Genet. Sel. Evol. 30 (Suppl. 1) (1998) S357-S372 © Inra/Elsevier, Paris

Original article

Evolution of Frankia–Casuarinaceae interactions

Elisabeth Navarro^a, Carole Rouvier^b, Philippe Normand^{b*}, Anne Marie Domenach^b, Pascal Simonet^b, Yves Prin^c

> ^a Laboratoire de microbiologie, Centre ORSTOM in Nouméa, BP A5 Nouméa cedex, New Caledonia

 ^b Laboratoire d'écologie microbienne du sol, UMR CNRS 5557, Université Lyon I, 43, boulevard du 11 novembre 1918, 69622 Villeurbanne cedex, France

^c Laboratoire des symbioses tropicales et méditerranéennes, Campus de Baillarguet, BP 5035, 34032 Montpellier cedex 1, France

Abstract – Nonisolated Frankia strains present in the root nodules of three of the four genera of the Casuarinaceae family (namely, Casuarina, Allocasuarina and Gymnostoma) have been characterised through polymerase chain reaction/restriction fragment length polymorphism (PCR/RFLP) analyses and sequencing of their nifD-nifK intergenic spacer (IGS). Analyses of the aligned sequences were used to deduce phylogenetic relations of these genes. Strains from Casuarina and Allocasuarina were found to be in the same cluster, while strains from *Gumnostoma* were closer to Elaeagnaceae strains. The relationships between IGS subgrouping and symbiotic (host spectrum) characteristics of the nonisolated strain confirmed the differences between Casuarina/Allocasuarina and Gymnostoma symbiosis. Genetic diversity among Casuarina and Allocasuarina microsymbionts seems to be host species-dependent. In contrast, no relation could be found between Gymnostoma microsymbionts and host species. The comparison between phylogenic analyses of the host plants and their microsymbionts suggests that the most coherent evolutionary scenario would be that an early split occurred in the evolution of Casuarinaceae, resulting into two distinct lines of descent. © Inra/Elsevier, Paris

diversity / Frankia / Casuarinaceae / coevolution

Résumé – Évolution des interactions Frankia–Casuarinaceae. En utilisant des analyses PCR/RFLP et le séquençage de l'intergène *nif*D-K, des souches non isolées de *Frankia* présentes dans les nodosités de trois des quatre genres constituant la famille des Casuarinacées (*Casuarina, Allocasuarina* and *Gymnostoma*) ont été caractérisées. L'analyse des séquences alignées a permis d'établir les relations phylogénétiques entre ces souches. Les souches infectives sur *Casuarina* et *Allocasuarina* appartiennent au

^{*} Correspondence and reprints

même groupe phylogénétique, alors que les souches infectives sur Gymnostoma sont regroupées avec les souches d'Elaeagnacées. Les relations entre le groupage moléculaire et les caractéristiques symbiotiques du micro-organisme (spectre d'hôte) confirment les différences entre les symbioses impliquant Casuarina/Allocasuarina et Gymnostoma. La diversité génétique des microsymbiotes de Casuarina et Allocasuarina semble être corrélée à l'espèce de la plante hôte. Au contraire, aucune relation n'a été trouvée entre le type de microsymbiotes de Gymnostoma et l'espèce végétale. La comparaison des arbres phylogénétiques des plantes hôtes et de leurs microsymbiotes suggère qu'une séparation précoce soit survenue dans l'évolution des Casuarinacées, entraînant l'existence de deux lignées de descendants. © Inra/Elsevier, Paris

diversité / Frankia / Casuarinacées / coévolution

1. INTRODUCTION

The actinomycete *Frankia* has established a nitrogen-fixing symbiosis with a wide range of dicotyledonous plants. This symbiosis is known to occur in more than 200 species of plants belonging to eight families (Betulaceae, Casuarinaceae, Myricaceae, Elaeagnaceae, Rhamnaceae, Rosaceae, Coriariaceae and Datiscaceae) (Benson and Silvester, 1993).

The Casuarinaceae family is composed of the four genera of tropical dicotyledonous plants Allocasuarina, Casuarina, Ceuthostoma and Gymnostoma, of which Gymnostoma is considered the most primitive (Johnson and Wilson, 1989). These plants are naturally confined to the Malaysian-Australian Melanesian region but some species, and particularly Casuarina equisetifolia, have been exported extensively to other tropical areas worldwide, to be used as windbreaks, to stabilise sand dunes or as a source of fuel wood (Diem et al., 1988; Diem and Dommergues, 1990). This is due in part to the nitrogen-fixing symbiosis that most of the 96 extant species from this family have established with the actinomycete Frankia permitting the plants to develop on poor soils.

No study has been carried out on the evolution of *Frankia*-Casuarinaceae relationships. Most of the genetic diversity work on Casuarina spe. and Allocasuarina spp. growing in areas where they are not native (Nazaret et al., 1991; Rouvier et al., 1992). Little is known about *Gymnostoma* microsymbionts, the only reports in the literature dealing with three successful isolations of *Frankia* strains in pure culture (Racette and Torrey, 1989; Savouré and Lim, 1991). Therefore, the phylogenetic relationships of *Frankia* strains infective on *Casuarina*, Allocasuarina and *Gymnostoma* genera from native areas have not been studied.

With this in mind, we studied the diversity of Casuarinaceae microsymbionts in northeastern Australia and New Caledonia, areas in the natural geographic range of the host plants. Using sequencing and polymerase chain reaction/restriction fragment length polymorphism (PCR/RFLP) analysis, we sought to compare Casuarinaceae microsymbionts and to determine the level of diversity among these strains and their relationships with host plant species.

2. MATERIALS AND METHODS

2.1. Nodules and bacterial strains

Nodules and reference strains used are described in table I.

2.2. DNA extraction from nodules

After peeling off the superficial layers, nodule lobes were disinfected with 30 % w/v H₂O₂ for 5 min, rinsed with sterile distilled water and kept at -20 °C. One nodule lobe was crushed in 500 µL of TCP buffer (100 mM Tris-HCl, pH 8, 1.4 M NaCl, 20 mM EDTA, 2 % w/v CTAB [Sigma, St Louis, MO, USA] and 3 % w/v PVPP [Sigma], pH 8). The mixture was incubated at 65 °C for 1 h and centrifuged at 3 000 g for 5 min (20 °C). The supernatant was chloroform-extracted and ethanol-precipitated. The DNA pellet was dissolved in 10 µL of TE buffer (pH 7.5).

2.3. PCR amplification of nifD-nifK intergene

For deoxyribonucleic acid (DNA) amplification of a region including the 3'end of nifD, the intergenic spacer (IGS), and the beginning of nifK, primers FGPD807 (5'-CACTGCTACCGGTCGATGAA-3') (Jamann et al., 1993) and FGPK333' (5'-CCGGGCGAAGTGGCT-3') (Nalin et al., 1995) were used. PCR amplification was performed in 0.5 mL Eppendorf tubes in a total volume of 50 μ L containing: template DNA (approximately 0.1 μ g), polymerase reaction buffer (10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 50 mM KCl, 0.01 % [w/v] gelatine, 20 μ M deoxynucleoside triphosphate [dNTP], 1 μ M each of the primers and 2.5 units of TaqI DNA polymerase [Gibco BRL, Gaithersburg, MD, USA]). DNA amplification was done in a thermocycler (Perkin Elmer, Norwalk, CT, USA) using the following programme: initial denaturation for 3 min at 95 °C, 35 cycles of denaturation (30 s at 95 °C), annealing (30 s at 63 °C) and extension (30 s at 72 °C), and a final extension (2 min at 72 °C). PCR amplification of DNA was checked by agarose gel electrophoresis (2 % w/v) in TBE buffer with 5 μ L of PCR product. The gel was stained in an aqueous solution of $1 \text{ mg} \text{ L}^{-1}$ ethidium bromide and photographed with HP5 film with a 302-nm ultraviolet source.

2.4. PCR amplification of 16S-23S intergene

Amplifications of a part of the 16S gene and the IGS were performed by using the standard conditions as described previously. Primers FGPS989e (5'-GGGGTCCTTAGGGGCT-3') (Bosco et al., 1992) and FGPL1973' (5'-ATCGGCTCGAGGTGCCAAGGGTC-3') (Navarro et al., 1992) were used for *Gymnostoma* DNA amplifications. Primers FGPS989ac (5'-GGGGTCCGT-AAGGGTC-3') (Bosco et al., 1992) and FGPL132' (5'-CCGGGTTTCCCATT-CGG-3') (Ponsonnet and Nesme, 1994) were used for *Casuarina/Allocasuarina* DNA amplifications.

	ence		(1997)						(1997)								(1997)			(1997)					(1997)								
	Strain refe		Navarro et al.	This study	This study	This study	This study	This study	Navarro et al.	This study	This study	This study	This study	This study	This study	This study	Navarro et al.	This study	This study	Navarro et al.	This study	This study	This study	This study	Navarro et al.	This study	This study	This study	This study				
23S IGS-types.	IGS-types		D	Ĺц	Ч	Н	D	В	D	D	E	D	C	D	Э	В	D	C	В	А	В	Α	В	Н	Ċ	В	D	D	в	D	в	Д	D
ed in this study and 16S-2	Strain or nodule ^b		TC23 - TC24	TC287	KC693	PMC753	PLD170 - PLD185	RBD570	MG59	DG251	EFG308	KI55	K172	DI247	DI249	PLL300	RPL161	RPL233	RPL526	CN61	CN721	TWN402	NHN750	KN301	KP54	KP80 - KP81	KP82	DP272	RBW161	RBW162 - RBW163	CW726	AW673	CHW746
nodules and isolated strains us	Geographical origin ^a		Tontouta (NC)	Tontouta (NC)	Kouaoua (NC)	Poum (NČ)	Plaine des Lacs (NC)	Rivière Bleue (NC)	Mé Aïu (NC)	Dzumac (NC)	Etoile Filante (NC)	Kouaoua (NC)	Kouaoua (NC)	Dzumac (NC)	Dzumac (NC)	Plum (NC)	Rivière des Pirogues (NC)	Rivière des Pirogues (NC)	Rivière des Pirogues (NC)	Canala (NC)	Canala (NC)	Tiwaka (NC)	Nehoue (NC)	Kouaoua (NC)	Kouaoua (NC)	Kouaoua (NC)	Kouaoua (NC)	Dzumac (NC)	Rivière Bleue (NC)	Rivière Bleue (NC)	Canala (NC)	Amieu (NC)	Chagrin (NC)
Table I. Origin of DNA	Host plant	<i>Gymnostoma</i> nodules	G chamaecyparis	$G.\ chamaecyparis$	$G.\ chamaecyparis$	$G.\ chamaecyparis$	$G. \ deplancheanum$	G. deplancheanum	$G. \ glaucescens$	$G. \ glaucescens$	$G. \ glaucescens$	$G. \ intermedium$	$G.\ intermedium$	$G.\ intermedium$	$G.\ intermedium$	$G. \ leucodon$	$G. \ leucodon$	$G. \ leucodon$	$G. \ leucodon$	$G. \ nodifforum$	$G. \ nodifforum$	$G. \ nodifforum$	$G. \ nodifforum$	$G. \ nodifiorum$	$G.\ poissonianum$	$G.\ poissonianum$	$G.\ poissonianum$	$G.\ poissonianum$	G. webbianum	$G. \ webbianum$	G. webbianum	G. we bbianum	$G. \ webbianum$

inued.	
Cont	
е I .	
Įđ	

Table I. Continued.					
Host plant	Geographical origin ^a	Strain or nodule ^b	IGS-types		Strain reference
Casuarina nodules					
C. equisetifolia/incana	Alva Beach (A)	1Cei1		1	This study
C. equisetifolia	Horseshoe Bay (A)	19 Ce1		1	This study
C. equisetifolia	Cow Bay (A)	CeCB		1	Rouvier et al. (1996)
C. equisetifolia	Garners Beach (A)	CeGB		1	Rouvier et al. (1996)
C. equisetifolia	Pallarenda (A)	CePall		1	Rouvier et al. (1996)
C. equisetifolia	Sauders Beach (A)	CeSB		1	Rouvier et al. (1996)
C. equisetifolia	Wangetti Beach (A)	CeWg2		1	Rouvier et al. (1996)
C. equisetifolia	Wongalinda Beach (A)	CeWB		1	Rouvier et al. (1996)
C. equisetifolia	Cape Hillsborough Beach (A)	14 Ce2		7	This study
C. equisetifolia	Wangetti Beach (A)	CeWg1		7	Rouvier et al. (1996)
C. equisetifolia	Mount Low Beach (A)	CeMLB		2	Rouvier et al. (1996)
$C.\ cunninghamania$	Cattle Creek (A)	7Cc1		33 S	This study
$C.\ cunninghamania$	Bakerville Creek (A)	CcBK		°.	Rouvier et al. (1996)
$C.\ cunninghamania$	Herbert River Crossing (A)	CcHRC		33	Rouvier et al. (1996)
$C.\ cunninghamania$	Jourama Falls (A)	CcJF		ŝ	Rouvier et al. (1996)
$C.\ cunninghamania$	Old Chinaman Creek (A)	CcOCC		°,	Rouvier et al. (1996)
$C.\ cunninghamania$	Tinaroo Creek (A)	CcTC		3	Rouvier et al. (1996)
$C.\ cunninghamania$	West Watsonville (A)	CcWW		ŝ	Rouvier et al. (1996)
Allocasuarina nodules					
A. torulosa	Paluma Road (A)	37At1		4	This study
A. torulosa	Atherton Rifle Range (A)	AltARR1 - AltARR2		4	Rouvier et al. (1996)
A. torulosa	Bluewater (A)	AltBW		4	Rouvier et al. (1996)
A. torulosa	North of Lawyer Creek (A)	AltNLC		4	Rouvier et al. (1996)
A. littoralis	Eungella Road (A)	11Al1		ъ	This study
A. littoralis	Atherton (A)	AllAT		പ	Rouvier et al. (1996)
A. littoralis	Kuranda (A)	AllKur1 - AllKur2		ъ	Rouvier et al. (1996)
A. littoralis	West of Herberton (A)	AllWH		ഹ	Rouvier et al. (1996)

Evolution of Frankia–Casuarinaceae interactions

S361

Table I. Continued.				
Host plant	Geographical origin ^a	Strain or nodule ^b	IGS-types	Strain reference
Casuarina strains C. equisetifolia C. cunninghamania C. equisetifolia	Dakar (Senegal) Florida (USA) Senegal	D11 (atypical) (U) CcI3 (9) CeD (9)	ND 1 1	Gauthier et al. (1981) Zhang et al. (1984) Diem and Dommergues (1983)
Elaeagnaceae strains Elaeagnus umbellata Hippophaë rhamnoides	Illinois (USA) Alps (France)	EUN1f (6) HRN18a (7)	DN UN	Lalonde et al. (1981) Moiroud and Faure-Reynaud (1083)
E. angustifolia E. angustifolia Sherpherdia canadensis	Ohio (USA) Ecully (France) Quebec (Canada)	EaN1-pec (5) Ea1-12 (4) SCN10a (U)	QN QN QN	Lalonde et al. (1981) Fernandez et al. (1989) Mort et al. (1983)
Alnus strain Alnus rubra	Oregon (USA)	ArI3 (1)	QN	Berry and Torrey (1979)
	-			

^a NC: New Caledonia, A: Australia; ^b genomic species numbers are in parentheses; U: undetermined species.

E. Navarro et al.

2.5. Sequencing of IGS amplicons

Before sequencing, the amplification reaction mix was purified by using Centricon-30 concentrators (Amicon-Grace Company, Epernon, France). The amplicons were sequenced using the Deaza G/A sequencing kit (Pharmacia Biotech SA, St-Quentin-Yvelines, France) and the direct DNA sequencing method described by Winship (1989). The fragments were sequenced in both directions. The sequences were determined for both strands.

2.6. Data analysis

The sequences were aligned with previously published sequences (Nalin et al., 1995; Navarro et al., 1997) using the multiple-alignment CLUSTA1V algorithm (Higgins and Sharp, 1988), with manual refinements in the noncoding regions. Distances were calculated according to Kimura's two-parameter model (Kimura, 1980) and phylogenetic analyses were made using neighbour-joining (N-J) (Saitou and Nei, 1987) and parsimony methods (Swofford, 1993). A bootstrap confidence analysis was performed with 1 000 replicates to determine the reliability of the distance tree topologies obtained (Felsenstein, 1985). The resulting tree was drawn by using the N-J plot sofware (Perrière and Gouy, 1996).

2.7. Amplicons restriction analysis

Restriction endonuclease digestions were done with 15 μ L of PCR reaction mixture for each reaction. The endonucleases, NcI1, MspI, HaeIII (all from Boehringer Mannheim, Meylan, France) and ScrF1 (Ozyme, Montigny Le Bretonneux, France) were used as specified by the manufacturers. Electrophoresis was carried out in a horizontal slab gel on a 4 % (w/v) Nusieve (FMC, Rockland, ME, USA) agarose gel containing 0.5 μ g mL⁻¹ ethidium bromide, using TBE electrophoresis buffer (89 mM Tris base, 89 mM boric acid, 2 mM EDTA). Gels were run at 4 V cm⁻¹ for 3 h and photographed as described previously.

3. RESULTS

3.1. Amplification

The method used for extracting DNA from Casuarinaceae nodule lobes provided endophyte DNA that was pure enough to be efficiently amplified with the sets of primers tested (data not shown). A DNA fragment corresponding to the 16S-23S IGS was obtained for all the 55 templates tested, including DNA from isolates and from nonisolated strains (data not shown).

3.2. Sequencing and comparison of the nifD-nifK intergenic spacer

By using the sequencing strategy shown in *figure 1*, a sequence that covers the 3' end of *nifD*, the IGS and the beginning of *nifK* was obtained for all 11 DNAs studied and aligned with the published sequences (Nalin et al., 1995; Navarro et al., 1997) (*figure 2*).



Figure 1. Sequencing strategy for the *nif*D-*nif*K intergenic spacers (IGS). The arrows show the direction of the sequencing reactions.

	1									
B T 2	1									* 100
A115	GATCTTCGCC	CGGGACATGO	ACATCGCCAT	CAACAGCCCG	GCCTGGGACC	TCCTCGAGGC	CCCCTGGTCG	AAGGCCGGCG	AGGTCGCCTG	ATCCACCGCC
RPLIGI	GGICTICGCC	CGCGACATGC	ACATGGCCAT	CAACAGCCCG	ACCTGGGACC	TGATOGAGAC	CCCGTGGTCG	AAGTCCGGAG	AGGTCTTCTG	ACCAGCTAGT
MG59	GGTCTTCGCC	CGCGACATGO	ACATGGCCAT	CAACAGCCCG	ACCTGGGACO	TGATOGAGAC	CCCGTGGTCG	AAGTOOGRAG	ACCTOTOTO	ACCACCTACT
TC24	GGTCTTCGCC	CGCGACATOG	ACATGGCCAT	CAACAGCCCG	ACCTOGGACC	TGATGGAGAC	CCCCTTCCTCC	AACTOCCAC	100000000000000000000000000000000000000	ACCAGETAGE
CN61	GGTCTTCGCC	CGCGACATGG	ACATGGCCAT	CAACAGCCCC	ACCTGGGACC	TATOCACAO	00000000000	ANOTOCOGAG	AGGICITCIG	ACCAGCTAGT
KP54	GGTCTTCGCC	CGCGACATGG	ACATCGCCAT	CAACAGCCCCG	ACCTGGGACC	TOMOGRAGIAC	00000000000	ANOTOCOGAG	AGGICTICIG	AGCAG-TAGT
EUN1f	GATCTTCGCC	CGCGACATGG	ACATOGOCAT	CAACAGOOOG	ACCTICICACC	TOATOCACAC	cccaraarca	ANGICLOUNG	AGGICTICIG	ACCAGTTAGT
SCIN10A	GATCTTOCCC	COCCACATOO	ACATOOCOAT	CAACAGCCCG	ACCIGGGACC	TGATGGAGAC	CCCGIGGICG	AAGICCUGAG	AGGTCTCCTG	ACCAGCTAGT
HRN1 8a	COTOTIOCOC	COCCACAMOO	ACATCOCOAT		ACCIGGGACC	TCATGGAGAC	CCCGIGGICG	AAGTCCGGAG	AGGTCTTCTG	ACCCGTTAG-
Ea1-12	GOTOTICOCC	COCCACATOO	ACATCOCOAT	CAACGGCCCG	ACCIGGGACC	TGATGGAGAC	CCCCTGGTCG	AAGGGTGGAG	AGGTCTTCTG	ACCCG
EaM1-pag	CONCINCIC	CGCGACATOG	ACATUGUGAT	CAACAGCCCG	ACCTGGGACC	TGATOGAGAC	CCCCTGGTCG	AAGGCCGGAG	AGGTGTGCTG	ACCGGATC-T
D11	GGICTICGCC	CUCUACATUG	ACATCGCGAT	CAACAGCCCG	ACCTGGGACC	TCATGGAGAC	CCCCTGGTCG	AAGGCCGGAG	AGGTGTGCTG	ACCGGATT-T
0-12	GGICFICGCC	CGTGACATGG	ACATCGCGAT	CAACAGCCCG	GCCTGGGACC	TGATOGAGAA	CCCCTGGTCG	AAGGCCTGGA	GAAAGGCCTG	ACCOGG
Cers	GGICTICGCC	COTGACATOG	ACATCGCGAT	CAACAGCCCG	ACCTGGGACC	TCATOGAGAC	CCCCTGGTCG	AAGGCCGGAG	ACCTUTTCTC	ACCOGCCOGT
Ced	GGTCTTCGCC	CGTGACATGG	ACATCGCGAT	CAACAGCCCG	ACCTGGGACC	TCATGGAGAC	CCCCTGGTCG	AAGGCCCGGAG	AGGTYTTTCTG	ACCOCCOCC
1Ceil	GGTCTTCGCC	COTGACATOG	ACATCGCGAT	CAACAGCCCG	ACCTGGGACC	TCATOGAGAC	COCOTOCTOC	AAGGCCGCAG	ACCTO TOTOT	Accorcoor
19Ce1	GGTCTTCGCC	CGTGACATGG	ACATCGCGAT	CAACAGCCCG	ACCTGGGACC	TCATOGAGAC	CCCCTCCTCC	ANOCOCCAG	2000000000	ACCOGCCOGT
14Ce2	GGTCTTCGCC	COTGACATOG	ACATTGCGAT	CAACAGCCCG	ACCTGGGACC	TCATCCACAC	CCCCTCCTCC	AAGGCCGGAG	AGGIGITUTG	ACCOGCCOGT
7Cc1	GGTCTTCGCC	CGTGACATGG	ACATTOCGAT	CAACAGCCCC	ACCTGGGACC	TOTTOGACAC	00000000000	AAGGCCGGGAG	AGGIGITUTG	ACCUGCCUGT
37At1	GGTCTTCGCC	CGCGaCaTCC	ACATO CONTAT	Characterooc	Accrossace	TCCTGGAGAC	CECETGGIEG	AAGGCCGGAG	AGGIGITUTG	ACCGGCCGGT
11A11	GOTOTTOGCO	COTGACATOO	ACATCOCOAL	CARCAGECEG	ACCTOGGACC	TCCTGGAGAC	CCCCIGGICG	AAGGCCGGAG	AGGTGTTCTG	ACCGGCCGGT
	0010110000	COLONCHIOG	ACATCOCOAT	CAACAGCCCG	ACCIGGACC	TCCTGGAGAC	CCCCTGGTCG	AAGGCCGGAG	AGGTGTTCTG	ACCGGCCGGT
	101									
N==T2	101									200
ALLS DOLLG	CICCGGGCAC	CCGGGGGGGGA	GGCGCCCAGC	AGTAGACGGC	CACTIGCCGT	-AGACOGCCA	CTCGCAGTA-	GACGGCCACT	CGCCGTAGAC	GGCCACTCGC
KPL101	-CCGGCACCG	GCCAACCA	CCOGCGACCA	CCGGTGATCA	CCGGTGGCTG	GCCGGCGGCC	ATCAGTGGGG	TCTTTTCCTT	CGCACGGTGT	ACCCCCCCCA
MG59	CGGCACGC	GCAACCA	CCGGCGACCA	CCGGTGATCA	CCGGTGGCTG	-CCGGCGGCC	ATCAGTGGG-	TOPPTTCOPP	COCACCETOT	ACCCCCACCA
TC24	-CCGGCACCG	CCGGCAACCA	CCGGCGACCA	CCGGTGATCA	CCGGTGGCTG	-CCGGCGGCC	ATCACING-	The state of the second	CCACCOCOTOT	ACCCCCCCC
CN61	-CCGGCACCG	CGCAACCA	CCGGCGACCA	CCGGTGATCA	CCGGTGGCTG	-000000000	ATCAGOGGG-	COTTOTO	COCHCOGIGI	ACCORCONCON
KP54	CGGCACCG	GCCAACCA	CCGGCGACCA	COGOGANCA	0000000000	100000000	Ancaom	concioni	COLACOGOGC	ACGGCCAGGA
EUN1 f	TCCGGCTCCG	CGCGCAACCA	CCGGCGACCA	COGGTGATCA	COOTOCOTO	-000000000	ARCACICAC			ACG
SCN10A	CCCGGCACCG	CCGCCGACCA	TOGGCGACCA	COGOGARCA	000000000		ATCAGTGAG-	CONCIONAT	CTCGCGGCGC	ACGGCCAGGA
HRN18a		-0000000	C CORCEACEA	CCOGCGATCA	CCGG10GCC-	GCCGGCGGCC	ATC			
Fa1-12	TOCOCTOTO	-0001000	C-GGIGACON	ccconcer		********				
Rall-nee	TCCGGTCTCC	GCCT	GUCUAT					• -		
Pant-bec	TUCOGIUTUU	GCC-	GGCGAT							
011	CCG	CCTCG-CA	CGCCGGG	TGGT	CCCCT				ACCGG	COGCGAGCG
CC13	-CCGGCCCCG	CCGCGGTAGC	CGCAGT							
Ced	-CCGGCCCCG	CCGCGCTAGC	CGCAGT							
lCeil	-CCOGCCCCG	CCGCGGTAGC	CGCAGT							
19Ce1	-CCGGCCCCG	CCGCGGTAGC	CGCAGT							
14Ce2	-CCGGCCCCG	CCGCGGCAGC	CGCAGT							•••••••
7Cc1	-CCGGCCCCG	CCGCGGTAGC	OGCAGT							
37At1	-CCGGCCCCG	COCOGTACC	CCCACT							
11A11	-00000000	COGCOGTAGC	COCACT							
		cededormee	COCHOI							
	201									
A+T3	201	-								300
ArI3	201 AGTAGACGGC	CACTIGCAGT	AGCTGACCAG	TAGTTCATCC	AGGCCTQGCC	GAACCAGCCG	CCAGCGACCC	COTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161	201 AGTAGACGGC ATGCCGTGAG	CACTTGCAGT	AGCTGACCAG CGACCGAAAA	TAGTTCATCC	AGGCCTGGCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA	AGCTGACCAG CGACCGAAAA CGACCGAAAA	TAGTTCATCC GG GG	AGGCCTGGCC	GAACCAGCCG	CCAGOGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA-	TAGTTCATCC GG GG GG	AGGCCTGGCC	GAACCAGCCG	CCAGOGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAACTGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGACCGAAA- CGATGGAAAA	TAGTTCATCC GG GG GG GG	AGGCCTGGCC	GAACCAGCCG	CCAGCGACCC	COTGACCOGA	GAAACGGTCG	300 CTGGCGGCTG
Ar13 RPL161 MG59 TC24 CN61 KP54	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG CA-GAAAGGC	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAACTGG TGATGTTACG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAAA CGATGGAAAA CGGTGCGTCG	TAGTTCATCC GG GG GG GG GT	AGGCCTQGCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGCCGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 EUN1f	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG CA-GAAAGGC CA-GAAAGGC	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAACTGG TGATGTTACG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGATGGAAAA CGGTGCGTCG CGATGGAAAA	TAGTTCATCC GG GG GG GG GT	AGGCCTGGCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 EUN1f SCN10A	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG CA-GAAAGGC ATGCCGTCAG	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA TGATGTTACG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGATGGAAAA CGATGGAAAA	TAGTTCATCC GG GG GG GT GG	AGGCCTOGCC	GAACCAGCCG		CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 EUN1f SCN10A HRN18a	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ACGTGAG CA-GAAAGGC ATGCCGTCAG	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAACTGG TGATGTTACG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGATGGAAAA CGGTGCGTCG CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTQGCC	GAACCAGCCG	CCAGCGACCC		GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161 M059 TC24 CN61 KP54 EUN1f SCN10A HRN18a ER1-12	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG CA-GAAAGGC ATGCCGTCAG	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAACTGG TGATGTTACG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGATGGAAAA CGATGGAAAA CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTQGCC	GAACCAGCCG	CCAGOGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
Ar13 RFL161 MG59 TC24 CCN61 KP54 EUN1f SCN10A HRN18a Ea1-12 EaN1-Dec	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG CA-GAAAGGC CA-GAAAGGC ATGCCGTCAG	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAACTGG TGATGTTACG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGATGGAAAA CGGTGCGTCG CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTGGCC	GAACCAGCCG		CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RFL161 MG59 TC24 CN61 KP54 EUN1f SCN10A HRN18a Ea1-12 EaN1-pec D11	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG CA-GAAAGGC ATGCCGTCAG 	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGATGGAAAA CGGTGCGTCG CGATGGAAAA	TAGTTCATCC GG GG	AGGCCTGGCC	GAACCAGCCG	CCAGOGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RFL161 MG59 TC24 CCN61 KP54 EUN1f SCN10A HRN18a Eal1-12 EaN1-pec D11 Cc13	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG CA-GAAAGGC ATGCCGTCAG 	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGCTGG TGATGTTACG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAAA CGATGGAAAA CGATGGAAAA 	TAGTTCATCC GG GG GG GG GG	AGGCCTOGCC	GAACCAGCCG	CCAGCGACCC	C9TGACCOGA	GAAACGOTCG	300 CTGGCGGCTG
ArI3 RFL161 MG59 TC24 CN61 EUN1f SCN10A HRN18a Ea1-12 EaN1-pec D11 Cci3 Ccd	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG CA-GAAAGGC ATGCCGTCAG 	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGATGGAAAA CGATGGACGTCG CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTQQCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGGTCG 	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 EUN1f SCN10A HRN18a Ea1-12 EaN1-pec D11 Cc13 Ced	201 AGTARACGCC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG CA-GAAAGGC ATGCCGTCAG CA-GAAAGGC ATGCCGTCAG 	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAACTGG CCCGAACTGG TGATOTTACG CCCGAGTCGA	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAAA CGATGGAAAA CGGTGCGTCG CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTGGCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGGTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 ELM1f SCN10A HRN18a Ea1-12 EaN1-pec D11 Cc13 Cced ICce1 ICc11	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG CA-GAAAGGC CA-GAAAGGC 	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG	AGCTGACCAG GGACCGAAAA CGACCGAAAA CGACCGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA 	TAGTTCATCC GG GG GG GG GG	AGGCCTOOCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGOTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 EUNIf SCNIDA HEN18a Eal-12 EaNI-pec D11 Cc13 Ced ICe11 19Ce1	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTCAG ATGCCGTCAG ATGCCGTCAG 	CACTIGCAGT CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAACTIGG CCCGAACTIGG CCCGAGTICGG CCCGGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGGAGTICGG CCCGGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICG CCCGAGTICG CCCGAGTICG CCCGAGTICG CCCGAGTICG CCCGAGTICG CCCGAGTICG CCCGAGTICG CCCGGAGTICG CCCGGAGTICG CCCGGAGTICG CCCGGAGTICG CCCGAGTICG CCCGAGTICG CCCGGAGTICG CCCGAGTICG CCCGGA	AGCTGACCAG GACCGAAAA GACCGAAAA GACCGAAA- CGATGGAAA CGATGGAAAA	TAGTTCATCC GG GG GT GG GG GG 	AGGCCTGOCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGOTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 EUN1f SCM10A HRN18a Eal-12 EaN1-pec D11 Cc13 Ced ICc11 19Ce1 14Ce2	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG CA-GGTGAG CA-GAAAGGC 	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAACTGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG	AGCTGACCAG GACCGAAAA CGACCGAAAA CGACCGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA	TAGTTCATCC GG GG GG GT GG 	AGGCCTOOCC	GAACCAGCCG	CCAGCGACCC	CGTGACCOGGA	GAAACGOTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 EUN1f SCN10A HEN18a Ea1-12 EaN1-pec D11 Cc13 Ccd 12Ce11 19Ce1 14Ce2 7Cc1	201 AGTAGACGGC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTCAG ATGCCGTCAG ATGCCGTCAG 	CACTIGCAGT CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAACTIGG CCCGAACTIGG CCCGAGTICGG	AGCTGACCAG GACCGAAAA GACCGAAAA GACCGAAA- CGATGGAAA CGATGGAAA CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTOOCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGOTCG	300 CTGGCGGCTG
ArI3 RPL161 MG59 TC24 CN61 KP54 EUM1f SCM10A HRN18a Ea1-12 EaN1-pec D11 Cc13 Cc4 ICc11 19Cc1 14Cc2 7Cc1	201 AGTAGACGCC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG ATGCCGTCAG 	CACTRICAGT CCCGAGTICGA CCCGAGTICGA CCCGAAGTICGA CCCGAAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA CGATGGAAA CGATGGAAA CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTOOCC	GAACCAGCCG	CCAGCGACCC	CGTGACCOGA	GAAACGOTCG	300 CTGGCGGCTG
ArI3 RPL161 NG59 TC24 CN54 EIN1f EIN1f EIN1f Eal-12 EaN1-pec D11 Cc13 Ccd 1Cc11 19Ce1 14Cc2 7Cc1 37Ac1 11A11	201 AGTAGACGGC ANGCCGTDAG ANGCCGTDAG ANGCCGTDAG ANGCCGTDAG ANGCCGTDAG ANGCCGTCAG 	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAAA CGATGGACAGAAA CGATGGCATCG CGATGGCATCG CGATGGCATCG CGATGGCATCG CGATGGCAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTOOCC	GAACCAGCCG	CCAGCGACCC	CGTGACCOGA	GAAAACGOTCG	300 CTGGCGGCTG
Ar13 RPL161 MG59 TC24 CCN61 EUN1f SCN10A HRN18a Eal-12 EC13 Cc13 Cc13 Cc13 Cc13 Cc11 19Ce1 14Ce2 7Cc1 37Ac1 11A11	201 AGTAGACGCC ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ACGTGAG CA-GAAAGC ATGCCGTCAG 	CACTIGCAGT CCCCAGTICGA CCCCAGTICGA CCCCAACTIGG CCCCAACTIGG CCCCAACTIGG CCCCAGTIGGG	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA- CGATGGAAA- CGATGGACATCA CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTOOCC	GAACCAGCCG	CCAGCGACCC	CGTGACCOGA	GAAACGOTCG	300 CTGGCGGCTG
ArI3 RPL161 NG59 TC24 CN61 KF54 EIR1f EIR1f Eal-12 EaN1-pec D11 Cc13 Ccd 1Cc11 19Ce1 14Cc2 7Cc1 37Ac1 11A11	201 AGTAGACGGC ANGCCGTCAMG ANGCCGTCAMG ANGCCGTCAMG ANGCCGTCAMG ANGCCGTCAMG 	CACTIGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGG CCCGAGTCGAGT	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAAA CGATGGACAGAAA CGATGGACAC CGATGGCATCG CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTOOCCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGOTCG	300 CTUGCGGCTG
Ar13 RPL161 MC59 TC24 CC61 KP54 EUN1f EUN1f SCN10A HRN18a Ea1-12 Ea1-12 Cc13 Ccd 10Ce11 19Ce1 14Ce2 7Cc1 37Ac1 11A11 Ar13	201 AGTAGACGGC ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCGTGAGA ANGCGTCAG 	CACTIGCAGT CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGG CCCGAGTICGA CCCGAGT	AGCTGACCAG CGACCGAAAA CGACCGAAAA CGACCGAAA CGATGGAAAA CGATGGACTCG CGCTGGCTCG CGATGGAAAA	TAGTTCATCC GG GG GG GG GG	AGGCCTOQCC	GAACCAGCCG	CCAGCGACCC	CGTGACCGGA	GAAACGOTCG	300 CTGGCGGCTG
Ar13 RPL161 MG59 TC24 CN61 ELM1f SCN10A HRN18a EAN1-Dec D11 Cc13 Ccd 12Ce11 13Ce1 14Ce2 7Cc1 14Ce2 7CC1 14Ce2 7CC1 14Ce2 7CC1 14Ce2 7CC1 14Ce2 7CC1 14Ce2 7CC1 14CC2 7CC1 14CC2 7CC1 14CC2 7CC1 14CC2 7CC1 14CC2 7CC1 14CC2 7CC1 14CC2 7CC1 14CC2 7CC2 7CC2 7CC2 7CC2 7CC2 7CC2 7CC2	201 AGTAGACGGC ANGCCGTGAG ATGCCGTGAG ATGCCGTGAG CA-GAAAGGC CGTGAG CGTGAG CGTGAG 	CACTTGCAGT CCCGAGTCGAGT	AGCTGACCAG GGACCGAAAA GGACCGAAAA CGACGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAACA CGATGGAACA CGATGGAACA CGATGGAACA CGATGGAACA CGACGGAAAA CGATGGAACA CGACGGAAAA CGATGGAACA CGACGGAAAA	TAGTYCATCC GG GG GG GG GG	AGGCCT00CC	GAACCAGCCG 	CCAGCGACCC	COTGACCOGA	GAAACGOTCG	300 CTGGCGGCTG
Ar13 RPL161 MG59 TC24 CN61 KP54 ERM16 ERM16 ERM162 EAN1-Dec D11 Cc13 Cc4 1Cc11 19Cc1 14Cc2 7Cc1 37At1 1At1 Ar13 RPL161 MG59	201 AGTNGACGGC MGCCOGGG MGCCOGGG AGCCOGGGG AGCCOGGGGGGG AGCCOGGGCAGG CA-GBAAGC 	CACTEGOACT CCCGAGTOCA CCCGAGTOCA CCCGAGTOCA CCCGAGTOCA CCCGACTOS TCATOTYACG CCCGACTOS	AGCTGACCAG CGACCGANA CGACCGANA CGACGANA CGACGGANA CGATGGANA CGATGGANA CGATGGANA CCATGGANA CCATGGANA CCATGGANA CCACTGGGAC CACTGGGAC	TAGTYCATCC GG GG GG GG GG	AGGCCTGGCC	GAACCAGCCG 		COTGACCOGA	GAAACGOTCG	300 Criaccaccia
Ar13 RPL161 MG59 TC24 ELM16 SCN10A HRN18a Eal-1-pec D11 Cc13 Ccd1 13Cel1 14Ce2 7Cc1 13Cel1 14Ce2 7Cc1 13TA11 Ar13 Ar13 Ar24 CC4 CC5 CC5 CC5 CC5 CC5 CC5 CC	201 AGTRACKGCC ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG 	CACTTGCAST CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCCAGGCCC	AGCTGACCAG CBACGGAAAA CGACCGGAAA CGACCGGAAA CGACCGGAAA CGACGGAAA CGACGGAAA CGACGGAAA CGACGGAAA CGACGGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGGAAAA CGACGGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGAAAA CGACGGAAAA CGACGGAAAA CGACGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAA CGACGGAAA CGACGGAAAA CGACGGAAAA CGACGGAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGGAAAA CGACGAAAAA CGACGAAAAA CGACGAAAAA CGACGAAAA CGACGAAAAA CGACGAAAAA CGACGAAAAA CGACGAAAAA CGACGAAAAA CGACGAACAAAA CGACGAAAAA CGACGAAAAA CGACGAAAAA CGACGAACGA	TAGTTCATCC GG GG	AGGCCTOGCC	GAACCAGCCG 	CCAGCGACCC	COTGACCOGA	GANACGOTCG	300 Criaccacrag
ArI3 RPL161 MG59 TC24 CC4 EIM16 EIM16 SCN10A EaN12 Ced 1Cei1 19Cel 1Ceci 14Ce2 7Cc1 37Ac1 1Al1 ArI3 RPL161 MC39 TC4 HC59 HC59 HC50	201 AGTACACGCC AGCCGGGAG AGCCGGGAG AGCCGGGAG AGCCGGGAG CA-GAAAGGC AGCCGGCAG 	CACTTGCAST CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCCGAGTCGA CCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA C	AGCTGACCAG CGACCGAAA GGACGGAAA CGACGGAAA CGACGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CCATGGGAC CCACCTGGGAC CACCTGGGAC CCTCA	TAGTTCAATCC CG	AGGCCTGOCC	GMACCAGCGG 		COTGACCGGA	GAAACGOTCG	300 CTGCCGCCG CTGCCGCGCGCGCGCGCGCGCGCGCGCGGGGGGGG
Ar13 RPL161 MG59 TC24 CN61 EUN1f SCN10A HRN18a Ea1-12 Ea1-12 Cc13 Ccd 1Cc11 19Cel 14Ce2 7Cc1 37At1 1A11 Ar13 RPL161 MG59 TC24 CK61 SCN10A HG59 CK61 SCN10A SCN10A HRN18a Ea1-12 Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a Ea1-12 SCN10A HRN18a CC13 CC13 CC13 HRN18a CC13 CC13 CC13 CC13 CC13 CC13 CC13 CC13 CC13 CC2 CC13 CC13 CC13 CC2 CC13 CC13 CC2 CC13 CC13 CC2 CC13 CC2 CC13 CC13 CC2 CC13 CC2 CC13 CC13 CC2 CC13 CC13 CC2 CC13 CC13 CC2 CC13	201 AGTRACKGCC ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGGCGCAG CCGCGGCGCAG CCGCGCGCAGCAG	CACTTGCACT CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCCCGAGTICGA CCCCCGAGTICGA CCCCCGAGTICGA CC	AGCTOACCAG CGACCGAAAA CGACCGAAAA CGACCGAAAA CGACCGAAAA CGACGGAAAA CGACGAAAAA CGACGAAAAA CGACGAAAAAA CGACGAAAAAA CGACGAAAAAA CGACGAAAAAA CGACGAAAAAA CGACGAAAAAAA CGACGAAAAAAA CGACGAAAAAAAA	TAGTYCATCC GG GG	AGGCCTOGCC	GAACCAGCCG 		COTGACCOGA	GAAACGOTCG	300 Criaccacrig
ArI3 RPL161 MG59 TC24 CN61 EUM16 EUM16 ECN10A HRV18a Eal-12 EAN1-pec D11 SCM10A ICeil 19Cel 14Cc2 7Cc1 37Ac1 11Al1 ArI3 RFL161 MG59 ArI3 KP54 KF KF KF KF KF KF KF KF KF KF	201 AGTRACKGCG ANGCCGTUAG ANGCCGTUAG ANGCCGTUAG ANGCCGTUAG CA-GRAAGGC A-GTUAG 	CACITISCAST CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCCCAGTICCA CCCCCAGTICCA CCCCCAGTICCA CCCCCCAGTICCA CCCCCCAGTICCA CCCCCCAGTICCA CCCCCCAGTICCA CCCCCAGTICCA CCCCCCAGTICCA CCCCCCCCAGTICCA CCCCCCAGTICCA CCCCCCCCCAGTICCA CCCCCCAGTICCA CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	AGCTGACCAG CGACCGAAA GGACGGAAA GGACGGAAA GGACGGAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CCATGGAAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAAA CCATGGAA CC	TAGTTCAATCC GG GG	AGGCCTGGCC 	GAACCAGCGG 			GAAACGOTCG	300 CTGRCGRCTG CTGRCGRCTG
ArI3 RFL161 MG59 TC24 CN61 ERN16 ERN16 SERN18a Eal-12 Eal-12 Eal-12 Eal-12 Cc13 Cc1 16ce1 14Ce2 7Cc1 37Ac1 14Cl2 ArI3 RFL161 MG59 TC24 CN61 ERN16 ENN16	201 AGTACACGCC ANGCCGGAG ANGCCGGAG ANGCCGGAG CA-GANAGC ANGCCGGAG C-CCGGGAG C-CCGGCGCAGG C-CCGGCGCAGG CCGCCGCAGGCAGG	CACTTSCAST CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCCGATCACCA CCCCGATCACCA CCCCGATCACCA CCCCGATCACCA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCGATCACA CCCCATCACA CCCCATCACACA CCCCATCACACA CCCCATCACACACA	AGCTOACCAG CGACCGANAA CGACCGANAA CGACCGANAA CGACGGANAA CGACGGANAA CGATGGANAA CGATGGANAA CGATGGANAA CGATGGANAA CGATGGAGAC CCATGGGAC CCATGGGAC CCATGGGAC CCATGGGAC CCCCTGGGAC CCCCTGGCCTGC CCCCTGGGAC CCCCTGGGAC CCCCTGGGAC CCCCTGGCCTGC CCCCTGGGAC CCCCTGGGAC CCCCTGGGAC CCCTGGCCTGC CCCTGGCC CCCCTGGCC CCCCCC CCCCCC CCCCC CCCCCCCC	TAGTYCATCC GG GG	AGGCCTOGCC	GAACCAGCCG 			GAAACGOTCG	300 Criaccacrag
ArI3 RPL161 MG59 TC24 CN61 EUM1f SCN10A HFM18a Eal-12 EaN1-pec D11 Cc13 Cc23 Cc43 14Cc2 7Cc1 37Ac1 11A11 ArI3 RFL161 MG59 ArI3 EUN1f EUN1f EUN1f EUN1f EUN1f EUN1f EUN1f EUN1f EUN1f EUN1f EUN1f	201 AGTRACACGCC ANGCCOGRUA ANGCCOGRUA ANGCCOGRUA ANGCCOGRUA CA-GARAAGCC 	CACITISCAST CCCGAGITCA CCCGAGITCA CCCGAGITCA CCCGAGITCA CCCGAGITCA CCCCCAGITCA CCCCCCAGITCA CCCCCCAGITCA CCCCCCAGITCA CCCCCAGITCA CCCCCCCCAGITCA CCCCCCAGITCA CCCCCCAGITCA CCCCCCCCAGITCA CCCCCCCCCCCCAGITCA CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	AGCTGACCAG CRACGRAAA GCACCGRAAA GCACCGRAA- GCACCGRAA- GCACGGRAA- CGATGGRAAA 	TAGTTCATCC GG GG GG GG G	AGGCCTGGCC	GAACCAGCCG			GANACOGOTCG	300 CTGRCGCGTG
ArI3 RPL161 MG59 TC24 CN61 KP54 ERM16 ERM16 ERM162 EAN1-Dec D11 Cc13 Ccd 1Cc11 19Cc1 14Cc2 7Cc1 37At1 1At1 ArI3 RPL161 MG59 TC24 CN61 KP54 ECN10A HRS18a	201 AGTACACGCC AGCCGGGA AGCCGGGA AGCCGGGA AGCCGGGA AGCCGGGA AGCCGGGA AGCCGGGA AGCCGGGA AGCCGGGA AGCCGGGA AGCCGA AGCCGGA AGCCGGA AGCCGA AGCCGGA AGCCGA AGCCGGA AGCCGGA AGCCGGA AGCCA AGCCA	CACTTSCAST CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCCGATCAC CCCCGATCCA CCCCA CCCA CCCCA CCCA CCCCA CCCA CCCA CCCA CCCA CCCCA CCCA CCC	AGCTGACCAG CGACCGAAA CGACCGAAA CGACCGAAA CGACGGAAA CGACGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CCACTGGGAC CCACTGGGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCCGAGAC CCCCGAGAC CCCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGAGAC CCCCGGAGAC CCCCGAGAC CCCCGGAC CCCCGGAC CCCCGGAGAC CCCCGGAC CCCCGGAGAC CCCCGGAC CCCCGGAGC CCCCGGAGC CCCCGGAC CCCCGGAC CCCCGGAC CCCCGGAC CCCCGGAC CCCCGCGAC CCCCGCGAC CCCCGCGAC CCCCGCGCGAC CCCCGCGCGAC CCCCGCGAC CCCCGCGCAC CCCCGCGCAC CCCCGCGCAC CCCCGCGCAC CCCCCCGCCGCC CCCCCCGCGCC CCCCCGCCCCCCC	TAGTTCATCC GG	AGGCCT9OCC	GMACCAGCCG			GAAACGOTCG	300 CTGACGGCTG CTGACGGCTGAGGT CCCTGAGGT CCCCTG CCCTG CCCTGAGGT CCCCCG CCCTG CCCTG CCCCCG CCCCCCCG CCCCCCCC
ArI3 RPL161 MG59 TC24 CN61 ELM1f SCN10A HRN18a Eal-12 EAM1-pec D11 Cc13 Ccd1 1Cc11 19Cc1 14Cc2 7Cc1 13Cc1 14Cc2 7Cc1 14Cc2 7Cc1 14Cc2 7Cc1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 14Cc2 7Cc2 CC1 11Cc2 11Cc2 CC2 CC1 11Cc2 CC2 CC2 CC2 CC2 CC2 CC2 CC2	201 AGTACACGCC ANGCCOGGA ANGCCOGGA ANGCCOGGA ANGCCOGGA CA-GANAGC 	CACTTGCAGT CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCGAGTICCA CCCCCCCAGTICCA CCCCCCAGTICCA CCCCCAGTICCA CCCCCCAGTICCA CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	AGCTGACCAG CBACGBAAA CGACCGBAAA CGACCGBAAA CGACCGBAAA CGACGGAAA CGACGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGACAA CGATGGAAA CGATGGACAAA CGATGGACAA CGATGGACAA CGATGGACAA CGATGGACAA CGATGGACAA CGATGGACAAA CGATGGACAAA CGATGGACAAA CGATGGACAAA CGATGGACAAA CGATGGACAAA	TAGTTCATCC GG GG	AGGCCTGGCC	GAACCAGCCG 		COTGACCOGA	GANACGOTCG	300 Crdaccacrag
ArI3 RPL161 MG59 TC24 CC4 EIM16 SCN10A EaN12 Ced ICei1 16Cei 16Cei 16Cei 16Cei 14Ce2 7Cc1 37Ac1 14Al1 ArI3 RFL161 KC59 CN64 SCN10A HC51 KC59 EIM1 f SCN10A KC54 EIM1 f SCN10A EIM1 f SCN10A KC54 EIM1 f SCN10A EIM1 f SCN10A	201 AGTACACGCC AGTCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CCGCGCCACAG CCGCGCCACAG CCGCGCCACAG CCGCGCCACAG CCGCGCCACAG	CACTTGCAST CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGCGAGTCGA CCCCGCGCCG CCCCGCGCGCCC CCCCCGCCG CCCCCGCCG	AGCTGACCAG CGACCGAAA GGACGGAAA CGACGGAAA CGACGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CCACTGGGAC CACCTGGGAC CACCTGGGAC	TAGTTCATCC CG CG GG GT GT GT GT GT GT GT GG GT GG	AGGCCTGOCC	GMACCAGCCG			GANACOGOTCG	300 CTGCCGCG
Ar13 RPL161 MG59 TC24 CC4 EUN1f SCN10A HRN18a Eal-12 Eal-12 EC13 Ccd1 13Ce1 13Ce1 14Ce2 7Cc1 13Ce1 14Ce2 7Cc1 13TA1 1AC1 1AC2 CC4 CC4 CC4 CC4 CC4 CC4 CC4 C	201 AGTRACACGCC ANGCCOGGA ANGCCOGA ANGC	CACTTGCAGT CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCGAGTICCA CCCCCGAGTICCA CCCCCGAGTICCA CCCCCGAGTICCA CCCCCGAGTICCA CCCCCGAGTICCA CCCCCGAGTICCA CCCCCGAGTICCA CCCCCGAGTICCA CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	AGCTOACCAG COACCEANAA	TAGTTCATCC GG GG	AGGCCTOGCC	GAACCAGCGG		COTGACCOGA	GANACGOTCG	300 Cr3aCcGcrG
ArI3 RPL161 MG59 TC24 CC81 EXM16 SCN10A HRN18a Eal-12 Eal-12 Eal-12 Cod ICeil 14Cc2 TCc1 37Ac1 11A11 ArI3 RFL161 MG59 TC24 CCN61 SCN10A HRN16a Eal-12 ECN61 CC13 SCN10A HRN16a Eal-12 SCN10A HRN16a Eal-12 CC13 SCN10A HRN16a Eal-12 CC13 SCN10A HRN16a Eal-12 SCN10A HRN16a Eal-12 CC13 SCN10A HRN16a Eal-12 CC13 SCN10A HRN16a Eal-12 SCN10A HRN16a HRN16a Eal-12 SCN10A HRN16a Eal-12 SCN10A HRN16a Eal-12 SCN10A HRN16a Eal-12 SCN10A HRN16a HRN16a Eal-12 SCN10A HRN16a HRN16A HRN16a HRN16A HR	201 AGTRACACGC AGTRACACGGTUBA ATGCCGTUBA ATGCCGTUBA ATGCCGTUBA ATGCCGTCAG 	CACTTGCAGT CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCCGAGTCCA CCCCGAGTCCA CCCCGAGTCCA CCCCGAGTCCA CCCCGAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCA CCCCCAGTCCAGT	AGCTGACCAG CRACGBAAA GGACGBAAA GGACGBAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CCATGGGACAAT CCTCA CCTC	TAGTTCAATCC GG GG	AGGCCTGGCC 	GAACCAGCGG 		COTGACCGGA	GANACGOTCG	300 CTGSCGCGTG CTGCGGGT CCCTGAGGGT CCCTGAGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGGT CCCTGCGGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGT CCCTGGGGGT CCCTGGGGGT CCCTGGGGGT CCCTGCGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGT CCCTGGGG CCCTGGGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGGT CCCTGCGGG CCCTGCGGG CCCTGGGG CCCTGGGG CCCTGGGG CCCTGGGG CCCTGGGG CCCCGGG CCCCGGG CCCCGGG CCCCGGG CCCCGG CCCCG CCCCGGG CCCCG CCCCGG CCCCG CCCCG CCCC CCCC CCC CC CCC CC
ArI3 RPL161 MG59 TC24 CN61 KP54 EXNIF EXNIF EXNIF EXNIF EXNIF EXNIF EXNIF EXNIF Ccd 1Ccil 14Cc2 TCc1 37Ac1 14Cc2 TCc1 37Ac1 1Ac1 ArI3 RPL161 MG59 TC24 CN61 KP54 EXNIF E	201 AGTACACGCC ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG CA-GAAAGC ANGCCGTCAG C-CCGGTGAG C-CCGGGCGCAG CCCCGCGCAGG CCCCGCGCAGG CCCCGCGCAGG CCCCGCGCAGG CCCCGCGCAGG	CACTTSCAST CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCGAGTCCA CCCCGAGTCCA CCCCGATCAC CCCCGATCAC CCCCGATCCA CCCCGATCAC CCCCGATCAC CCCCGATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCCCATCAC CCCATCAC CCCCATCAC CCCATCAC CCCCATCAC CCCCATCAC CCCATCAC CCCATCAC	AGCTGACCAG CGACCGAAA CGACCGAAA CGACCGAAA CGACCGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGACAA CGATGGACAAC CTCA CTCA	TAGTTCATCC GG	AGGCCT90CC	GMACCAGCCG		COTGACCOGA	GAAACGGOTCG	300 CTG3CG2CTG
ArI3 RPL161 MG59 TC24 CN61 EUM1f SCN10A HFM18a Eal-12 EaN1-pec D11 Cc3 Cc3 Cc3 Cc3 Cc3 Cc3 14Cc2 7Cc-1 37Ac1 11A11 ArI3 RFL161 MG59 EUN1f SCN10A HRM18a Eal-12 EUN1f SCN10A HRM18a Eal-12 EUN1f Cc3 Cc3 Cc3 Cc3 Cc3 Cc3 Cc4 Cc3 Cc4 Cc3 Cc4 Cc3 Cc4 Cc3 Cc4 Cc3 Cc4 Cc3 Cc4 Cc4 Cc3 Cc4 Cc4 Cc4 Cc4 Cc4 Cc4 Cc4 Cc4	201 AGTRACACGCC ANGCCGTGAG	CACTTGCAGT CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCCGCCG	AGCTGACCAG CRACGRAAA GCACCGRAAA GCACCGRAA GCACCGRAA GCACGGRAA CGATGGRAAA CGATGGRAAA CGATGGRAAA CGATGGRAAA CGATGGRAAA CCACTGGGRAC CCACCTGGGRAC CCACCTGGGRAC CCACCTGGGRAC CCACCTGGGRAC CCACCTGGGRAC CCACCTGGGRAC CCACCTGGGCACATC CCACCTGGGCACATC CCACCTGGGCACATC CCACCTGGGCACATC CCACCTGGGCCACATC CCACCTGGCCACATC	TAGTTCAATCC GG GG	AGGCCTGGCC 	GAACCAGCCG 		COTGACCOGA	GAAACGG9TCG 	300 CTGSCGCGG
ArI3 RPL161 MG59 TC24 CN61 KP54 EIM16 SCN10A SCN10A EAN1-Dec D11 Cc13 Ccd 1Cc11 19Cc1 14Cc2 7Cc1 37Ac1 1A11 ArI3 RPL161 MG59 TC24 CN61 KP54 SCN10A HRN18a EAN1-Dec D11 Cc13 Cc14 CC16 CC	201 AGTRACACCG AGTRACACGTUA ATACCGTUA ATACCGTUA ATACCGTUA ATACCGTUA ATACCGTUA ATACCGTUA CA-GANAGC ATACCGTUA CA-GANAGC ATACCGTUA COCGGUA COCGGUA CCGCGGCAGG CCGCGCCAGG CCGCGCCAGG CCGCGCCAGG	CACTTGCAST CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGCG CCCCGAGTCGA CCCCGCG CCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCGCG CCCCGCG CCCCGCG CCCGCG CCCGCG CCCGCG CCCGCG CCCGCG CCCGCG CCCGCG CCCGCG CCCGG CCCGCG CCCG CCCGG CCCG CCCGG CCCG CC	AGCTGACCAG CRACCGAAA GCACCGAAA CGACCGAAA CGACCGAAA CGACGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGACAAC CCTG CGGCCCACC CGGGCCCACC CGGCCCACC	TAGTTCAATCC GG	AGGCCT9OCC	GMACCAGCCG		COTGACCGGA	GAAACGOTCG 	300 CTGCCGCCTG
ArI3 RPL161 MG59 TC24 CN61 ELM1f ELM1f SCN10A HFN18a Eal-12 EaN1-pec D11 Cc13 Ccd1 1Cc11 11A11 1Cc13 TCc1 37Ac1 11A11 ArI3 RFL161 MG59 ELM1f SCN10A HRN18a Eal-12 EAN1-pec D11 Cc13 Ccd1 10A11 EAN1-pec D11 Cc13 Cc13 Ccd1 10A11 EAN1-pec D11 Cc13 Cc23 Cc13 Cc13 Cc23 Cc13 Cc13 Cc23 Cc13 Cc23 Cc13 Cc23 Cc13 Cc23 Cc13 Cc23 Cc13 Cc23 Cc13 Cc23 Cc23 Cc13 Cc23	201 AGTRACAGGC AGTRACAGGC AGTRACAGGCGGGGG AGTRACAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	CACTTGCAGT CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCGAGTICGA CCCCGCG CCCCGAGTICGA CCCCGCG CCCCGAGTICGA CCCCGAGTICGA CCCCGCG CCCCGAGTICGA CCCCGCG CCCCGAGTICGA CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCGCG CCCCCG CCCG CCCCG CCCG CCCCG CCCG CCCG CCCG CCCG CCCCG CCCG CCCCG CCCG C	AGCTGACCAG CBACGRAAA GCACCGRAAA GCACCGRAAA GCACCGRAAA GCACCGRAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGAAAA CGATGGACAC CCACCGRAAAA CCACCGRAAA CCACCGRAAA CCACCGRAAA CCACCGRAAA CCACCGRAAA CCACCGRAAA CCACCGRAAA CCACCGRAAA CCACCGRAAA CCACCGRAAAA CCACCGRAAAA CCACCGRAAAA CCACCGRAAAA CCACCGRAAAAAA CCACCGRAAAA CCACCGRAAAAAA CCACCGRAAAAAAAAAA	TAGTTCATCC GG GG GG GG GG </td <td>AGGCCTOGCC </td> <td>GAACCAGOCG </td> <td></td> <td>CGTGACCGGA</td> <td>GANACGOTCG</td> <td>300 CTG3CG3CGG</td>	AGGCCTOGCC 	GAACCAGOCG 		CGTGACCGGA	GANACGOTCG	300 CTG3CG3CGG
ArI3 RPL161 MG59 TC24 CC4 EIM16 EIM16 EIM16 EIM162 EIM162 Ced 1Cei1 13Cel 1Cei2 Ced 1Cei1 14Ce2 7Cc1 37Ac1 1Al1 ArI3 RPL161 KD54 EIM16 SCN10A HC52 CC4 Ci3 Cc4 Cc53 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc4 Cc53 Cc54 Cc55	201 AGTRACACGC AGTRACACGGTUBA ATROCCGTUBA ATROCCGTUBA ATROCCGTUBA ATROCCGTUBA CA-GRANAGC 	CACTTGCAGT CCCGAGTCGA CCCCGAGTCGA CCCCGA CCCGAGTCGA CCCCGAGTCGA CCCCGAGTCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCGA CCCCG CCCC	AGCTGACCAG CRACGRAA GGACGRAA GGACGRAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CCATGGAAA CCATGGAAA CCATGGACAAC CCATGGACAAC CCATGCAAC CGGGCCAAC CGGGCCAAC CGGGGCCAAC CGGGGCCAAC CGGGGCCAAC CGGGGCCAAC	TAGTTCATCC CG CG CG GG CG GT CG GG GG GG	AGGCCTGOCC	GMACCACCCG 			GANACOGOTCG	300 CTGCCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
ArI3 RPL161 MG59 TC24 CN61 KP54 ELM1f SCN10A HRN18a Eal-12 EaM1-pec D11 Cc13 Ccd1 13Ce1 13Ce1 13Ce1 14Ce2 7751 13A11 14Ce2 7751 13A11 14Ce2 7751 13A11 14Ce2 7751 13A11 14Ce2 7751 14Ce2 7751 13A11 14Ce2 7751 13A11 14Ce2 7751 14Ce2 7751 13A21 14Ce2 7751 13A21 14Ce2 7751 13A21 14Ce2 7751 13A21 14Ce2 7751 13A21 14Ce2 7751 13A21 14Ce2 7751 13A21 13A21 13A21 14Ce2 14Ce2 7751 13A21 14Ce2	201 AGTACACGCC ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG ANGCCGTGAG CA-GANAGC C	CACITISCAST CCCGAGNICCA CCCGAGNICCA CCCGAGNICCA CCCGAGNICCA CCCGAGNICCA CCCGAGNICCA CCCGAGNICCA CCCCGAGNICCA CCCCGAGNICCA CCCCGAGNICCA CCCCGAGNICCA CCCCGAGNICCA CCCCCAGNICCA CCCCAGNICCA CCCCAGNICCA CCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCAGNICCA CCCCCCAGNICCA CCCCCCAGNICCA CCCCCCCCCA CCCCCCA CCCCCCCCCCCCCCC	AGCT0ACCAG CBACCGAAA CGACCGAAA CGACCGAAA CGACCGAAA CGACGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CCATCOCGACAC CTCA CTCA CTCA CTCA CTCA CTCA CTC	TAGTTCATCC GG GG	AGGCCTOGCC 	GAACCAGCCG 		COTGACCOGA	GANACGOTCG	300 CTG3CG3CGTG CTG3CG3CGG CTG3CG3CGGGG CCCCAAGGT CCCCGAGGGGGGGGGGGGGGGGGG
ArI3 RPL161 MG59 TC24 CC4 EXMIDA EANIDA EANIDA EANIDA EANIDA EANIDA Ced ICeil Idce2 TCc1 37Ac1 IANI ArI3 RPL161 MG59 CC4 ICeil IAC2 CC4 ICeil IAC2 CC5 CC4 ICeil IAC2 CC5 CC4 ICeil IAC2 CC5 CC5 CC5 CC5 CC5 CC5 CC5 C	201 AGTACACGCC AGTCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTGAG ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG CA-GRAAGCC ATGCCGTCAG ATGCGTCAG ATGCCGTCAG AT	CACTTGCAGT CCCGAGTCGA CCCGCGAGTCGA CCCGAGTCGA CCCGCAGTCGA CCCGCAGTCGA CCCGCAGTCGA CCCGCAGTCGA CCCGCAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGAGTCGA CCCGCG CCCGCG CCCGCG CCCGAGTCGA CCCGCG CCCGCG CCCGCG CCCGCG CCCCG CCCGCG CCCG CCCCG CCCG CCCCG CCCG CCCCGC	AGCTGACCAG CRACGRAA GRACGRAA GGACGRAA GGACGRAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CGATGGAAA CCATGGAAC CCATGGAAA CCATGGAAC CCATGGAAA CCATGGAAC	TAGTTCAATCC CG CG	AGGCCTGOCC 	GAACCACCCG 			GANACOGOTCG	300 CTGSCGCTG CTGSCGGT CGCTGCGGGT CGCTGGGGT CGCTGCGGGT CGCTGCGGGT CGCTGGGGT CGCTGGGGGT CGCTGGGGT CGCGGGG CGCGGGG CGCGGGG CGCGGGG CGCGGGG CGCGGGG CGCGGGG CGCGGGG CGCGGG CGCGGG CGCGGG CGCGG CGCGGG CGCGGG CGCGG CGCGGG CGCGG CGCGGG CGCGGG CGCGG CGCGG CGCGG CGCG CGCGG CGCG CGCGG CGCG CGCGG CGCG CGC CGC CGCG CGC

Figure 2. Alignment of the nifD-nifK intergenic spacer (IGS). The stop codon of nifD and the start codon of nifK are marked by an asterisk.

	401									500
ArI3	CCTTGACCAC	AACGACATCT	TCCGCGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGCGGCCC	CGAAGGAAGA	GGTCCAGCGG
RPL161	CCTCAACCAC	ACCGAGCTGT	TCAAGGACGA	GGTCTACACC	AAGCAGTTCG	AGACGAAGOG	CGAGTTCGAG	AACGOTGOGA	TOGACGCCCGA	GOTCARCCOG
MG59	CCTCNACCAC	ACCGAGCTGT	TCAAGGACGA	GGTCTACACC	AAGCAGTTCG	AGACCAAGCG	CGAGTTCGAG	AACGGTGCGG	ATCACGCCGA	GGTCAACCOG
TC24	CCTCGACCAC	ACCGAGCTGT	TCAAGGACGA	GGTCTACACC	AAGCAGTTCG	AGACCAAGCG	CGAGTTCGAG	AACGGTGCCG	ATCACGCCGA	GGTCAACCGG
CN61	CCTCGACCAC	ACCGAGCTGT	TCAAGGACGA	GGTCTACACC	AAGCAGTTCG	AGACCAAGCG	CGAGTTCGAG	AACGGTGCGG	ATGACGCCGA	GGTCAACCGG
KP54	CCTCGACCAC	ACCGAACTGT	TCAAGGACGA	GGTCTACACC	AAGCAGTTCG	AGACCAAGCG	CGAGTTCGAG	AACGGTGCGG	ATGACGCCGA	GGTCAACCGG
EUN1 f	CCTCGACCAC	ACCGAGCTGT	TCAAGGACGA	GGTCTACACC	AAGCAGTTCG	AGACCAAGCG	CGAGTTCGAG	AACGGCGCCG	ATGACGCCGA	GGTCAACCGG
SCN10A	CCTCGACCAC	AACGAGCTGT	TCAAGGACGA	GGTCTACACC	AAGCAGTTCG	AGACCAAGCG	CGAGTTCGAG	AACGGTGCGG	ACGACGCCGA	GGTCAACCGG
HRN18a	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACAGG	AAGCAGTTCG	AGACCAAGCG	CGAGTTCGAG	AACGGCGCCG	ATGACGCCGA	GGTCAACCGG
Eal-12	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGGG	CGAGTTCGAG	AACGGCACGG	ACGCCGCCGA	GGTCAACCGG
EaN1-pec	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGGC	CGAGTTCGAG	AACGGCAGTG	AATCCGCCGA	GGTTTCCCCGA
D11	GCTCGACCAC	AACGAGCTGT	TCAAGGGCGA	GGCCTACCAC	AAGCAGTTCG	AGGGAAAGAC	CGAGTTCGAG	AACGGCAGTG	ACTCCGCCGA	GGTTGCCCGC
Cci3	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGAG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTGCAGCGG
Ced	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTCGCCCGG
1Ceil	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTCGCCCGG
19Cel	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTCGCCCGG
14Ce2	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	GAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTCGCCCGG
7Cc1	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	GAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTAGCCCGG
37At1	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTCGCCCGG
11A11	CCTCGACCAC	AGCGAGCTGT	TCAAGGACGA	GGTCTACCAG	AAGCAGTTCG	AGGGCAAGCG	CGAGTTCGAG	AACGGCAGCG	ACGCCGCCGA	GGTCGCCCGG
	501			540	544					
Aris	GICCICGACT	GGACCCGCGG	GTGGGAGTAC	CGGGAGAAGA	ACTT					
MG59	GTCCTGGAGT	GGACCCGTAC	CTGGGAGTAC	CGCGAGAAGA	ACTT					
RPL161	GTCCTGGAGT	GGACCCGTAC	CTGGGAGTAC	CGCGAGAAGA	ACTT					
1024	GICCIGGAGT	GGACCCGTAC	CTOGGAGTAC	CGCGAGAAGA	ACTT					
CN61	GTCCTGGAGT	GGACCCGTAC	CTCGGAGTAC	CGCGAGAAGA	ACTT					
KP54	GTTCTGGAGT	GGACCCGTAC	CTGGGAGTAC	CGCGAGAAGA	ACTT					
EUNIE	GTCCTGGAGT	GGACCCGCAC	CTGGGAGTAC	CGCGAGAAGA	ACTT					
SCN10A	GTCCTGGAGT	GGACCCGCAC	CTOGGAGTAC	CGGGAGAAGA	ACTT					
HRNISa	GTTCTGGAGT	GGACCCGCGG	CTGGGAGTAC	CGCGAGAAGA	ACTT					
Eal-12	GTCCTCGAGT	GGACCCGCGG	CTOGGAGTAC	CGGGAGAAGA	ACTT					
EaN1-pec	GTCCTCGAGT	GGACCCGCGG	CIGGGAGTAC	CGGGAGAGA	ACTT					
DII	GTTCTCGAAT	GGACCCGCGG	CTGGGAGTAC	CGGGAGAAGA	ACTT					
Cers	GICCIGGAGT	GGACCCGCGG	CTGGGAGTAC	CGGGAGAAGA	ACTT					
Ced	GILCIGGAGT	GGACCCGCGG	CIGGGAGTAC	CUGGAGAAGA	ACTT					
100-1	GICCIGGAGT	GGACCCGCGG	CTOGGAGTAC	CGGGAGAAGA	ACTT					
19001	GIUCIGGAGT	GGACCCGCGG	CIGGGAGTAC	CUCUGAGAAGA	ACTT					
14082	GTCCTGGAGT	GGACCCGCGG	CIGGGAGTAC	CUCUGAGAAGA	ACTT					
273+1	GILCIGGAGT	GGACCCCCCC	CIGGGAGTAC	LUGGAGAAGA	ACTT					
11311	CECCEGAGE	GUALCEGEGG	CIGGGAGTAC	COOGAGAAGA	ACTT					
11411	GICCIGGAGT	JUMILLUGUUG	CIGGGAGIAC	COOCHGAAGA	ACTT					

Figure 2. (Continued).

In the group of *Casuarina/Allocasuarina*-infective strains, up to six differences were observed (corresponding to 0.018 substitutions/site). In the group of Elaeagnaceae-infective strains, between 6 and 71 differences were observed (corresponding to 0.014–0.232 substitutions/site). The *Alnus*-infective strains exhibited between 84 and 175 differences with *Casuarina*-infective strain CcI3 and *Gymnostoma* microsymbiont MG59, respectively.

Two clusters were identified by using the distance matrix (*table II*) and the resulting phylogenetic tree (*figure 3*).

Cluster 1 is a very tight group that contains *Casuarina* and *Allocasuarina* microsymbionts. Two *C. equisetifolia* microsymbionts, 1Cei1 and 19Ce1, have identical sequence with that of the reference *Casuarina* infective strains (*table II*). The remaining *C. equisetifolia* microsymbiont, 14Ce2, is very closely related to the *C. cunninghamiana* microsymbiont. This grouping was also detected by parsimony analysis. *A. torulosa* and *A. littoralis* microsymbionts form a group not confirmed by parsimony analysis. Nevertheless, 37At1 and 11Al1 sequences had a very low level of divergence (*table II*).

The Elaeagnaceae-infective strains form a broad group designated cluster 2. This coherent cluster (100 % of the bootstrap replicates) was also detected by parsimony analysis. The level of sequence divergences was higher in this cluster than in cluster 1. *Gymnostoma* microsymbionts belonged to cluster 2. They form a coherent group (100 % of the bootstrap replicates, parsimony analysis) with EUN1f and SCN10a.

3.3. PCR/RFLP on 16S-23S IGS

Using two restriction enzymes, IGS-types were determined (table III; figure 4). All reference Casuarina and Allocasuarina infective strains have

S366	

Table II. Distance matrix. Calculated distances (expressed as substitutions/100 sites, below diagonal) as described by Kimura (1980) and observed number of differences between pairs of sequences (above diagonal).

	Arl3	RPL16	MG59	TC24	CN61	KP54	EUN1f	SCN10	HRN18	Eal	EaN1	D11 (CcI3 (CeD 1	lCei 1	[9Cel	14Ce2	7Cc1	37At1 :	11A11
ArI3	0	178	175	174	171	147	174	111	105	66	66	119	84	84	84	84	86	87	84	84
RPL161d	573	0	œ	7	22	48	25	25	50	58	59	71	60	60	60	60	61	64	61	62
MG59	563	18	0	9	19	46	24	24	49	54	57	71	61	61	61	61	62	65	62	63
TC24	552	16	14	0	17	43	22	18	48	55	58	71	57	57	57	57	58	61	58	59
CN61	548	52	45	40	0	42	18	21	48	56	59	66	59	59	59	59	60	63	60	61
KP54	487	127	122	113	111	0	46	17	46	55	58	63	58	58	58	58	60	63	59	60
EUN1f	551	59	56	51	42	121	0	21	44	54	56	68	58	58	58	58	60	63	59	60
SCN10A	379	20	68	50	59	48	58	0	42	49	50	58	48	48	48	48	50	53	49	50
HRN18a	392	162	159	155	155	149	141	135	0	37	40	51	46	46	46	46	49	51	47	48
Ea1-12	358	196	182	185	189	186	180	164	121	0	6	49	35	35	35	35	37	39	36	37
EaN1-pec	360	201	194	197	201	198	189	169	133	26	0	49	32	32	32	32	34	36	33	34
D11	431	231	232	230	215	212	219	197	173	168	169	0	50	50	50	50	52	54	53	52
CcI3	297	201	206	188	197	194	192	158	153	113	103	170	0	0	0	0	3	ŋ	n	2
CeD	297	201	206	188	197	194	192	158	153	113	103	170	0	0	0	0	ŝ	ŝ	e	2
1Ceil	297	201	206	188	197	194	192	158	153	113	103	170	0	0	0	0	ŝ	5	ŝ	7
19Ce1	297	201	206	188	197	194	192	158	153	113	103	170	0	0	0	0	ŝ	S	33	2
14Ce2	305	205	210	192	201	202	199	165	164	120	110	178	6	6	6	6	0	4	9	IJ
7Cc1	310	216	222	203	213	214	211	176	172	127	117	186	15	15	15	15	12	0	9	ß
37At1	296	204	210	192	201	198	196	161	156	117	106	182	6	6	6	6	18	18	0	1
11A11	296	208	214	196	205	202	199	165	160	120	110	178	9	9	9	9	15	15	ŝ	0



0.028 s/s

Figure 3. Phylogenetic neighbour-joining tree based on *nifD-nifK* intergenic spacer sequences. The bars represent 0.028 substitutions/site. The numbers are the percentages of bootstrap replicates in which the cluster was found. An asterisk indicates that the group was found by parsimony analysis.

IGS-types	Restricti	on patterns	of the 16	S-23S IGS
	MspI	HaeIII	NcI1	ScrF1
A	M1	H1		
В	M1	H2		
\mathbf{C}	M2	H1		
D	M2	H3		
\mathbf{E}	M3	H3		
\mathbf{F}	M3	H4		
G	M4	H1		
Н	M5	H5		
1			N1	SF1
2			N2	$\mathbf{SF2}$
3			N3	$\mathbf{SF3}$
4			N1	$\mathbf{SF4}$
5			N4	$\mathbf{SF4}$

Table III. IGS-type determination.

similar patterns and were grouped together (IGS-type 1), whereas the *Casuarina/Allocasuarina* microsymbionts from Australia are distributed in five IGS-types (*table I*). *C. equisetifolia* microsymbionts belong to two IGS-types: 1 and 2, the majority being in group 1. IGS-type 3 contains *C. cunninghamiana* microsymbionts. *Allocasuarina* spp. microsymbionts belong to the remaining IGS-types. IGS-type 4 includes *A. torulosa* microsymbionts, whereas IGS-type



Figure 4. PCR/RFLP patterns of the 16S-23S intergenic spacer (IGS) after digestion with *Nci*1, *Scr*F1. Lanes 1 to 5, IGS-types 1 to 5; lane L, 1-kb ladder.

5 includes A. littoralis microsymbionts. The genotypic grouping of the Casuarina and Allocasuarina infective strains was consistently associated with the host plant species (table IVa).

The 35 Gymnostoma microsymbionts were distributed into eight IGS-types (*table I*). For each Gymnostoma species, microsymbionts belong to two or three IGS-types. Conversely, microsymbionts belonging to six IGS-types were associated with several Gymnostoma species. IGS-type D is the most promiscuous group, being found with seven of the eight Gymnostoma species tested.

The grouping of the strains according to molecular criteria was not related to the grouping based on the host plant species, since each species was found to have established naturally a symbiosis with microsymbionts classified in several IGS- types (*table IVb*).

4. DISCUSSION

Coevolution has been found in several host-pathogen systems (Futuyama, 1986) and in highly specific obligate mutualism such as endosymbiosis (Futuyama, 1986; Moran et al., 1993). Coevolution can be either loose or strict, but this must be confirmed by evidence for the congruence of the two partners' phylogenetic trees.

Comparative studies of the phylogenies of host plant and symbionts constitute a promising approach for the elucidation of the evolution of actinorhizal **Table IV.** Distribution of *Frankia* 16S-23S IGS-types as a function of Casuarinaceae species.

IVa. Casuarina and Allocasuarina species.

		IG	S-type	s	
Host plant	1	2	3	4	5
C. equisetifolia C. cunninghamania A. torulosa A. littoralis	+	+	+	+	+

IVb. Gymnostoma species.

				IGS-t	$_{\rm ypes}$			
Host plant	A	В	С	D	E	F	G	Н
G. chamaecyparis				+		+		+
G. deplancheanum		+		+				
G. glaucescens				+	+			
G. intermedium			+	+	+			
$G. \ leucodon$		+	+	+				
$G. \ nodiflorum$	+	+						+
G. poissonianum		+		+			+	
$G.\ webbianum$		+	+	+				

symbiosis. Of the 21 dicotyledonous genera described as actinorhizal (Benson and Silvester, 1993), strains capable of fulfilling Koch's postulates or present as microsymbionts in 11 of these have had their 16S determined and compared (Nick et al., 1992). This analysis has shown that the genus *Frankia* is coherent and that isolated strains infective on *Casuarina* are phyletically close to those infective on *Alnus*. Study of the plant phylogeny, on the other hand, has shown that *Casuarina* and the other Casuarinaceae genera *Allocasuarina* and *Gymnostoma* formed a phyletically coherent family in the Hammamelidae (Maggia and Bousquet, 1994). It was thus expected that the microsymbionts present in the nodules of these three genera would be phyletically close.

The present work on the *nifD-nifK* intergenic spacer has shown that, on the contrary, the nonisolated strains present in *Gymnostoma* nodules and *Casuarina/Allocasuarina* nodules belonged to different clusters. *Casuarina/Allocasuarina* microsymbionts form a tight group with a very low level of sequence divergence (*figure 3*; *table II*). The phylogenetic tree outlined in this cluster is similar to the trees obtained by PCR/RFLP analysis (*figure 5*). *Gymnostoma* microsymbionts were in the cluster of Elaeagnaceae-infective strains. Cross-inoculation studies have confirmed that *Gymnostoma*-infective strains are *Elaeagnus*-infective and not *Casuarina*-infective (Navarro et al., 1997).

Differences between Casuarina/Allocasuarina and Gymnostoma microsymbionts were confirmed by PCR/RFLP analysis of the 16S-23S intergenic spacer (table IV). Genetic diversity among Casuarina and Allocasuarina microsym-



Figure 5. Dendrogram of genotypic relationship of intergenic spacer (IGS) sequences derived by PCR/RFLP analysis according to Rouvier (1995).

bionts seems to be host-species dependent. Cross-inoculation studies support this finding (Reddell and Bowen, 1985; Sellstedt, 1995). This differentiates this group from *Gymnostoma* microsymbionts for which no relation between genetic diversity and host species was observed. Identical results were obtained by cross-inoculation studies (Gauthier, personal communication).

These results mean that in the Casuarinaceae line of descent the host plants did not evolve gradually as proposed by Maggia and Bousquet (1994). These authors suggested that evolution of the symbiosis had been from a promiscuous ancestor identified as comparable to present-day *Gymnostoma* to the very restrictive descendant (Allocasuarina) with Casuarina in between. Instead, the most coherent scenario would be that an early split occurred in the evolution of Casuarinaceae, resulting in two distinct lines of descent. In each of them evolution of the symbiosis has occurred in two different ways. Evolution of the Casuarina/Allocasuarina symbiosis has proceeded towards a greater specificity and specialisation, and is presumably an example of coevolution. These host plants are in symbiosis with hard-to-isolate and slow-growing Frankia strains, suggesting that this symbiosis is becoming obligate. More saprophytic Frankia strains (Nalin et al., 1997), faster growing and easier to isolate, have established a nonspecific symbiotic association with *Gymnostoma*. No evolutionary relationships could be evidenced in this interaction. These hypotheses could be confirmed by comparing the phylogeny of the two symbiotic partners, using sequencing of host plant and microorganism DNA from the same nodule, from a larger sample of Casuarinaceae species.

ACKNOWLEDGEMENTS

Thanks are expressed to J. Briolay (Centre d'analyse moléculaire de la biodiversité) for technical assistance.

REFERENCES

Benson D.R., Silvester W.B., Biology of *Frankia* strains, Actinomycete symbionts of Actinorhizal plants, Microbiol. Rev. 57 (1993) 293–319.

Berry A.M., Torrey J.G., Isolation and characterization *in vivo* and *in vitro* of an actinomycetous endophyte from *Alnus rubra* Bong, in: Gordon J.C., Wheeler C.T., Perry D.A. (Eds.), Symbiotic Nitrogen Fixation in the Management of Temperate Forests, Oregon State University, Corvallis, OR, 1979, pp. 69–83.

Bosco M., Fernandez M.P., Simonet P., Materassi R., Normand P., Evidence that some *Frankia* sp. strains are able to cross boundaries between *Alnus* and *Elaeagnus* host specificity groups, Appl. Environ. Microb. 58 (1992) 1569–1576.

Diem H.G., Dommergues Y., The isolation of *Frankia* from nodules of *Casuarina*, Can. J. Bot. 61 (1983) 2822–2825.

Diem H.G., Dommergues Y., Current and potential uses and management of Casuarinaceae in the tropics and subtropics, in: Schwintzer C.R., Tedpkema J.D. (Eds.), The Biology of *Frankia* and Actinorhizal Plants, Academic Press, Inc., New York, 1990, pp. 317–342.

Diem H.G., Duhoux E., Simonet P., Dommergues Y., Actinorhizal symbiosis biotechnology: the present and the future, in: Durand G., Bodichon L., Florent J. (Eds.), Proceedings of the 8th International Biotechnology Symposium, Société française de microbiologie, Paris, 1988, pp. 984–995.

Felsenstein J., Confidence limits on phylogenies: an approach using the bootstrap, Evolution 39 (1985) 783–791.

Fernandez M.P., Meugnier H., Grimont P.A.D., Bardin R., Deoxyribonucleic acid relatedness among members of the genus *Frankia*, Int. J. Syst. Bacteriol. 39 (1989) 424-429.

Futuyama D.J., The evolution of interactions among species, in: Davis A., Vesely J. (Eds.), Evolutionary Biology, Sinauer Associates, Inc., Sunderland, MA, 1986, pp. 482–504.

Gauthier D., Diem H.G., Dommergues Y., Infectivité et effectivité des souches de *Frankia* isolées de nodules de *Casuarina equisetifolia* et *Hippophae rhamnoides*, C.R. Acad. Sci. 293 (1981) 489–491.

Higgins D.G., Sharp P.M., CLUSTAL: a package for performing multiple sequence alignments on a microcomputer, Gene 73 (1998) 237–244.

Jamann S., Fernandez M.P., Normand P., Typing method for N₂-fixing bacteria based on PCR-RFLP – application to the characterisation of *Frankia* strains, Mol. Ecol. 2 (1993) 17–26.

Johnson L.A.S., Wilson K.L., Casuarinaceae: a synopsis, in: Crane P.R., Blackmore S. (Eds.), Evaluation, Systematics and Fossil History of the Hamamelidae, vol. 2: Higher Hamamelidae, Systematics Association, Claredon Press, London, special vol. 40B, 1989, pp. 67–188.

Kimura M., A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences, J. Mol. Evol. 16 (1980) 111-120.

Lalonde M., Calvert H.E., Pine S., Isolation and use of *Frankia* strains in actinorhizae formation, in: Gibson A.H., Newton W.E. (Eds.), Current Perspectives in Nitrogen Fixation, Australian Academy of Science, Canberra, 1981, pp. 296–299.

Maggia L., Bousquet J., Molecular phylogeny of the actinorhizal Hamamelidae and relationships with host promiscuity towards *Frankia*, Mol. Ecol. 3 (1994) 459–467.

Moiroud A., Faure-Reynaud M., Influences de quelques herbicides à large spectre sur la croissance et l'infectivité de cultures pures de *Frankia*, Plant Soil 74 (1983) 133-136. Moran N.A., Muson M.A., Baumann P., Ishikawa H.I., A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts, Proc. R. Soc. Lond. [Biol.] 253 (1993) 167–171.

Mort A., Normand P., Lalonde M., 2-O-methyl-D-mannose, a key sugar in the taxonomy of *Frankia*, Can. J. Microbiol. 29 (1983) 993-1002.

Nalin R., Normand P., Domenach A.-M., Characterization of a *Frankia* community in a soil column, Physiol. Plant. 99 (1997) 732–738.

Nalin R., Domenach A.-M., Normand P., Characterization of an intergenic spacer between *nifD* and *nifK* in *Frankia*, Mol. Ecol. 4 (1995) 483–491.

Navarro E., Simonet P., Normand P., Bardin R., Characterization of natural population of *Nitrobacter* spp. using PCR/RFLP analysis of the ribosomal intergenic spacer, Arch. Microbiol. 157 (1992) 107–115.

Navarro E., Nalin R., Gauthier D., Normand P., The nodular microsymbionts of *Gymnostoma* spp. are *Elaeagnus*-infective strains, Appl. Environ. Microb. 63 (1997) 1610–1616.

Nazaret S., Cournoyer B., Normand P., Simonet P., Phylogenetic relationships among *Frankia* genomic species determined by use of amplified 16S rDNA sequences, J. Bacteriol. 173 (1991) 4072–4078.

Nick G., Paget E., Simonet P., Moiroud A., Normand P., The nodular endophytes of *Coriara* spp. form a distinct lineage within the genus *Frankia*, Mol. Ecol. 1 (1992) 175–181.

Perrière G., Gouy M., WWW-Query: an on-line retrieval system for biological sequence banks, Biochimie 78 (1996) 364–369.

Ponsonnet C., Nesme X., Identification of *Agrobacterium* strains by PCR-RFLP analysis of pTi and chromosomal regions, Arch. Microbiol. 161 (1994) 300–309.

Racette S., Torrey J.G., The isolation, culture and infectivity of a *Frankia* strain from *Gymnostoma papuanum* (Casuarinaceae), Plant Soil 118 (1989) 165–170.

Reddell P., Bowen G.D., Do single nodules of Casuarinaceae contain more than one *Frankia* strain? Plant Soil 88 (1985) 275-279.

Rouvier C., Caractérisation génétique de *Frankia* nodulant les Casuarinacées : influence de la plante-hôte et de sa zone d'origine, l'Australie, thesis, Université Lyon I, France, 1995.

Rouvier C., Nazaret S., Fernandez M.P., Picard B., Simonet P., Normand P., rrn and nif intergenic spacers and isoenzyme patterns as tools to characterize *Casuarina*-infective *Frankia* strains, Acta Oecol. 13 (1992) 487–495.

Rouvier C., Prin Y., Reddell P., Normand P., Simonet P., Genetic diversity among *Frankia* strains nodulating members of the family Casuarinaceae in Australia revealed by PCR and restriction fragment length polymorphism analysis with crushed root nodules, Appl. Environ. Microb. 62 (1996) 979–985.

Saitou R.R., Nei M., A Neighbor-Joining method: a new method for reconstructing phylogenetic trees, Mol. Biol. Evol. 44 (1987) 406-425.

Savouré A., Lim G., Characterization of an infective Frankia (ISU 0224887) isolated from nodules of Gymnostoma sumatranum, Plant Soil 131 (1991) 21–27.

Sellstedt A., Specificity and effectivity in nodulation by *Frankia* on southern hemisphere actinorhiza, FEMS Microbiol. Lett. 125 (1995) 231–236.

Swofford D.L., PAUP - phylogenetic analysis using parsimony, version 3.1, Illinois Natural History Survey, Champaign, IL, 1993.

Winship P.R., An improved method for directly sequencing PCR amplified material using dimethyl sulfoxide, Nucleic Acids Res. 17 (1989) 1266.

Zhang Z., Lopez M.F., Torrey J.G., A comparison of cultural characteristics and infectivity of *Frankia* isolates from root nodules of *Casuarina* species, Plant Soil 78 (1984) 79–90.