilog taksonomiji i horologiji nekih vrsta roda Crocus L. iz Jugoslavije (C. weldenii, C. alexandri i C. adamii sens. lat.). Glas. Republičkog Zavoda Zaštitu Prir. Prir. Zbirke Titogr. 1977, 10, 81–99.
- 17. Ranđelović, N.; Hill, D.A.; Ranđelović, V. *The Genus Crocus L. in Serbia*; The Serbian Academy of Sciences and Arts: Belgrade, Serbia, 1990; p. 49.
- 18. Rakaj, M. Floristic and chorological news from north Albania. Bot. Serbica 2009, 33, 177–183.
- Ranđelović, N.; Sekovski, Ž.; Dimeska, G. Systematic, chorological and genetic research into the genus *Crocus* L. in Macedonia. In *Collection of Papers Devoted to Academician Kiril Micevski*; Cyril and Methodius University in Skopje: Skopje, North Macedonia, 2007; pp. 98–131.
- 20. Brighton, C.A.; Mathew, B.; Marchant, C. Chromosome counts in the genus Crocus (Iridaceae). *Kew Bull.* **1973**, *28*, 451–464. [CrossRef]
- 21. Goldblatt, P.; Takei, M. Chromosome cytology of Iridaceae, patterns of variation, determination of ancestral base numbers, and modes of karyotype change. *Ann. Mo. Bot. Gard.* **1997**, *84*, 285–304. [CrossRef]
- 22. Schneider, I.; Kerndorff, H.; Pasche, E. Chromosome numbers of Turkish *Crocus* (Liliiflorae, Iridaceae) and their geographical distribution. *Feddes Repert.* 2012, 123, 73–79. [CrossRef]

- Harpke, D.; Peruzzi, L.; Kerndorff, H.; Karamplianis, T.; Constantinidis, T.; Ranđelović, V.; Ranđelović, N.; Jušković, M.; Pasche, E.; Blattner, F.R. Phylogeny, geographic distribution and new taxonomic circumscription of the *Crocus reticulatus* species group (Iridaceae). *Turk. J. Bot.* 2014, *38*, 1182–1198. [CrossRef]
- 24. Harpke, D.; Kerndorff, H.; Pasche, E.; Peruzzi, L. Neotypification of the name *Crocus biflorus* Mill. (Iridaceae) and its consequences in the taxonomy of the genus. *Phytotaxa* **2016**, *260*, 131–143. [CrossRef]
- 25. Mosolygó, Á.; Sramkó, G.; Barabás, S.; Czeglédi, L.; Jávor, A.; Molnár, A.; Surányi, G. Molecular genetic evidence for allotetraploid hybrid speciation in the genus *Crocus* L. (Iridaceae). *Phytotaxa* **2016**, 258, 121–136. [CrossRef]
- 26. NCBI GenBank Nucleotide Database. Available online: https://www.ncbi.nlm.nih.gov/genbank/ (accessed on 10 January 2022).
- Yüzbaşioğlu, S.; Aslan, S.; Özhatay, N. Crocus thracicus (Iridaceae), a new species from north-western Turkey. Phytotaxa 2015, 197, 207–214. [CrossRef]
- Harpke, D.; Kerndorff, H.; Raca, I.; Pasche, E. A new Serbian endemic species of the genus *Crocus* (Iridaceae). *Biol. Nyssana* 2017, 8, 7–13. [CrossRef]
- 29. Kerndorff, H.; Pasche, E.; Harpke, D. The Genus *Crocus* (Liliflorae, Iridaceae): Life-cycle, morphology, phenotypic characteristics, and taxonomical relevant parameters. *Stapfia* **2015**, *103*, 27–65.
- Aykurt, C.; Deniz, I.G.; Başaran, M.A.; Genç, G.E. Taxonomic and GIS-based ecologic studies on critically endangered *Crocus* wattiorum B. Mathew (Olympos saffron). *Thaiszia-J. Bot.* 2015, 25, 73–84.
- 31. Hršak, V.; Lukač, G. The genus *Crocus* L. in the Paklenica National Park. *Acta Bot. Croat.* **2001**, *60*, 97–101. Available online: https://hrcak.srce.hr/160820 (accessed on 10 February 2021).
- 32. Milović, M. Rod *Crocus* L. (Iridaceae) u flori Hrvatske. *Glas. Hrvat. Bot. Društva* **2016**, *4*, 4–20. Available online: https://hrcak.srce.hr/167881 (accessed on 28 January 2021).
- 33. Kerndorff, H.; Pasche, E. Two New Taxa of Crocus (Liliiflorae, Iridaceae) from Turkey. Stapfia 2011, 95, 2–5.
- Leibniz Institute of Plant Genetics and Crop Plant Research (IPK). Herbarium GAT. Occurrence Dataset. GBIF.org. Available online: https://www.gbif.org/dataset/5d47cb09-3ef2-477d-a539-dd7f6955aefd (accessed on 20 January 2022).
- Royal Botanic Gardens, Kew. Royal Botanic Gardens, Kew-Herbarium Specimens. Occurrence Dataset. GBIF.org. 2021. Available online: https://www.gbif.org/dataset/cd6e21c8-9e8a-493a-8a76-fbf7862069e5 (accessed on 20 January 2022).
- Steiermärkisches Landesmuseum Joanneum-Herbarium GJO. Steiermärkisches Landesmuseum Joanneum-Herbarium GJO. Occurrence Dataset. GBIF.org. Available online: https://www.gbif.org/dataset/85a214d6-f762-11e1-a439-00145eb45e9a/ (accessed on 13 February 2022).
- Bijmoer, R.; Scherrenberg, M.; Creuwels, J. Naturalis Biodiversity Center (NL)-Botany. Naturalis Biodiversity Center. Occurrence Dataset. GBIF.org. 2022. Available online: https://www.gbif.org/dataset/15f819bd-6612-4447-854b-14d12ee1022d (accessed on 13 February 2022).
- University of Oslo. Vascular Plant Herbarium, Oslo (O) UiO. Version 1.1754. Occurrence Dataset. GBIF.org. 2022. Available online: https://www.gbif.org/dataset/e45c7d91-81c6-4455-86e3-2965a5739b1f/ (accessed on 13 February 2022).
- 39. Van Westen, M. Micam 2.4. Available online: http://science4all.nl (accessed on 10 October 2019).
- Tirel, C.; Jérémie, J.; Lobreau-Callen, D. Corchorus neocaledonicus (Tiliaceae), véritable identité de l'énigmatique Oceanopapaver. Bull. Du Muséum Natl. D'histoire Nat. 1996, 18, 35–43.
- 41. NCBI. Standard Nucleotide BLAST. Available online: https://blast.ncbi.nlm.nih.gov/Blast.cgi (accessed on 10 January 2022).
- Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994, 22, 4673–4680. [CrossRef]
- Ronquist, F.; Teslenko, M.; Van der Mark, P.; Ayres, D.L.; Darling, A.; Hohna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 2012, *61*, 539–542. [CrossRef]
- Stöver, B.C.; Müller, K.F. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinform.* 2010, 11, 7. Available online: https://bmcbioinformatics.biomedcentral.com/track/pdf/10.1186/1471-2105-11-7.pdf (accessed on 10 June 2019). [CrossRef]
- CLC Sequence Viewer 8.0, QIAGEN. Available online: https://clc-sequence-viewer.software.informer.com/8.0/ (accessed on 27 October 2021).