## **Original Article**

Cytogenetic and Genome Research

Cytogenet Genome Res 2009;124:81–89 DOI: 10.1159/000200091 Accepted after revision: September 23, 2008 by B. Friebe

# Characterization of RUSI, a telomere-associated satellite DNA, in the genus *Rumex* (Polygonaceae)

R. Navajas-Pérez<sup>a</sup> T. Schwarzacher<sup>b</sup> M. Ruiz Rejón<sup>c</sup> M.A. Garrido-Ramos<sup>c</sup>

<sup>a</sup>Plant Genome Mapping Laboratory, University of Georgia, Athens, GA (USA); <sup>b</sup>Department of Biology, University of Leicester, Leicester (UK); <sup>c</sup>Departamento de Genética, Facultad de Ciencias, Universidad de Granada, Granada (Spain)

#### Abstract

A satellite-DNA family (RUSI) has been isolated and characterized in Rumex induratus Boiss and Reuter (Polygonaceae), an Iberian endemic polygamous sorrel. The RUSI repeats are 170 bp in length and  $\sim$  68% AT-rich containing different variants of degenerate telomere motifs – (TT)<sub>n</sub>AN(GG)<sub>n</sub> –, a typical feature of subtelomeric DNA repeats adjacent to telomeres, which have been referred to as telomere-associated sequences or TASs. In fact, fluorescent in situ hybridization showed that this satellite DNA is located in subtelomeric positions of most of the chromosomes of R. induratus, with some centromeric loci. PCR and Southern-blot hybridization assays for sequence conservation in the genus Rumex, indicated that the RUSI sequences are restricted to the genomes of R. induratus and R. scutatus, both species of the section Scutati, suggesting that they are recently evolved. Sequence variation within the two species is high (mean value of sequence differences between repeats of 15% for R. induratus and 7.5% for R. scutatus) and the degree of sequence differentiation between species is low with no species-specific variants, postulated to be due to slowed rates of spreading of sequence variants by molecular homogenizing mechanisms. Characteristics of RUSI sequences are discussed in the light of their chromosomal location and analyzed for their evolutionary and phylogenetic implications.

Copyright © 2009 S. Karger AG, Basel

## KARGER

Fax +41 61 306 12 34 E-Mail karger@karger.ch www.karger.com © 2009 S. Karger AG, Basel 1424–8581/09/1241–0081\$26.00/0

Accessible online at: www.karger.com/cgr

Rumex (Polygonaceae) is a plant genus containing dioecious, gynodioecious, polygamous and hermaphroditic species (Rechinger, 1964). Many scientific programs have focused on this genus because of its biological and evolutionary significance in sexual dimorphism. Within the genus Rumex three phylogenetic clades can be defined (Navajas-Pérez et al., 2005a; Fig. 1). The basal clade is composed of hermaphroditic docks, while the second comprises polygamous/gynodioecious sorrels and the third dioecious sorrels except for one hermaphrodite, this latter clade containing species with different sexdetermination systems (XX/XY and XX/XY<sub>1</sub>Y<sub>2</sub>). Polygamous and gynodioecious sorrels have important evolutionary significance because they are intermediate forms on the way toward the evolution of separate sexes (Navajas-Pérez et al., 2005a) and could help unravel the origin of dioecy and sex chromosomes in Rumex species. Furthermore, these species have applied interests in medicine (Rivera and Obón, 1995), as a dietary source of potential bioactive compounds (Ferreres et al., 2006) or have been proposed as candidates for Hg phytoremedia-

This research was supported by grant CGL2006-00444/BOS awarded by the Ministerio of Educación y Ciencia (Spain), and Fondo Europeo de Desarrollo Regional (FEDER, EU). R.N.-P. is a Fulbright postdoctoral scholar (FU-2006-0675) supported by Spanish MEC.

Rafael Navajas-Pérez Current address: Departamento de Genética, Facultad de Ciencias Universidad de Granada, Granada (Spain) telephone: +34 958 24 30 80; fax: +34 958 24 40 73; e-mail: rnavajas@ugr.es

	Species	Subgenus/section	Mating/sex chromosome system	х	Origin
Rumex acetosa		Acetosa/Acetosa	Dioecy (XX/XY <sub>1</sub> Y <sub>2</sub> )	7	Capileira, Granada (Spain)
	Rumex papillaris	Acetosa/Acetosa	Dioecy (XX/XY <sub>1</sub> Y <sub>2</sub> )	7	La Benajara, S <sup>a</sup> Baza, Granada (Spain)
	Rumex tuberosus	Acetosa/Acetosa	Dioecy (XX/XY <sub>1</sub> Y <sub>2</sub> )	7	Sinjar (Iraq)
	Rumex intermedius	Acetosa/Acetosa	Dioecy (XX/XY <sub>1</sub> Y <sub>2</sub> )	7	Vollubilis (Morocco)
	Rumex thyrsoides	Acetosa/Acetosa	Dioecy (XX/XY <sub>1</sub> Y <sub>2</sub> )	7	Vollubilis (Morocco)
	Rumex acetosella	Acetosella/Acetosella	Dioecy (XX/XY)	7	Capileira, Granada (Spain)
	Rumex hastatulus (RNC)	Acetosa/Americanae	Dioecy (XX/XY <sub>1</sub> Y <sub>2</sub> )	4	Cumberland County, North Carolina (USA)
	Rumex hastatulus (RTX)	Acetosa/Americanae	Dioecy (XX/XY)	5	Masan County, Texas (USA)
	Rumex suffruticosus	Acetosa/Scutati	Dioecy (XX/XY)	8	Pto. Navacerrada, Segovia (Spain)
	Rumex bucephalophorus	Platypodium/Platypodium	Hermaphroditism	8	Padul, Granada (Spain)
	Rumex lunaria	Acetosa/Hastati	Polygamy-Gynodioecy	9	Gáldar, Gran Canaria (Spain)
	Rumex vesicarius	Acetosa/Vesicarii	Hermaphroditism-Polygamy	9	Presa de Ayagures, Gran Canaria (Spain)
	Rumex roseus	Acetosa/Scutati	Hermaphroditism-Polygamy	10	Tarancón, Cuenca (Spain)
	Rumex induratus	Acetosa/Scutati	Hermaphroditism-Polygamy	10	Padul, Granada (Spain)
	Rumex scutatus	Acetosa/Scutati	Hermaphroditism-Polygamy	10	S <sup>a</sup> Mágina, Jaén (Spain)
	Rumex maderensis	Acetosa/Hastati	Polygamy-Gynodioecy	10	Real Jardín Botánico de Madrid (Spain)
	Rumex crispus	Rumex/Rumex	Hermaphroditism	10	Atarfe, Granada (Spain)
	Rumex patientia	Rumex/Rumex	Hermaphroditism	10	Real Jardín Botánico de Madrid (Spain)
	Rumex pulcher	Rumex/Rumex	Hermaphroditism	10	Guadarrama, Madrid (Spain)
	Rumex conglomeratus	Rumex/Rumex	Hermaphroditism	10	Atarfe, Granada (Spain)
	Rumex obtusifolius	Rumex/Rumex	Hermaphroditism	10	Guadarrama, Madrid (Spain)

**Fig. 1.** List *Rumex* species analyzed in this paper, indicating their affiliations, mating/sex chromosome system, basic chromosome number (x), and their phylogenetic relationship (left – based on Navajas-Pérez et al., 2005a).

tion of contaminated soils (Moreno-Jiménez et al., 2006). In this context, we are interested here in the cytogenetic and molecular analysis of *Rumex induratus* Boiss and Reuter (Polygonaceae), an Iberian endemic polygamous sorrel. Ultimately, we aim to understand the phylogenetic aspects of sequence and chromosomal evolution in *Rumex* using DNA markers to study the processes involved in karyotype and sex-chromosome evolution. In particular, satellite-DNA and other repetitive-DNA sequences have proved to be powerful markers for these types of evolutionary studies (Heslop-Harrison, 2000; Navajas-Pérez et al., 2005b, 2006; Mariotti et al., 2006; Cuñado et al., 2007).

Satellite-DNA sequences are non-coding highly repetitive tandem arrayed sequences originally so named because they form shoulders – satellites – in density buoyant gradients if their AT versus GC content is significantly different from the bulk DNA, but later extended to all highly repeated tandemly organized sequences (Beridze, 1986). Together with retrotransposon-like dispersed sequences, they represent most of the DNA content in plants (Heslop-Harrison, 2000). Satellite-DNA sequences are located at heterochromatin which is found mostly in centromeric and subtelomeric regions in the chromosomes, but also at intercalary positions (reviewed in Sharma and Raina, 2005). Within satellite-DNA families, subtelomeric satellites are repetitive sequences adjacent to telomeres that have been referred to as telomere-associated sequences or TASs (Louis and Vershinin, 2005) which in addition to location have several specific characteristics such as the presence of telomere degenerated motifs within the repeat sequence and the ineffectiveness of sequence homogenization events (Contento et al., 2005).

In the present study, we analyze a new satellite-DNA family isolated from *R. induratus*, RUSI, comprised of 170-bp repeats. RUSI sequences are tandemly located in subtelomeric positions of most of the chromosomes of the species, with centromeric sites in some of them, and have main features to be regarded as TASs.

### **Materials and methods**

Seeds and leaves of *R. induratus* were collected from a natural population in Padul, Granada (Spain). Seeds were kept in a cold and dry place until germination and leaves sampled from up to ten individuals were stored at -80°C for further analysis. DNA isolation was performed using the Plant DNAzol kit (Invitrogen) following the manufacturer's recommendations. DNA from the following species was isolated previously and includes *R. acetosa*, *R. intermedius*, *R. thyrsoides*, *R. tuberosus*, *R. suffruticosus*, *R. acetosella*, *R. hastatulus* (Texas race and North Carolina race), *R. bu* 

cephalophorus, R. scutatus, R. lunaria, R. maderensis, R. roseus, R. vesicarius, R. conglomeratus, R. crispus, R. patientia, R. obtusifolius and R. pulcher, collected from different sources (for systematics and locations, see Fig. 1 and Navajas-Pérez et al., 2005a).

RUSI satellite-DNA was isolated after restriction analysis from *R. induratus* total genomic DNA using *Bgl*I restriction endonuclease and electrophoresis in agarose gel with ethidium bromide. The most prominent bands visible under UV light were excised from the gel and purified using GFX<sup>TM</sup> PCR DNA and Gel Band Purification Kit (Amersham Biosciences). Purified *Bgl*I fragments were then ligated to the *Bam*HI compatible ends of the pUC18 vector and cloned in competent DH5- $\alpha$  cells (Gibco BRL). Minipreparations were made using Perfectprep<sup>®</sup> Plasmid Mini (Eppendorf) and recombinant plasmids carrying the monomeric sequences were identified after screening with the purified band by dot-blot hybridization.

For Southern-blot (all the species listed in the first paragraph of this section were checked for the presence of RUSI sequences by the Southern-blot hybridization technique) and dot-blot hybridization and fluorescent in situ hybridization (FISH), a plasmid insert of clone RUSI\_19 was used as a probe. Southern-blot and dot-blot hybridizations were carried out following Garrido-Ramos et al. (1999). For FISH, chromosome preparations were made following Schwarzacher and Heslop-Harrison (2000). Probe was labeled with biotin-dUTP by random priming according to the specifications of the Random Primer Labelling System (Invitrogen). Labeled probe (25-50 ng) was added to the hybridization mixture (50% formamide, 2× SSC, 20% dextran sulphate, 0.125% SDS and 0.125 mM EDTA) (see Schwarzacher and Heslop-Harrison, 2000). Combined denaturation of the probe and chromosomal DNA was performed at 80°C for 8 min using a Thermo-Hybaid HyPro-20 and re-annealed at 37°C overnight. Stringent washes (20% formamide and 0.1× SSC at 42°C) were made prior to detection. Probe was detected with Alexa594/streptavidin (Molecular Probes, 0.5 ng/ml) in 5% (w/v) Bovine Serum Albumin (BSA) in 4× SSC, 0.2% Tween 20 following Schwarzacher and Heslop-Harrison (2000). Preparations were counterstained with 4,6-diamidino-2-phenylindole (DAPI, 2 µg/ml) and mounted in antifade solution. Preparations were analyzed with a Zeiss Axioplan 2 epifluorescence microscope (Oberkochen, Germany) with suitable filters and photographed with a CCD camera (Optronics, model s97790). Colour figures and overlays were prepared with Adobe Photoshop 7.0 software, using only those processing functions that are applied to all pixels of the image. Karyotyping was performed according to a classical orcein staining method for 4-5 h followed by mounting preparations in 45% acetic acid.

For amplification of RUSI sequences in *R. scutatus*, we designed the pair of specific primers RUSI-A, 5'-CGAGTTTCACT-TATTTGTCC-3', and RUSI-B, 5'-AAGTATACAAATTCAACC-TT-3', from the sequences isolated in *R. induratus*. Presence/absence status in the rest of species was also screened using the same primer combination. PCR amplifications were carried out in 50  $\mu$ l reactions containing 10 ng of purified DNA, 2 mM of dNTPs, 2 mM of each primer and 1.25 units of Taq polymerase in 10 mM Tris HCl at pH 8.3, 5 mM KCl, 2 mM MgCl<sub>2</sub> reaction buffer. Thermal cycles consisted of 1 min at 94°C, 1 min at 55°C and 1 min at 72°C. The PCR products were electrophoresed in agarose gels, thereafter the bands were cut out of the gel, purified and ligated to the cloning plasmid pGEM-Teasy (Promega) and cloned in

*Escherichia coli* JM109 competent cells (Promega) following the manufacturer's instructions.

Recombinant clones belonging to both a pUC18 and a pGEM library were sequenced by the dideoxy-sequencing method using the automatic ABI-Prism 377 sequencer (Applied Biosystems). The EMBL accession numbers for all the sequences analyzed in this paper are: AM398567 to AM398599.

For sequence analysis, multiple alignments were performed using Clustal X (Thompson et al., 1997) followed by manual adjustments. Basic sequence evolutionary calculations, as well as the detection of transition stages of satellite-DNA evolution (Strachan et al., 1985) and shared and non-shared polymorphisms were performed by the software satDNA Analyzer (Navajas-Pérez et al., 2007). Additionally, a Perl script able to recognize regular expressions in both strands was written to search specifically for degenerated telomere motifs.

### Results

# *Isolation and organization of a new satellite-DNA family in R. induratus*

When total genomic DNA of *R. induratus* was digested with *BglI* and electrophoresed in agarose gels stained with ethidium bromide, several prominent bands of 170-bp fold units were observed (not shown). We isolated and cloned these 170-bp fragments as candidate repeat units of a satellite-DNA family and 16 clones were selected for sequencing. One of these clones (RUSI\_19) was used as a probe for Southern blot hybridization against the genomic DNA of *R. induratus* digested with different restriction enzymes. A typical ladder pattern with a repeat unit of 170 bp was detected for enzymes *DraI* and *BglI* suggesting tandem repeat organization typical for satellite-DNA. In the case of enzymes *Eco*RI and *Hin*dIII a type B ladder-pattern was detected (Fig. 2, lanes 1–4).

DNA sequence analysis, showed that the 16 *Bgl*I repeats analyzed ranged from 167 to 182 bp in length, with an AT content of 67%. The mean percentage of variation among these sequences is 15%. Significantly, degenerated variants of the *Arabidopsis*-like TTTAGGG telomere repetitive motif matching the formula (TT)nAN(GG)n, were detected in both forward and reverse sense in most of the monomeric sequences (Fig. 3). No significant positive matches were detected when sequences were contrasted with the NCBI/EMBL database (July, 2008 release). Thus, they were considered to be a new satellite-DNA family, that we named RUSI (**RU**mex **S**cutati section Induratus species sequence).

RUSI, a telomere-associated satellite DNA in *Rumex* 



**Fig. 2.** Southern blot hybridization of RUSI sequences against total genomic DNA of *Rumex induratus* and *Rumex scutatus* digested with *Eco*RI (1, 5), *Hind*III (2, 6), *Dra*I (3, 7) and *BgI*I (4, 8) restriction enzymes. Size markers are indicated on the left. (\*) indicate monomeric repetitive unit (170 bp).

# Chromosomal analysis and location of RUSI in R. induratus karyotype

R. induratus is tetraploid, with a karyotype composed of 2n = 40. According to the chromosomal morphology and the fluorescent in situ hybridization (FISH) pattern of RUSI sequences, the karyotype of R. induratus can be grouped in two complements of ten chromosome pairs, six meta-/submetacentric and four subtelocentric pairs each confirming the basic chromosome number of the group, x = 10 (Figs. 4 and 5) (Navajas-Pérez et al., 2005a). Clusters of RUSI sequences are located in most of the larger chromosomes of R. induratus, mainly in subtelomeric positions. Subtelomeric/distal FISH signals were found either in one arm or in both arms depending on the chromosome. Additionally, a few chromosomes show hybridization signals indicating the presence of RUSI sequences in the centromeric region, exclusively or in addition to the subtelomeric sites. We then found subtelomeric/distal FISH signal of: (1) both arms in the two homeologous chromosome pairs I; (2) in the short arm of the

two homeologous pairs III and IV; (3) in the long arm of the two homeologous pairs VI and VII. Pairs II and V in both sets and pair III of one set show paracentromeric hybridization signal indicating the presence of RUSI sequences in the centromeric region exclusively, but strongly (pairs II) or weaker and in addition to the subtelomeric sites (pairs III and V). Homeologous chromosome pairs VIII to X have no evidence of hybridization (Fig. 5).

### Interspecific analysis of RUSI sequences

We screened the distribution of this new satellite-DNA family in representative species of the four subgenera of genus *Rumex* (López González, 1990). For that, we blotted onto a nylon membrane total genomic DNA digested with *Eco*RI of *Rumex* species listed in Fig. 1. Southern-blot hybridization with the RUSI\_19 clone gave a positive signal only in *R. scutatus* and in the positive control *R. indura-tus*. These results agree with additional PCR experiments using specific primers for RUSI sequences (not shown). The pattern of hybridization in *R. scutatus* was slightly different from that of *R. induratus* since all restriction enzymes used in Southern-blot hybridization experiments gave a type B ladder pattern (Fig. 2, lanes 5–8).

From PCR experiments, a total of 17 RUSI monomeric sequences were obtained from *R. scutatus*. These sequences showed the same features described for *R. induratus*, including an average size of 170 bp and a high AT content (66.7%). The degree of intraspecific identity was 7.5%. As in *R. induratus*, we also detected the presence of (TT)nAN(GG)n degenerated telomere motifs (Fig. 3).

For interspecific sequence comparison monomeric sequences isolated from R. induratus and R. scutatus were aligned requiring a 188-character dataset due to the inclusion of some indels as a consequence of insertions/deletions. The degree of interspecific divergence among sequences belonging to both species was 12%. A tree generated using an NJ approach does not reveal the existence of highly supported clades of sequences, which do not appear clustered according to taxonomic affinity (not shown). The study of the alignment position per position for the analysis of the different stages toward the homogenization according to Strachan's model (Strachan et al., 1985; Navajas-Pérez et al., 2007) revealed the lack of positions between transition stages IV to VI. Up to 65 (35%) positions were detected to be among II and III transition stages, while 35 (19%) of the positions represented shared polymorphic positions between the two species (Table 1). Therefore, no species-specific variants were detected and, in fact, consensus sequences were identical between the two species.

					~ ~ ~ ~ ~ ~						
RI_30B	TTTTCGAG-T-TTC	ACTTATTGT	CCGAA	AAA AGGTT-O	JAA- 11	I'TGTATACT	TATATATAA-	- TAATTTTATC	AAGT-TT	ATATT-TTTA	CGGTAT-TAT
RI_30	TTTACGAG-1-IIC	ACTIATIGI	CCGAA	AAA AGGIIIO	באאק II – אאק מידי איני	I IGIAIACI	TITATIGA-	- GAAIIIAAIC	AAGI-II	ATATI-TATA	CGGIAI-IAI
RI_3/B DT 10B	TTTTTCGAG-A-CIC	ACTIATIGI	CCGA	AAA AGGII-C	נו - אא <i>נ</i> ריד – גגי	TIGIAIACI PTGTATACT	TTTATATGA	- GAAIIIIAIC	AAII-II AATT-TT	ATATT-TGTA	CGGGAT-TAT
RT 37	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA AGGTT-	3AA- TT	TTGTATACT	TATT-ATTA-	- GAAATTTGTG	AGGT-TT	ATATT-TATA	TGGTAT-TCT
RI 19	TTTACGAGCT-TTC	ACTTATTTGT	CCGAZ	AAA AGGTT-0	3AA- T1	TTGTATACT	TATT-ATTA-	- GAAATTTGTG	AGGT-TT	ATATT-TATA	TGGTAT-TCT
RI 17B	TTTACGAG-T-TTC	ACTTATTTGT	CCAAA	ACA AGGTT-C	GAA- TI	TGTATCTT	TTTTTATGA-	- GAAATTTATC	ATGT-TT	TTATT-TGAA	AGGGAT-ACT
RI_17	TTTTCGAGCTCTTC	ACTTATTTGT	CCGTACA	ACA AGGTT-C	GAAA TI	TTGTATACT	TTTATATC	- GAGAGATATG	ATCAAGTATT	ATATTC <b>TGTA</b>	CGGTATTGAT
RI_50	TTTACGAG-T-TTC	CCTTATTTGT	GGCGA	AAA AGGAT-O	GAA- TI	TTCTATAGT	TTTATATCCA	A GAATTTTGTC	AAGT-TT	ATATT-TGTA	TTATAG-TAT
RI_49	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA A <mark>GGTT-0</mark>	GAAA -1	ITGTATACT	TTTATAAGA-	- GAATTTTATC	AAGT-TT	ATATT-TGTC	GGGTAT-TAT
RI_10	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA A <mark>GGTT-</mark> (	GAA- TI	ITGTATACT	TTTATATGA	- GAATTTGATC	AAGT-TT	ATATT-TGTA	CTATAT-TAT
RI_11	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA A <mark>GGTT-</mark> (	GAA- TI	ITGTATACT	TATATATAA-	- GAAAATTTTC	AAGT-TT	ATGTT-TGTA	CATTAT-TCT
RI_46	TTTACGAG-T-TTC	ACTTATTTGC	CCAAA	AAA AGGTT-1	raa- ti	TTGTATCTT	TTTTTATGA-	- GAAATTTGTC	AAGT-TT	TTATT-TGAA	AGGTAT-ACT
RI_4	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA AGGTT-0	3AA- T1	ITGTATACT	TTTATATGA-	- GAAATTTGTC	TAAT-TT	ATATT-TGTA	TGGTAT-TCT
RI_44	TTTACGAG-T-TTC	ACTTATTGT	CCGAA	AAA -GGTT-C	JAA- 'I''	I'I'I'I'A'I'AC'I'	GTTATATAA-	- GAAATTTGTC	AAGA-TT	TTATT - TGTA	CGGTAT-TCT
RI_20 PC 16	TTTACGAG-T-TTC	ACTIATIGI	CCGA	AAA -GGTT-C	רר – AAt	TIGIAIACI PTCTATACT	TTTTTATTA-	- GAAATTIGGG	AGGI-II	ACATT-TATA	CCCTAT-TCT
RS_10	TTTACGAG-T-TTC	ACTIATIO		AAA AGGII-C	388- IJ 388- TT	TIGIAIACI PTGTATACT	TTTATATGA-	- GAAIIIGAIG	ATCAAGI-II ATCAAGT-TT	ATATT-TTIA ATATT-TTIA	CGGTAT-TAT
RS 1	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA AGGTT-	3AA- TT	TTGTATACT	TTTTATATGA-	- GAATTTGATG	ATCAAGT-TT	ATGTT-TTTA	CGGTAGGTAT
RS 37	TTTACGAG-T-ATC	ACTTATTTGT	CCGA	AAA AGGTT-T	FAA- TI	TTGTATACT	TTTTTTATGA-	- GAATTTGATC	AAGT-TT	ATATT-TATA	CGGTAT-TAT
RS_11	TTTATGAG-T-TTC	ACTTATTTGT	CCGA	AAA AGGTT-O	GAA- TI	TTGTATACT	TTTTTATGA-	GAATTTGATG	ATCAAGT-TT	ATAAT- <b>TGTA</b>	CGGTAT-TAT
RS_17a	TTTACGAG-T-TTC	ACTTATTTGT	GGGA	AAA AGGTT-C	GAA- TI	ITGTATACT	TTTATATGA-	- GAATTTGATC	AAGT-TT	ATATT-TGTA	CGTTAT-TAT
RS_17b	TTGACGAG-T-TTC	ACTTATTTGT	CAGAA	AAA AGGTT-O	GAA- TI	FTGTATACT	TTTTTATGA	- GAATTTGATC	AAGT - TT	ATATT-TATA	CGGTAT-TAT
RS_35a	TTTACGAG-T-TTC	ACTTATTTGT	CAGAA	AAA AGGTT-O	GAA- TI	FTGTATACT	TTTATATGA-	- GCCTTTGATC	AAGT-TT	ATATT- <b>TGTA</b>	CGGTAT-TAT
RS_35b	TTTTCGAG-T-TTC	ACTTATTTGT	CAGAA	AAA AGGTT-C	GAA- TI	FTGTATACT	TTTATATGA-	- GAATTTGATC	AAGT-TT	ATATT-TGTA	CGTTAT-TAT
RS_T37	TTTACGAG-T-ATC	ACTTATTTGT	CCGA	AAA AGGTT-1	ГАА- Т'I	ITGTATACT	TTTTTATGA-	- GAATTTGATC	AAGT-TT	ATATT-TATA	CGGTAT-TAT
RS_18b	TTTACGAG-T-TTC	ACTTATTGT	CCGA	AAA AGGTT-C	JAA- 'I''I	I'TATATACT	'I''I''A'I'A'I'GA-	- GAATTITATC	AGGT-TT	ATATT TGTA	CGGTAT-T-T
RS_111 PC 10	TTCATCAC-T-TTC	ACTIATIGI	CCGA	AAA AGGII-( NNN NCCTT-(	ב באאכ חידי – האכ	TIGIAIACI PTGTATACT	TTTATATGA-	- GAAIIIIGIC	AAGI-II	ATATI-IGIA	CGGTAC-ICC
RS 5	TTTACGAG-T-TTC	ACTTATTGT	ACGA	AAA AGGII ( AAA AGGTT-(	3AA 11 3AA- T7	TTGTAAACT	TTTTAAATGA-	- GAATTTGATC	AA GI II AAGT-TT	ATATT-TATA	CGGTAT-TAT
RS 34	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA AGGTT-/	AGA- TI	TTGTATACT	TTTATTTGA-	- GAATTTGATC	AAGT-TT	ATATT-TGTA	CGGTAT-TAT
RS_21	TTTACGAG-T-TTC	ACTTATTTGT	CCGA	AAA A <mark>GGTT-</mark>	FAA- TI	TTGTATACT	TTTATATGA-	GAATTTGATC	AAGT-TT	ATATT-TGTA	CGTTAT-TAT
RS_21b	TTTACGAG-T-TTC	AGTTATTTGT	CCGA	AAA AGGTT-O	GAA- TI	TTGTATACT	TTTATATGA-	- GAATTTGATC	AAGT-TT	ATATT-TTTA	CGGTAT-TAT
RI_30B RI 30	A-TGACTGGT AGAI A-AGACTGAT AGAI	ATAAATT ATA'	T-TGAAT T	TTTGTTTATC	TTAAAA TTCAAA	AATAT GCA	A-TAAGT TCA	AAAGGGGG -TA	GTACAAA CGT GTACACG GTT		
RI_30B RI_30 RI_37B	A-TGACTGGT AGA A-AGACTGAT AGA A-AGACTGAT AGA	ATAAATT ATA' ATAAATT ATA' -TAAATT ATA'	T-TGAAT T T-TGAAA T T-TGAAT T	TTTGTTTATC TTTTTTTATC TTGTTTTATA	TTAAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA	A-TAAGT TCA A-AATGT TCA A-AAAGT TCA	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA	GTACAAA CGT GTACACG GTT GTACATG GTT		
RI_30B RI_30 RI_37B RI_19B	A-TGACTGGT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-AGACTGAT AGA	ATAAATT ATA' ATAAATT ATA' -TAAATT ATA' -TAAATT ATA'	T-TGAAT T T-TGAAA T T-TGAAT T T-TGAAT T	FTTGTTTATC FTTTTTTATC FTGTTTTATA FTGTTTTATA	TTAAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA	A-TAAGT TCA A-AATGT TCA A-AAAGT TCA A-AAAGT TCA	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA	GTACAAA CGT GTACACG GTT GTACATG GTT GTACATG GTT		
RI_30B RI_30 RI_37B RI_19B RI_37	A-TGACTGGT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCA	ATAAATT ATA ATAAATT ATA -TAAATT ATA -TAAATT ATA ATAAAAT ATA	T-TGAAT T-TGAAA T-TGAAT T-TGAAT T-TGAAT	FTTGTTTATC FTTTTTTATC FTGTTTTATA FTGTTTTATA FTGTTTTATA	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA. AAAAA ACA. ATTAT GCA. ATTAT GCA. AATAT TCC.	A-TAAGT TCA A-AATGT TCA A-AAAGT TCA A-AAAGT TCA A-AAAGT TCA	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGA -TA	GTACAAA CGT GTACACG GTT GTACATG GTT GTACATG GTT GTACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_37 RI_19	A-TGACTGGT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCA A-ATATCAAT GCA	ATAAATT ATA ATAAATT ATA -TAAATT ATA -TAAATT ATA ATAAAAT ATA ATAAAAT ATA	T-TGAAT T-TGAAA T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTGTTTATC TTTTTTTATC TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTTTTTATA	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC.	A-TAAGT TCA A-AATGT TCA A-AAAGT TCA A-AAAGT TCA A-AAAGT TCA A-AAAGT TCA	AAAGGGGG -TAA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TAA AAAGGGTA -TA AAAGGGTA -TA	GTACAAA CGT GTACACG GTT GTACATG GTT GTACATG GTT GTACACG CTT GTACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_37 RI_19 RI_17B	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGATTAAT AGAI	ATAAATT ATA' TAAATT ATA' TAAATT ATA' TAAATT ATA' ATAAAAT ATA' ATAAAAT ATA' ATAAAAT TAA'	T-TGAAT T-TGAAA T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTTTTTATA TTTGTTTATA	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT TCC.	A-TAAGT TC/ A-AATGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGTA -TA AAAGGGTA -TA AAAGGGGG -TA	GTACAAA CGT GTACACG GTT GTACATG GTT GTACATG GTT GTACACG CTT GTACACG CTT GTACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_37 RI_19 RI_17B RI_17B RI_17	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGATTAAT AGAI ACAGACTGAT AGA	ATAAATT ATA' ATAAATT ATA' -TAAATT ATA' -TAAATT ATA' ATAAAAT ATA' ATAAAAT ATA' ATAAATT ATA' -TAAATT ATA'	T-TGAAT T T-TGAAT T T-TGAAT T T-TGAAT T T-TGAAT T T-TGAAT T TATGAAT T	TTTGTTTATC TTGTTTTATA TGTTTTATA TTTTTTTATA TTTTTTTATA TTTGTTTATC TTGTTTTATC	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC. AATAT TCC. AATAT GCT. ATTAT GCT.	A-TAAGT TC/ A-AAGT TC/ A-AAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ AGAAAGT TC/ AGAAAGT TC/	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGTA -TA AAAGGGGG -TA AAAGGGGG -TA	GTACAAA CGT GTACACG GTT GTACATG GTT GTACACG CTT GTACACG CTT GTACACG CTT GTACACC CTT		
RI_30B RI_30 RI_37B RI_19B RI_17 RI_17B RI_17 RI_50 PT 49	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGATTAAT AGAI ACAGACTGAT AGAI A-AGACTAAT AGAI	ATAAATT ATA' ATAAATT ATA' -TAAATT ATA' -TAAATT ATA' ATAAAAT ATA' ATAAAAT ATA' ATAAATT TAA' -TAAATT ATA' ATAAAATT ATA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT TATGAAT T-TGAAT	FTTGTTTATC TTGTTTTATA TTGTTTTATA FTTTTTTATA FTTGTTTATA FTTGTTTATC FTGTTTTATC FTTGTTTTATC	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA. AAAAA ACA. ATTAT GCA. AATTAT GCA. AATAT TCC. AATAT TCC. AATAT GCT. AATAT GCA.	A-TAAGT TCA A-AATGT TCZ A-AAAGT TCZ A-AAAGT TCZ A-AAAGT TCZ A-AAAGT TCZ AGAAAGT TCZ AGAAAGT TCZ A-AAATT TCZ	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGT -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA	STACAAA CGT STACACG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_37 RI_19 RI_17 RI_17 RI_17 RI_50 RI_49 RT_10	A-TGACTGGT AGAI A-AGACTGAT AGA: A-AGACTGAT AGA: A-AGACTGAT AGA: A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGATTAAT AGAI ACAGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT ATA' -TAAATT ATA' -TAAATT ATA' -TAAATT ATA' ATAAAAT ATA' ATAAAAT ATA' -TAAATT ATA' ATAAATT ATA' ATAAATT ATA' -TAAATT ATA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT TATGAAT T-TGAAT T-TGAAT	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTATA TTTTTTTATA TTTGTTTATC TTGTTTATC TTGTTTATC TTTGTTTATC	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTTAAA	AATAT GCA. AAAAA ACA ATTAT GCA. ATTAT GCA. AATAT TCC. AATAT TCC. AATAT GTT. AATAT GCA. AATAT GCA. AATAT GCA.	A-TAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ AGAAAGT TC/ A-AAAGT TC/ A-AAATT TC/ A-AAATT TC/ A-AAAGT GC/	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGTA -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGTG -TA AAAGGGTG -TA	STACAAA CGT STACACG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_37 RI_19 RI_17 RI_17 RI_50 RI_49 RI_10 RI_11	A-TGACTGGT AGAI A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           -TAAATT         ATA'           ATAAATT         ATA'           -TAAATT         ATA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	ITTGTTTATC ITGTTTTATA ITGTTTTATA ITTTTTTATA ITTGTTTATA ITTGTTTATC ITGTTTATC ITTGTTTATC ITTGTTTATC ITTGTTTATC	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTTCAAA TTTCAAA TTCAAA	AATAT GCA. AAAAA ACA ATTAT GCA. AATAT TCC. AATAT TCC. AATAT GTT. AATAT GCA. AATAT GCA. AATAT GCA. AATAT GCA.	A-TAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ AAAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/	AAAGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGG -TA AAAGGGTA -TA AAAGGGTA -TA AAAGGGG -TA AAAGGGG -TA AAAGGGGG -TA AAAGGGG -TA	STACAAA CGT STACACG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_19B RI_17 RI_17 RI_17 RI_17 RI_50 RI_49 RI_10 RI_11 RI_46	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGA: A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTTAAT AGAI	ATAAAATT         ATA'           ATAAAATT         ATA'           -TAAAATT         ATA'           -TAAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTTTTTATA TTTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTTGTTTATC TTTGTTTATC TTTGTTTATC	ТТАААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC AATAT TCC AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT TTA	A-TAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT GC2 A-AAAGT TC2	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGT -TA AAAGGGTA -TA AAAGGGG -TA AAAGGGGG -TA AAAGGGG -TA AAAGGGGG -TA AAAGGGGG -TA	STACAAA CGT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG CTT STACACG CTT STACACG CAT		
RI_30B RI_30 RI_37B RI_19B RI_19B RI_19B RI_19B RI_17 RI_50 RI_49 RI_10 RI_11 RI_46 RI_14	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTAAT AGAI A-AGACTAAT AGAI A-AGATTAAT GGAI	ATAAAATT         ATA'           ATAAAATT         ATA'           -TAAAATT         ATA'           -TAAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTATA TTTTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TATTATAATC	ТТАААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC. AATAT TCC. AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT TTA AATAT TCA	A-TAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/ A-AAATT TC/ A-AAATT TC/ A-AAAGT GC/ A-AAAGT TC/ A-AAAGT TC/ A-AAAGT TC/	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGG -TA AAAGGGTA -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA	STACAAA CGT STACACG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG CAT STACACG CAT STACACA CTT		
RI_30B RI_30 RI_37B RI_19B RI_17 RI_19 RI_17 RI_50 RI_49 RI_10 RI_49 RI_10 RI_14 RI_44 RI_44	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGA- A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGATTAAT GGAI A-AGATTAAT GGAI A-AAATTAAC AGAI	ATAAATT         ATA           ATAAAATT         ATA           -TAAAATT         ATA           -TAAAATT         ATA           ATAAAAT         ATA           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         AAA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTTTTTATC TTGTTTTATA TTGTTTTATA TTTTTTTATA TTGTTTTATA TTGTTTATC TTGTTTATC TTTGTTTATC TTTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC	ТТАААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC. AATAT TCC. AATAT GCA. AATAT GCA. AATAT GCA. AATAT GCA. AATAT TCA. AATAT TCA.	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG -TA4 AAAGGGGG -TA4 AAAGGGGG -TA4 AAAGGGGT -TA4 AAAGGGGT -TA4 AAAGGGGG -TA4 AAAGGGGG -TA4 AAAGGGGG -TA4 AAAGGGGG -TA4 AAAGGGGG -TA4 AAAGGGGG -TA4 AAAGGGGG -TA4	STACAAA CGT STACACG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CAT STACACG CAT STACACA CTT		
RI_30B RI_30 RI_37B RI_19B RI_17 RI_17B RI_17B RI_17D RI_49 RI_10 RI_11 RI_46 RI_44 RI_44 RI_44 RI_42	A-TGACTGGT AGAI A-AGACTGAT AGA: A-AGACTGAT AGA: A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGATTAAT AGAI A-AGACTGAT AGA: A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGATTAAT AGAI A-AGATTAAT AGAI A-AGATTAAT GGAI A-ATATCAAT TCAI	ATAAATT         ATA           ATAAAATT         ATA'           -TAAAATT         ATA'           -TAAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         TAA'           ATAAAATT         TAA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           -TAAAATT         ATA'           -TAAAATT         ATA'           -TAAAATT         ATA'           ATAAAATT         GTA'           ATAAAATT         AA'	T-TGAAT T-TGAA	FTTGTTTATC TTGTTTTATA TTGTTTTATA FTTGTTTATA TTGTTTATA TTGTTTATC TTGTTAC TTGTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTAC TTGTTCAC TTGTTCAC TTGTTTAC TTGTTTAC TTGTTCAC TTGTTCAC TTGTTCAC TTGTTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGTCAC TTGCC	ТТАААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД ТТСААД	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT TCA AATAT TCA	A-TAAGT TC2 A-AATGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAATT TC2 A-AAATT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2 A-AAAGT TC2	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGTA -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA	STACAAA CGT STACACG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_19 RI_17 RI_50 RI_17 RI_50 RI_49 RI_10 RI_11 RI_46 RI_44 RI_44 RI_44 RI_26 RS_16 PE_6	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGATTAAT GGAI A-AGATTAAT GGAI A-AATATCAAT TCAI A-AGACTGAT AGAI	ATAAATT         ATA'           ATAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'           ATAAATT         ATA'           ATAAATT         TAA'           ATAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'           ATAAATT         ATA'           ATAAATT         ATA'           ATAAAATT         ATA'	T-TGAAT T-TGAT	FTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTATA TTGTTTTATA TTGTTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTAC TTGTTTAC TTGTTTAC TTGTTTAC	ТТАААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT GTT. AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT TCA AATAT TCA AATAT TCA AATAT CCA	A-TAAGT TC2 A-AATGT TC2 A-AAAGT TC2	AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGTA -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA	STACAAA CGT STACACG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CAT STACACA CTT STACACG CTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19 RI_17B RI_17B RI_17 RI_50 RI_49 RI_10 RI_11 RI_46 RI_44 RI_44 RI_26 RS_16 RS_16 RS_9 PC 1	A-TGACTGGT AGAI A-AGACTGAT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT GGA A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGATTAAT GGAI A-AGATTAAT GGAI A-AATTAAC AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA           ATAAATT         ATA           -TAAATT         ATA           -TAAATT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'           ATAAAATT         TA'           ATAAATT         TA'           ATAAAATT         AA'           ATAAAATT         AA'           ATAAATT         AA'           ATAAATT         AA'           ATAAATT         ATA'	T-TGAAT ( T-TGAAT ( T-AGAAT (	ITTGTTTATA TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTGTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTAGTTTATC TTAGTTTATC TTAGTTTATC TTAGTTTATA TTATTTATC	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC AATAT TCC AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT TCC AATAT CCA AATAT CCA AATAT GCA	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG -TA AAAGGGGG -TA AAAGGGGG -TA AAAGGGGT -TA AAAGGGTA -TA AAAGGGGT -TA AAAGGGGG -TA	STACAAA CGT 3TACACG GTT 3TACACG GTT 3TACACG CTT 3TACACG GTT 3TACACG GTT 3TACACG GTT 3TACACG GTT		
RI_30B RI_30 RI_37B RI_19B RI_17B RI_17B RI_17B RI_17 RI_50 RI_49 RI_10 RI_41 RI_46 RI_41 RI_46 RI_41 RI_46 RS_16 RS_9 RS_1 RS_37	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGATTAAT GGAI A-AGATTAAT GGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA'           ATAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAATT         AAA'           ATAAATT         AAA'           ATAAATT         AAA'           ATAAATT         ATA'	T-TGAAT ( T-TGAAT ( T-AGAAT (	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTTTTTATA TTTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TATATAATC TTAGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC	ТТАААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР АТСААР АТСААР	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA. AATAT TCA. AATAT CCA. AATAT GCA AATAT GCA.	A-TAAGT TC2 A-AAAGT TC2	AAAGGGGG         -TAi           AAAGGGGGG         -TAi           AAAGGGGGG         -TAi           AAAGGGGGG         -TAi           AAAGGGGGG<	STACAAA CGT STACATG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT		
RI_30B RI_30 RI_37B RI_19B RI_17B RI_17B RI_17B RI_17B RI_10 RI_10 RI_10 RI_11 RI_46 RI_4 RI_44 RI_44 RI_44 RI_44 RS_16 RS_16 RS_9 RS_11 RS_37 RS_11	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGATTAAT AGAI A-AGATTAAT AGAI A-AAATTAAC AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAAATT         ATAA           -TAAAATT         ATA           -TAAAATT         ATA           -TAAAATT         ATA           -TAAAATT         ATA           ATAAAAAT         ATA           ATAAAAT         ATA           ATAAAATT         TAA           ATAAAATT         ATA           ATAAAATT         ATA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-AGAAT T-TGAAT T-TGAAT	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTATA TTTTTTATA TTGTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTTTTTATC TTATTTATC TTATTTATC TTATTTAT	ТТАААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР ТТСААР АТСААР АТСААР АТСААР	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC. AATAT TCC. AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT TCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGG         -TAi	STACAAA CGT STACACG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CAT STACACG CAT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_17 RI_17 RI_17B RI_17 RI_170 RI_10 RI_10 RI_11 RI_46 RI_46 RI_46 RI_46 RS_9 RS_1 RS_17 RS_17a	A-TGACTGGT AGAI A-AGACTGAT AGA: A-AGACTGAT AGA: A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGATTAAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA           ATAAAATT         ATA           -TAAAATT         ATA           -TAAAATT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAATT         ATA           ATAAAATT         ATA           ATAAAATT         ATA           ATAAAATT         ATA           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         AAA'           ATAAAATT         ATA'           ATAAAATT         ATA' </th <th>T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-AGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT</th> <th>TTTTTTTATC TTTTTTTATA TTTTTTTATA TTTTTTTATA TTTTTT</th> <th>ТТАААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА АТСААА АТСААА ТТСААА ТТСААА</th> <th>AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT TCA AATAT CCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA</th> <th>A-TAAGT TC/ A-AAAGT TC/</th> <th>AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGG         -TAi           AAGGGGGG         -TAi           AAGGGGGG         -TAi           AAGGGGGG         -TAi           AAGGGGGG</th> <th>STACAAA CGT STACACG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CAT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG TT STACACG TT STACACG TT STACACG TT STACACG TT</th> <th></th> <th></th>	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-AGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTTTTTATC TTTTTTTATA TTTTTTTATA TTTTTTTATA TTTTTT	ТТАААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА ТТСААА АТСААА АТСААА ТТСААА ТТСААА	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT TCA AATAT CCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGG         -TAi           AAGGGGGG         -TAi           AAGGGGGG         -TAi           AAGGGGGG         -TAi           AAGGGGGG	STACAAA CGT STACACG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CAT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG TT STACACG TT STACACG TT STACACG TT STACACG TT		
RI_30B RI_30 RI_37B RI_19B RI_17B RI_17B RI_17 RI_50 RI_49 RI_10 RI_11 RI_46 RI_4 RI_44 RI_46 RS_16 RS_16 RS_16 RS_16 RS_17 RS_171	A-TGACTGGT AGAI A-AGACTGAT AGA: A-AGACTGAT AGA: A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAAATT         ATA'           ATAAAATT         ATA'           -TAAAATT         ATA'           -TAAAATT         ATA'           -TAAAATT         ATA'           ATAAAAT         ATA'           ATAAAATT	T-TGAAT [ T-TGAAT ] T-TGAAT ] T-AGAAT ] T-AGAAT ] T-TGAAT ] T-TGAAT ] T-TGAAT ] T-TGAAT ] T-TGAAT ] T-TGAAT ] T-TGAAT ] T-TGAAT ]	TTTTTTATC TTTTTTTATA TTTTTTATA TTTTTTATA TTTTTTATA TTTTTT	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC. AATAT TCC. AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT TCA. AATAT TCA. AATAT TCA. AATAT GCA. AATAT GCA. AATAT GCA. AATAT GCA. AATAT GCA. AATAT GCA.	A-TAAGT TC2 A-AAAGT TC2	AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGA -TAA AAAGGGGA -TAA AAAGGGGG -TAA	STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG TTT STACACG TTT STACACG TTT STACACG TTT STACACG TTT		
RI_30B RI_30 RI_37B RI_19B RI_17B RI_17B RI_17 RI_50 RI_10 RI_10 RI_10 RI_10 RI_11 RI_46 RS_16 RS_16 RS_16 RS_16 RS_16 RS_17 RS_11 RS_37 RS_11 RS_175 RS_175	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT GCAI A-AGACTGAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA'           ATAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAATT         ATA'           ATAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'      ATAAATT         ATA' <t< th=""><th>T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-TGAAT T-GAAT</th><th>ITTGTTTATA TTGTTTTATA TTGTTTTATA TTTGTTTATA TTTGTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATA TAGTTATA TAGTTTATC TTATTTATC TTATTTATC TTATTTATC TTATTTAT</th><th>TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA</th><th>AATAT         GCA.           AAAAA         ACA.           AATAT         GCA.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         GCA.           AATAT         TCA.           AATAT         GCA.           AATAT         GCA.  </th><th>A-TAAGT TC/ A-AAAGT TC/</th><th>AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGT -TAA AAAGGGTA -TAA AAAGGGGT -TAA AAAGGGGG -TAA</th><th>STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG TTT STACACG STT</th><th></th><th></th></t<>	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-TGAAT T-GAAT	ITTGTTTATA TTGTTTTATA TTGTTTTATA TTTGTTTATA TTTGTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATA TAGTTATA TAGTTTATC TTATTTATC TTATTTATC TTATTTATC TTATTTAT	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT         GCA.           AAAAA         ACA.           AATAT         GCA.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         GCA.           AATAT         TCA.           AATAT         GCA.	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGT -TAA AAAGGGTA -TAA AAAGGGGT -TAA AAAGGGGG -TAA	STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG TTT STACACG STT		
RI_30B RI_30 RI_37B RI_37B RI_19 RI_17B RI_17B RI_17 RI_50 RI_49 RI_10 RI_11 RI_46 RI_44 RI_46 RI_44 RI_46 RS_16 RS_9 RS_11 RS_37 RS_171 RS_172 RS_172 RS_350 RS_350 RS_350 RS_350	A-TGACTGGT AGAI A-AGACTGAT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT GGA A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA           ATAAATT         ATA           -TAAATT         ATA           -TAAATT         ATA           ATAAAAT         ATA           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-AGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTGTTTATA TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTTTTTATA TTTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTATTTATC TTATTTATC TTATTTATC TTATTTAT	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA ATCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AATAT GCA ATTAT GCA ATTAT GCA AATAT TCC AATAT TCC AATAT TCC AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT TCA AATAT TCA AATAT TCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT GCA	A-TAAGT TC2 A-AAAGT CC2 A-AAAGT CC2 A-AAAAGT CC2 A-A	AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGT -TAA AAAGGGTA -TAA AAAGGGGT -TAA AAAGGGGG -TAA	STACAAA CGT STACATG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG TTT STACACG TTT STACACG TTT STACACG TTT STACACG TTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_17B RI_17B RI_17B RI_17 RI_50 RI_17 RI_49 RI_10 RI_11 RI_46 RI_44 RI_46 RI_44 RI_46 RS_16 RS_16 RS_17 RS_17 RS_172 RS_172 RS_172 RS_172 RS_35b RS_737 RS_35b	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAAATT         ATAA           ATAAAATT         ATA           -TAAAATT         ATA           -TAAAATT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAATT         ATA           ATAAATT         ATA	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-TGAA	TTTGTTTATA TTGTTTTATA TTGTTTTATA TTTTTTATA TTTTTTATA TTTTTTATA TTTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTATTTATC TTATTTATC TTATTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT TCC. AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT TCA. AATAT TCA. AATAT TCA. AATAT GCA AATAT GCA	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG -TAA AAAGGGGG -TAA	STACAAA CGT STACATG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG TTT STACACG TTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_17 RI_17 RI_17 RI_17 RI_17 RI_17 RI_10 RI_10 RI_10 RI_10 RI_10 RI_11 RI_46 RI_46 RS_16 RS_17 RS_17 RS_17 RS_17 RS_177 RS_135 RS_135 RS_137 RS_135 RS_137 RS_135 RS_155 RS_155 R	A-TGACTGGT AGAI A-AGACTGAT AGA: A-AGACTGAT AGA: A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAAATT         ATAA           -TAAAATT         ATA           -TAAAATT         ATA           -TAAAATT         ATA           -TAAAATT         ATA           ATAAAAAT         ATA           ATAAAATT         TAA           ATAAAATT         ATA           ATAAATT         ATA           ATAAAATT         ATA	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-AGAAT T-AGAAT T-AGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTGTTTTATA TTTTTTATA TTGTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTTTTTATC TTTTTTATC TTATTTATC TTATTTAT	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA AATAT GCA AATAT TCC. AATAT TCC. AATAT GCA AATAT GCA	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGA -TAA AAAGGGGG -TAA	STACAAA CGT STACACG GTT STACATG GTT STACATG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CAT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG TTT STACACG TTT STACACG TTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_17B RI_17B RI_17 RI_50 RI_17 RI_10 RI_11 RI_46 RI_44 RI_44 RI_46 RS_16 RS_16 RS_16 RS_16 RS_17 RS_111 RS_178 RS_175 RS_175 RS_18b RS_T11 RS_118 RS_	A-TGACTGGT AGAI A-AGACTGAT AGA: A-AGACTGAT AGA: A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGAI A-AGACTGAT GGAI A-AGACTGAT GGAI	ATAAATT         ATA'           ATAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'	T-TGAAT [ T-TGAAT ] T-TGAAT ]	TTTTTTTATC TTTTTTTATA TTTTTTTATA TTTTTTTATA TTTTTT	TTAAAA TTCAAA	AATAT GCA AAAAA ACA ATTAT GCA ATTAT GCA AATAT TCC. AATAT TCC. AATAT GCT. AATAT GCA AATAT GCA AATAT GCA AATAT GCA AATAT TCA AATAT TCA AATAT CCA AATAT GCA AATAT GCA	A-TAAGT TC/ A-AATGT TC/ A-AAAGT TC/	AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGG         -TAi           AAAGGGGGG         -TAi           AAAGGGGGG <th>STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG CTT</th> <th></th> <th></th>	STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_17B RI_17B RI_17 RI_50 RI_10 RI_10 RI_11 RI_44 RI_44 RI_44 RI_44 RI_44 RI_46 RS_16 RS_16 RS_16 RS_16 RS_17 RS_112 RS_17 RS_111 RS_178 RS_111 RS_178 RS_111 RS_118 RS_111 RS_118 RS_111 RS_110 RS_511 RS_511	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT GGAI A-AGACTGAT AGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT AGAI A-AGACTGAT GGAI	ATAAATT         ATA'           ATAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'	T-TGAAT [ T-TGAAT ] T-TGAAT ]	ITTGTTTATA TTGTTTTATA TTGTTTTATA TTTGTTTATA TTTGTTTATA TTGTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTATTTATC TTATTTATC TTATTTATC TTATTTAT	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT         GCA.           AAAAA         ACA.           AATAT         GCA.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         GCA.           AATAT         GCA. </th <th>A-TAAGT TC/ A-AAAGT TC/</th> <th>AAAGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGT -TAA AAAGGGTA -TAA AAAGGGGT -TAA AAAGGGGG -TAA</th> <th>STACAAA CGT JTACACG GTT JTACATG GTT JTACATG GTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG GTT JTACACG CTT JTACACG CTT</th> <th></th> <th></th>	A-TAAGT TC/ A-AAAGT TC/	AAAGGGG -TAA AAAGGGGG -TAA AAAGGGGG -TAA AAAGGGGT -TAA AAAGGGTA -TAA AAAGGGGT -TAA AAAGGGGG -TAA	STACAAA CGT JTACACG GTT JTACATG GTT JTACATG GTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG CTT JTACACG GTT JTACACG CTT JTACACG CTT		
RI_30B RI_30 RI_37B RI_37B RI_17B RI_17B RI_17 RI_50 RI_49 RI_10 RI_11 RI_46 RI_44 RI_46 RS_16 RS_16 RS_16 RS_16 RS_17 RS_17a RS_17b RS_35a RS_35b RS_177 RS_18b RS_711 RS_10 RS_110 RS_34	A-TGACTGGT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-AGACTGAT AGAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT GGAI A-AGACTGAT AGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT GGAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA           ATAAATT         ATA           -TAAATT         ATA           -TAAATT         ATA           -TAAATT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAAAT         ATA           ATAAATT         ATA           ATAAAATT         ATA           ATAAATT         ATA           ATAAAATT         ATA           ATAAAATT         ATA           ATAAA	T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-TGAAT T-GAAT T-TGAAT	ITTGTTTATA TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTGTTTATA TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTAGTTATA TATTTATC TTAGTTATA TTATTTATC TTGTTTATC TTGTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTGTTTATC	TTAAAA TTCAAA	AATAT GCA AATAT GCA ATTAT GCA AATAT TCC AATAT TCC AATAT TCC AATAT TCC AATAT GCA AATAT GCA	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG         -TAI           AAAGGGGG         -TAI           AAAGGGGG         -TAI           AAAGGGGA         -TAI           AAAGGGGA         -TAI           AAAGGGGA         -TAI           AAAGGGGA         -TAI           AAAGGGGA         -TAI           AAAGGGGG	STACAAA CGT STACAAG GTT STACAAG GTT STACAAG GTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG TTT STACACG TTT STACACG TTT STACACG TTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_37B RI_19 RI_17B RI_17B RI_17 RI_50 RI_49 RI_10 RI_41 RI_46 RI_44 RI_46 RS_16 RS_9 RS_11 RS_376 RS_17a RS_17a RS_35a RS_35b RS_777 RS_18b RS_771 RS_100 RS_51 RS_34 RS_21	A-TGACTGGT AGAI A-AGACTGAT AGA A-AGACTGAT AGA A-AGACTGAT AGA A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT GGAI A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATAA           ATAAATT         ATA           -TAAATT         ATA           -TAAATT         ATA           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'	T-TGAAT [ T-TGAAT ] T-TGAAT ]	TTTGTTTATA TTGTTTTATA TTGTTTTATA TTTTTTTATA TTTTTTTATA TTTTTTTATA TTTTTT	TTAAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA ATCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA TTCAAA	AATAT         GCA.           AAAAA         ACA.           AATAT         GCA.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         GCA.           AATAT         TCA.           AATAT         GCA.           AATAT         GCA. </th <th>A-TAAGT TC/ A-AAAGT TC/</th> <th>AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGTA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGG         -TAi           AAAGGGGGG<th>STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG CTT</th><th></th><th></th></th>	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGTA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGA         -TAi           AAAGGGGG         -TAi           AAAGGGGGG <th>STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG CTT</th> <th></th> <th></th>	STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG CTT		
RI_30B RI_30 RI_37B RI_19B RI_17 RI_17 RI_50 RI_17 RI_17 RI_50 RI_10 RI_10 RI_10 RI_10 RI_10 RI_10 RI_10 RI_11 RI_46 RI_44 RI_44 RI_44 RI_46 RS_17 RS_	A-TGACTGGT AGAI A-AGACTGAT AGA- A-AGACTGAT AGA- A-AGACTGAT AGA- A-ATATCAAT GCAI A-ATATCAAT GCAI A-ATATCAAT GCAI A-AGACTGAT AGA- A-AGACTGAT AGA- A-AGACTGAT AGAI A-AGACTGAT AGAI	ATAAATT         ATA'           ATAAATT         ATA'           -TAAATT         ATA'           -TAAATT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAAT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'           ATAAAATT         ATA'	T-TGAAT T-TGAA	TTTGTTTATC TTGTTTTATA TTGTTTTATA TTTTTTATA TTTTTTATA TTTTTTATA TTTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTTTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC TTGTTTATC	TTAAAA TTCAAA	AATAT         GCA.           AAAAA         ACA.           AATAT         GCA.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         TCC.           AATAT         GCA.           AATAT         GCA.           AATAT         GCA.           AATAT         GCA.           AATAT         GCA.           AATAT         GCA.           AATAT         TCA.           AATAT         GCA.           AATAT         GCA. </th <th>A-TAAGT TC/ A-AAAGT TC/</th> <th>AAAGGGGG         -TAi           AAAGGGGG         -TAi           AAAGGGGG</th> <th>STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG TTT STACACG TTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG TTT STACACG TTT STACA</th> <th></th> <th></th>	A-TAAGT TC/ A-AAAGT TC/	AAAGGGGG         -TAi           AAAGGGGG	STACAAA CGT STACACG GTT STACACG GTT STACACG CTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG GTT STACACG CTT STACACG TTT STACACG TTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG CTT STACACG TTT STACACG TTT STACA		

**Fig. 3.** Alignment of *R. induratus* and *R. scutatus* RUSI sequences. Shaded nucleotides indicate degenerated variants of the *Arabidopsis*-like TTTAGGG telomere repetitive motif in both forward (grey) and reverse (black) senses.



**Fig. 4.** Karyotype of *R. induratus*. Note the presence of 20 chromosome pairs ordered in two ten-pair complements (A, B), corresponding to the basic chromosome number of the group, x = 10.



**Fig. 5.** Composite karyotype of *R. induratus* metaphase chromosomes (A, B) analyzed after FISH with RUSI sequences.

**Table 1.** Statistics for RUSI satellite-DNA sequences, including intraspecific and interspecific mean distances between monomeric sequences. The differences between species are stated as Fixed positions for fixed or nearly fixed polymorphisms while Transitional positions represent intermediate stages in the process of homogenization toward the fixation, according to the Strachan et al. (1985) method (Navajas-Pérez et al., 2007).

Mean di	stance	Differences between species		
Intraspecific (R. induratus/R. scutatus)		Inter- specific	Fixed (IV–V)	Transitional (II–III)
0.15	0.075	0.12	0	65

### Discussion

*R. induratus* is an endemic polygamous species of the Iberian Peninsula. This species is tetraploid (Fernández-Casas, 1977) while the widespread closely related *R. scutatus* is diploid (Rechinger, 1964). In the present study, we have confirmed a karyotype of 2n = 4x = 40 for *R. induratus*, with a basic chromosome number of x = 10 (Figs. 4 and 5). This basic chromosome number has been found to be the ancestral number in all docks and in most of the hermaphroditic and polygamous species of sorrels (De-

graeve, 1975; Navajas-Pérez et al., 2005a). Recently, we have identified that, within the clade of polygamous and hermaphroditic sorrels, there are two major subclades, one of them contains the species *R. scutatus* and *R. induratus* in the *Scutati* section (subgenus *Acetosa*) and *R. maderensis* in the *Hastati* section (*Acetosa*) (Navajas-Pérez et al., 2005a). The macaronesian endemic *R. maderensis* is diploid and a basal species of the phylogenetic clade containing these three species. Thus, the tetraploidy of *R. induratus* can be interpreted as a derived character state that evolved in this clade independently of other polyploidization events occurring in other *Rumex* lineages.

In this species we describe here for the first time a new satellite-DNA family, RUSI. This satellite-DNA is located at the subtelomeric regions of most chromosomes of the karyotype (Fig. 5). Within the distal regions of the chromosomes in most eukaryotic species we can distinguish two differentiated parts. The more distal regions represent the true telomere and, except in rare exceptions, are formed of a variable number of short tandem repeats (Blackburn and Greider, 1995). In vertebrate species, the sequence TTAGGG is conserved at telomeres while in invertebrate species-specific variations of the sequence TT-GGGG first described in *Tetrahymena* are present. In *Arabidopsis*, as well as in the vast majority of flowering plants analyzed up to date, TTTAGGG is the most com-

mon repetitive telomere sequence (Richards and Ausubel, 1988). The subterminal sequences adjacent to the telomeres have been referred to as telomere-associated sequences or TASs (Louis and Vershinin, 2005) and are usually composed of a variety of highly tandem repetitive DNA sequences or satellite DNAs. In plants, TASs are often structurally more variable and polymorphic than any other regions of the genome even in species where they were described as not highly repeated, such as in rice and maize for example (Burr et al., 1992; Ashikawa et al., 1994). These subtelomeric sequences do not necessarily participate in telomere function but, for example, can facilitate meiotic pairing or protect terminal genes against the loss and gain processes at the chromosome ends (Henderson, 1995; Kipling, 1995). They could also be involved in the regulation of subtelomeric genes (Sykorova et al., 2003). On the other hand, these sequences have been postulated to acquire vital importance when telomere sequences disappear. In fact, chromosomes with depleted telomeres have a propensity to be lost (Sandell and Zakian, 1993). In this sense, several cases of a satellite DNA adopting the telomere-repeat role have been reported, as for instance the substitution of most telomere repeats of Alliaceae by rDNA sequences, retrotransposable elements and satellite-DNA sequences (Barnes et al., 1985; Pich and Schubert, 1998; Do et al., 2001). Subtelomeric sequences are not conserved, but are often species specific with a variety of lengths and degrees of repetitiveness (Kipling, 1995; Pryde et al., 1997), even existing as different satellite-DNA families (Brown, 1989; Brown et al., 1990; Weber et al., 1990, 1991) or subfamilies of the same satellite DNA (Kazama et al., 2006) in different subsets of chromosomes within a species. A common characteristic of subtelomeric satellite DNAs is the direct proximity to telomere repeats as in tobacco (Fajkus et al., 1995), Silene (Garrido-Ramos et al., 1999), or Triticeae (Contento et al., 2005) which lead to the intermixing of subtelomeric repeats with telomere sequence that are often degenerate. Also, subtelomeric repeats often show inverted organization with head to head orientation (Vershinin et al., 1995; Contento et al., 2005). The intimate contact of RUSI repeats with the telomere is corroborated by the presence of degenerated telomere motifs intercalated with them and might represent a common feature that occurs as a consequence of dynamic processes taking place at chromosome ends.

Interestingly, the RUSI repeats are not only located at the telomeres but also at centromeric positions of some chromosomes (Fig. 5). Also interestingly, there are differences between chromosomes having subtelomeric RUSI repeats on one or both arms. These observations lead to the consideration of chromosome reorganization implying that some loss/gain processes may have occurred. This report represents an important exception to the equilocality principle of satellite-DNA distribution (Thomas, 1971). Presence of the same type of satellite-DNA sequences at both centromeric and subtelomeric sequences can be explained by chromosome reorganization such as Robertsonian translocations (Garagna et al., 2002; Castiglia et al., 2006) or by transposition (Abad and Villasante, 2000; Tek and Jiang, 2004; Bao et al., 2006). Reported cases of alteration of the equilocal distribution of satellite DNAs also included satellite-DNA amplifications in specific regions of particular chromosomes such as sex chromosomes (Cuñado et al., 2007), B chromosomes (Alfenito and Birchler, 1993) or microchromosomes (Kuhn et al., 2007) or exceptional cases of independent evolution of interstitial heterochromatic regions of one chromosome pair with respect to the rest of the heterochromatic regions of the karyotype, leading to the characteristic asymmetries found in the genus Muscari (de la Herrán et al., 2001). In the case described here, the appearance of RUSI sequences in two non-equilocal different regions of the chromosomes of R. induratus could imply chromosome reorganization such as inversions and/or transpositions, but also Robertsonian translocations.

Satellite-DNA families have been found to be irregularly distributed within the different groups of plants, including chromosome-specific ones, such as the WE35 family of Triticum aestivum (Ueng et al., 2000), speciesspecific families, as RAYSI Y-specific satellite DNA in several Rumex species (Navajas-Pérez et al., 2006) or satellite DNA borne by species from a whole tribe, such as the subtelomeric 120-bp repeat family of Triticeae (Contento et al., 2005). Assuming that two species bearing the same sequence would be more related to each other than those lacking these sequences, satellite DNA can be used as a cladistic marker by merely analysing the presence/absence status. Here, we have detected that the RUSI satellite-DNA family is exclusively distributed in two species of the genus Rumex, R. induratus and R. scutatus. Both classical and more novel systematics of the genus (López González, 2000; Navajas-Pérez et al., 2005a) describe these two species as closely related to each other as part of the section Scutati of the subgenus Acetosa. Since we have not found evidence of RUSI sequences in any other Rumex species analyzed (not shown), including those more phylogenetically related as R. maderensis, the present study demonstrates that this satellite DNA originated very recently (between 2.5 and 4.5 mya – Navajas-Pérez et al., 2005a). This is not an uncommon situation due to the high turnover rate of this part of the eukaryotic genome. Studies on satellite DNA dynamics explain loss or increase in copy number through not always well-known amplification mechanisms in a relatively short evolutionary time (Charlesworth et al., 1994).

Repetitive-DNA families are influenced by several molecular mechanisms of non-reciprocal exchange such as unequal crossing-over or gene conversion (Ohta and Dover, 1984) that can gradually spread a variant sequence throughout a family within a sexual population in a population-genetic process called molecular drive (Dover, 2002). Molecular drive leads to high homogeneity levels in a repetitive DNA family for species-diagnostic mutations, which subsequently is the origin of inter-species genetic divergence, an evolutionary pattern of repetitive sequences known as concerted evolution (Dover, 2002). Many empirical observations indicate that the rate of production of new sequence variants (mutations) in satellite-DNA sequences is a slower process than their rate of spread, while the general paucity of transition stages indicates also that the replacement is relatively fast (Ugarković and Plohl, 2002). However, this is not occurring for RUSI sequences since we detected high levels of intraspecific sequence variation (i.e., low rates of intraspecific sequence homogenization; Table 1). The intraspecific sequence variation has two sources: on one hand, a high percentage of shared polymorphisms between *R*. induratus and R. scutatus (19%), which might represent ancestral polymorphisms; on the other hand, a much higher percentage of new species-specific mutations (35% of transition stages II+III) representing an initial process of sequence divergence between the two species (mean sequence differences of 12%). However, spreading rates do not appear to operate at higher rates than mutation rates, a fact that could be explained as a consequence of the multichromosomal and multiloci distribution (i.e., chromosomal rearrangements) of RUSI sequences. Multichromosomal distribution and chromosomal rearrangements together with polyploidy (Krieger and Fuerst, 2002) might be among the possible causes for extensive intraspecific variability of repetitive sequences of various organisms. Polyploidization is also a fact to be considered here since we have found higher levels of variation in the tetraploid R. induratus than in the diploid R. scutatus.

#### Acknowledgements

We thank Dr John Bailey, University of Leicester, for checking the English of the manuscript.

#### References

- Abad JP, Villasante A: Searching for a common centromeric structural motif: *Drosophila* centromeric satellite DNAs show propensity to form telomeric-like unusual DNA structures. Genetica 109:71–75 (2000).
- Alfenito MR, Birchler JA: Molecular characterization of a maize B chromosome centric sequence. Genetics 135:589–597 (1993).
- Ashikawa I, Kurata N, Nagamura Y, Minobe Y: Cloning and mapping of telomere-associated sequences from rice. DNA Res 1:67–76 (1994).
- Bao W, Zhang W, Yang Q, Zhang Y, Han B, et al: Diversity of centromeric repeats in two closely related wild rice species, *Oryza officinalis* and *Oryza rhizomatis*. Mol Genet Genomics 275:421–430 (2006).
- Barnes SR, James AM, Jamieson G: The organisation, nucleotide sequence, and chromosomal distribution of a satellite DNA from *Allium cepa*. Chromosoma 92:185–192 (1985).
- Beridze TG: Satellite DNA (Springer, Berlin 1986).

- Blackburn EH, Greider CW: Telomeres (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1995).
- Brown WRA: Molecular cloning of human telomeres in yeast. Nature 38:774–776 (1989).
- Brown WRA, MacKinnon PJ, Villasanté A, Spurr N, Buckle VJ, et al: Structure and polymorphism of human telomere-associated DNA. Cell 63:119–132 (1990).
- Burr B, Burr FA, Matz EC, Romero-Severson J: Pinning down loose ends: mapping telomeres and factors affecting their length. Plant Cell 4:953–960 (1992).
- Castiglia R, Garagna S, Merico V, Oguge N, Corti M: Cytogenetics of a new cytotype of African *Mus* (subgenus *Nannomys*) *minutoides* (Rodentia, Muridae) from Kenya: C- and Gbanding and distribution of (TTAGGG)n telomeric sequences. Chromosome Res 14: 587–594 (2006).
- Charlesworth B, Sniegowski P, Stephan W: The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220 (1994).

- Contento A, Heslop-Harrison JS, Schwarzacher T: Diversity of a major repetitive DNA sequence in diploid and polyploid Triticeae. Cytogenet Genome Res 109:34–42 (2005).
- Cuñado N, Navajas-Pérez R, de la Herrán R, Ruiz Rejón C, Ruiz Rejón M, et al: The evolution of sex chromosomes in the genus *Rumex* (Polygonaceae): identification of a new species with heteromorphic sex chromosomes. Chromosome Res 15:825–832 (2007).
- de la Herrán R, Robles F, Cuñado N, Santos JL, Ruiz Rejón M, et al: A heterochromatic satellite DNA is highly amplified in a single chromosome of *Muscari* (Hyacinthaceae). Chromosoma 110:197–202 (2001).
- Degraeve N: Contribution à *l'étude* cytotaxonomique des *Rumex*. I. Le genre *Rumex* L. sensu stricto. Caryologia 28:187–201 (1975).
- Do GS, Seo BB, Yamamoto M, Suzuki G, Mukai Y: Identification and chromosomal location of tandemly repeated DNA sequences in *Allium cepa*. Genes Genet Syst 76:53–60 (2001).

Dover G: Molecular drive. Trends Genet 18:587–589 (2002).

- Fajkus J, Kovarík A, Královics R, Bezděk M: Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. Mol Gen Genet 247:633–638 (1995).
- Fernández-Casas J: Números cromosómicos de plantas españolas, IV. Anales Inst Bot Cavanilles 34:335–349 (1977).
- Ferreres F, Ribeiro V, Izquierdo AG, Rodrigues MA, Seabra RM, et al: *Rumex induratus* leaves: interesting dietary source of potential bioactive compounds. J Agric Food Chem 54:5782–5789 (2006).
- Garagna S, Zuccotti M, Capanna E, Redi CA: High-resolution organization of mouse telomeric and pericentromeric DNA. Cytogenet Genome Res 96:125–129 (2002).
- Garrido-Ramos MA, de la Herrán R, Ruiz Rejón M, Ruiz Rejón C: A subtelomeric satellite DNA family isolated from the genome of the dioecious plant *Silene latifolia*. Genome 42: 442–446 (1999).
- Henderson ER: Telomere DNA structure, in Blackburn EH, Greider CW (eds): Telomeres, pp 11–34 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1995).
- Heslop-Harrison JS: Comparative genome organization in plants: from sequence and markers to chromatin and chromosomes. Plant Cell 12:617-635 (2000).
- Kazama Y, Sugiyama R, Suto Y, Uchida W, Kawano S: The clustering of four subfamilies of satellite DNA at individual chromosome ends in *Silene latifolia*. Genome 49:520–530 (2006).
- Kipling D: The Telomere (Oxford University Press, New York 1995).
- Krieger J, Fuerst PA: Evidence for a slowed rate of molecular evolution in the order Acipenseriformes. Mol Biol Evol 19:891–897 (2002). Erratum in: Mol Biol Evol 19:1644 (2002).
- Kuhn GCS, Franco FF, Manfrin MH, Moreira-Filho O, Sene FM: Low rates of homogenization of the DBC-150 satellite DNA family restricted to a single pair of microchromosomes in species from the *Drosophila buzzatii* cluster. Chromosome Res 15:457–469 (2007).
- López González G: Género *Rumex* L., in Castroviejo S, Laínz M, López González G, Montserrat P, Muñoz Garmendia F, Paiva J, Villar L (eds): Flora Iberica, vol II, pp 595–634 (CSIC, Real Jardín Botánico de Madrid, Madrid 1990).

- Louis EJ, Vershinin AV: Chromosome ends: different sequences may provide conserved functions. Bioessays 27:685–697 (2005).
- Mariotti B, Navajas-Pérez R, Lozano R, Parker JS, de la Herrán R, et al: Cloning and characterization of dispersed repetitive DNA derived from microdissected sex chromosomes of *Rumex acetosa*. Genome 49:114–121 (2006).
- Moreno-Jiménez E, Gamarra R, Carpena-Ruiz RO, Millán R, Peñalosa JM, et al: Mercury bioaccumulation and phytotoxicity in two wild plant species of Almadén area. Chemosphere 63:1969–1973 (2006).
- Navajas-Pérez R, de la Herrán R, López González G, Jamilena M, Lozano R, et al: The evolution of reproductive systems and sex-determining mechanisms within *Rumex* (Polygonaceae) inferred from nuclear and chloroplastidial sequence data. Mol Biol Evol 22: 1929–1939 (2005a).
- Navajas-Pérez R, de la Herrán R, Jamilena M, Lozano R, Ruiz Rejón C, et al: Reduced rates of sequence evolution of Y-linked satellite DNA in *Rumex* (Polygonaceae). J Mol Evol 60:391– 399 (2005b).
- Navajas-Pérez R, Schwarzacher T, de la Herrán R, Ruiz Rejón C, Ruiz Rejón M, Garrido-Ramos MA: The origin and evolution of the variability in a Y-specific satellite-DNA of *Rumex acetosa* and its relatives. Gene 368: 61–71 (2006).
- Navajas-Pérez R, Rubio-Escudero C, Aznarte JL, Ruiz Rejón M, Garrido-Ramos MA: SatDNA Analyzer: a computing tool for satellite-DNA evolutionary analysis. Bioinformatics 23:767–768 (2007).
- Ohta T, Dover GA: The cohesive population genetics of molecular drive. Genetics 108:501– 521 (1984).
- Pich U, Schubert I: Terminal heterochromatin and alternative telomeric sequences in *Allium cepa*. Chromosome Res 6:315–321 (1998).
- Pryde FE, Gorham HC, Louis EJ: Chromosome ends: all the same under their caps. Curr Opin Genet Dev 7:822–828 (1997).
- Rechinger KH Jr: The genus *Rumex* L., in Tutin TG, Heywood VH, Burges NA, Valentine DH, Walters SM, Webb DA (eds): Flora Europaea, vol 1, pp 82–89 (Cambridge University Press, Cambridge 1964).
- Richards EJ, Ausubel FM: Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. Cell 53:127–136 (1988).
- Rivera D, Obón C: The ethnopharmacology of Madeira and Porto Santo Islands, a review. J Ethnopharmacology 46:73–93 (1995).

- Sandell LL, Zakian VA: Loss of a yeast telomere: arrest, recovery, and chromosome loss. Cell 75:729–739 (1993).
- Schwarzacher T, Heslop-Harrison JS: Practical in situ Hybridization. (BIOS Scientific Publishers Limited, Oxford 2000).
- Sharma S, Raina SN: Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. Cytogenet Genome Res 109:15–26 (2005).
- Strachan T, Webb D, Dover GA: Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. EMBO J 4:1701–1708 (1985).
- Sykorova E, Lim KY, Chase MW, Knapp S, Leitch IJ, et al: The absence of *Arabidopsis*-type telomeres in *Cestrum* and closely related genera *Vestia* and *Sessea* (Solanaceae): first evidence from eudicots. Plant J 34:283–291 (2003).
- Tek AL, Jiang J: The centromeric regions of potato chromosomes contain megabase-sized tandem arrays of telomere-similar sequence. Chromosoma 113:77–83 (2004).
- Thomas CA Jr: The genetic organization of chromosomes. Annu Rev Genet 5:237–256 (1971).
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG: The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882 (1997).
- Ueng PP, Hang A, Tsang H, Vega JM, Wang L, et al: Molecular analyses of a repetitive DNA sequence in wheat (*Triticum aestivum* L.). Genome 43:556–563 (2000).
- Ugarković D, Plohl M: Variation in satellite DNA profiles – causes and effects. EMBO J 21: 5955–5959 (2002).
- Vershinin AV, Schwarzacher T, Heslop-Harrison JS: The large-scale genomic organization of repetitive DNA families at the telomeres of rye chromosomes. Plant Cell 7:1823–1833 (1995).
- Weber B, Collins C, Robbins C, Magenis RE, Delaney AD, et al: Characterization and organization of DNA sequences adjacent to the human telomere associated repeat (TTAGGG)n. Nucleic Acids Res 18:3353– 3361 (1990).
- Weber B, Allen L, Magenis RE, Hayden MR. A low-copy repeat located in subtelomeric regions of 14 different human chromosomal termini. Cytogenet Cell Genet 57:179–183 (1991).