Nopp140 Shuttles on Tracks between Nucleolus and Cytoplasm

U. Thomas Meier and Günter Blobel Laboratory of Cell Biology Howard Hughes Medical Institute The Rockefeller University New York, New York 10021

Summary

Nopp140 is a nucleolar phosphoprotein of 140 kd that we originally identified and purified as a nuclear localization signal (NLS)-binding protein. Molecular characterization revealed a 10-fold repeated motif of highly conserved acidic serine clusters that contain an abundance of phosphorylation consensus sites for casein kinase II (CK II). Indeed, Nopp140 is one of the most phosphorylated proteins in the cell, and NLS binding was dependent on phosphorylation. Nopp140 was shown to shuttle between the nucleolus and the cytoplasm. Shuttling is likely to proceed on tracks that were revealed by immunoelectron microscopy. These tracks extend from the dense fibrillar component of the nucleolus across the nucleoplasm to some nuclear pore complexes. We suggest that Nopp140 functions as a chaperone for import into and/or export from the nucleolus.

Introduction

The nucleolus is the site of ribosomal RNA synthesis and preribosome assembly. Ribosomal proteins are synthesized in the cytoplasm and imported into the nucleolus, where they are assembled into preribosomal particles that in addition contain nonribosomal proteins (Kumar and Warner, 1972; Prestayko et al., 1974; Hügle et al., 1985). Upon export into the cytoplasm, the preribosomal particles lose the nonribosomal proteins and mature into functional ribosomal subunits (for a review of ribosome biogenesis see Warner, 1989).

Previously we identified and purified a nuclear localization signal (NLS)-binding protein of 140 kd (p140) that was primarily localized to the nucleolus (Meier and Blobel, 1990). In a ligand blot assay we showed that this protein binds specifically to a synthetic peptide representing the NLS of SV40 T antigen (Kalderon et al., 1984a, 1984b; Lanford and Butel, 1984) but not to a peptide of an importincompetent mutant NLS. Based on its NLS-binding ability and nucleolar localization, we suggested that p140 may shuttle between the nucleolus and the cytoplasm.

Other nuclear proteins have been shown to shuttle between nucleus and cytoplasm (Rechsteiner and Kuehl, 1979; Goldstein and Ko, 1981; Madsen et al., 1986; Bachmann et al., 1989; Borer et al., 1989; Mandell and Feldherr, 1990; Guiochon-Mantel et al., 1991; Pinol-Roma and Dreyfuss, 1992). Some of these proteins may function in nucleocytoplasmic transport, i.e., in import into and/or export from the nucleus (for a recent review of nucleocytoplasmic transport see Nigg et al., 1991). In addition, several groups have identified NLS-binding proteins by various methods (Adam et al., 1989; Benditt et al., 1989; Lee and Mélèse, 1989; Li and Thomas, 1989; Silver et al., 1989; Yamasaki et al., 1989; Imamoto-Sonobe et al., 1990; Stochaj et al., 1991). Of these only the yeast nuclear protein NSR1 has been characterized on a molecular level (Lee et al., 1991). Moreover, two NLS-binding proteins of bovine erythrocytes have been shown to stimulate cytosol-dependent nuclear import in digitonin-permeabilized cells (Adam and Gerace, 1991).

In this article, we present the primary, cDNA-deduced, structure of p140. Because of its predominant nucleolar localization and its high degree of phosphorylation we named it Nopp140 (nucleolar phosphoprotein of 140 kd). Nopp140 shuttles between the nucleus and the cytoplasm and appears to do so on striking curvilinear tracks emanating from the dense fibrillar component of the nucleolus.

Results

Cloning and Nucleotide Sequence Analysis

To generate probes for cloning the cDNA of Nopp140, the purified protein (Meier and Blobel, 1990) was subjected to chemical and proteolytic cleavage and to amino-terminal sequencing. Degenerate oligonucleotides designed according to the peptide sequences were used in a polymerase chain reaction with rat cDNA to generate a specific nucleotide probe for Nopp140. Screening rat cDNA libraries with this probe resulted in the identification of two species of cDNA, pTM17 and pTM6 (schematically depicted in Figure 2A), that shared a virtually identical open reading frame but differed in their 5'- and 3'-untranslated regions (UTRs). Figure 1 delineates the nucleotide sequence and the predicted amino acid sequence of the cDNAs pTM17 (A) and pTM6 (B). The coding sequences were identical except for the CAG triplet encoding glutamine 150 in pTM17 (Figure 1A), which was absent in pTM6 as confirmed by sequencing of several independent clones from two different cDNA libraries. Interestingly, these three nucleotides constitute the highly conserved 3' splice site of vertebrate introns (Padgett et al., 1986) and could, therefore, represent part of an intron in pTM17 that was spliced out in the case of pTM6. The difference in length of the 3'-UTRs of the two cDNAs can readily be explained by alternative use of polyadenylation sites. Thus, clone pTM6 contained a less used site that was skipped in favor of the most common site in pTM17 (Figure 1A, shaded boxes; Wickens and Stephenson, 1984).

To demonstrate that these two species of cDNA represented cellular mRNAs, Northern blot analysis was performed by hybridizing cDNA-specific probes to rat and human poly(A)⁺ RNA (Figures 2A and 2B). The two detected bands of approximately 3.6 and 3.2 kb (Figure 2B) corresponded well to the size of the clones pTM17 and pTM6, respectively, indicating that the isolated clones represented full-length mRNAs. The total amount and the ra-

A pTM17

| - | | | |
 | | | ÷.,
 | | | 2 | inc | 110
 | AGT | | ncc. | 666
 | iluc. | UC A | ACG |
 | 610 | MC G | | 1114
 | ACC | COUNCI | |
|--|--|--|---
---|--|--
--|---|---|---|--
--|--
---|---|---
---|--|---|--
---|--|---
--|---|--|---|
| | - | - | |
 | - | |
 | · _ | | | |
 | | | |
 | 20 | | | _
 | | | |
 | | 30 | |
| MAD | т | G | r. | R
 | R | v | v
 | Р | s | D | r | Ŷ
 | P | L | v | L
 | G | F | ι | R
 | D | N | Q | L
 | s | EV | |
| ATGGCGGA | TACC | eec | TTG | CCC
 | cec | GTC | GGT1
 | tcco | ACC | GAC | CTT | TAT
 | CCC | CTTO | STG | стс
 | GGC | TTT | CTG | CGA
 | GAT | AAC | CAG | CTC
 | TCA | GAGGTO | : 14: |
| | | | |
 | | | 40
 |) | | | |
 | | | |
 | 50 | | |
 | | | |
 | | 60 | |
| ASK | F | A | ĸ | A
 | т | G | A
 | 7 | a | 0 | D | ٨
 | N | | s | s
 | τ. | 1. | n | T
 | ۷ | s | F | H
 | t | ĸs | |
| GCCAGTAA | ATTT | GCA | AAA(| CG
 | ACA | 000 | coca
 | AC. | CAG | CAG | GAC | acc
 | | GCC- | тст | тсс
 | CTC | тто | cac. | A TT
 | тат | acc | | TCC
 | стс | AACTCC | |
| | | | |
 | | | 20
 | \
\ | | | |
 | | | |
 | | | 0110 |
 | | | |
 | | 0.00100 | |
| | | | |
 | | |
 | · | | | |
 | | | - 1 |
 | | | |
 | | | |
 | | 90 | |
| TKA | P | ĸ | v | ĸ
 | ÷. | Q | \$
 | N | G | P | v | •
 | <u>κ</u> . | <u>к</u> | Δ. | _ K_
 | ĸ | E | т | s
 | s | 5 | D | 5
 | 5 | ЕD | |
| ACCARAGC | CCCG | AAG | GTG/ | ***
 | CTG | CAC | GTC
 | | rcc A | CCA | 07 6 | GCC
 | *** | AAGO | GC T. | AAG
 | AAA | GAG. | ACT | TCA
 | TCC | AGT | GAC | AGC
 | AGT | GAGGAG | 32 |
| | | | |
 | | | 100
 | > | | | |
 | | | |
 | 110 | | |
 | | | |
 | | 120 | |
| SSE | ю | к | D | к
 | ۰ | 0 | v
 | Р | т | 0 | к | А
 | | | Р |
 | ĸ | R | A | s
 | 1. | P | 0 | н
 | | G . K | | |
| ACCACTON | | | |
 | | ~ |
 | | | ~~~ | | ~~~
 | ~~~~ | | |
 | | | |
 | | ~~~ | - ¥- |
 | ~~~~ | | |
| Auchoron | 0000 | 040 | Unur | ~~~
 | u.c. | C.A. | 1 1 1 1
 | L.C.C. | | C AU | nno | 0¢ 1
 | occi | occi | | ucc
 | | con | ucci | H G 1
 | 110 | cer | CAG | CAI
 | GCI | GOGAAA | 41. |
| | | | |
 | | | 130
 | | | | |
 | | | |
 | 140 | | |
 | | | - |
 | | 150 | |
| | ĸ | A | s | Е
 | s | \$ | s
 | s | E | ε | s | s
 | E | E | 3 | Е
 | ε | ĸ | D | ĸ
 | ĸ | ĸ | κi | ۶
 | ۷ | 0:0 | |
| GCAGCAGC | CAAA | GCT | TCAC | SAG
 | AGC | AGC | CAGI
 | AGT | GAA | GAG | TCC | AGT
 | GAC | GAAC | JAG | GAG
 | GAG | AAG | GAC | ÀAĀ
 | AAG | 242 | AAG | сст
 | атс | CAGCAG | 50: |
| | | | |
 | | | 160
 | | | | |
 | | | |
 | 170 | | |
 | | | |
 | | 180 | |
| ¥ 8 V | × | D | ^ |
 | v | |
 | | | | 6 |
 | | | |
 | | | | P
 | | | ~ | ~
 | ~ | | | |
| | | | |
 | | <u>.</u> |
 | | . . . | | |
 | | ^ | ÷. | 9
 | 2 | E | • | E.
 | 2 | D | 2 | 2
 | 2 | E D | |
| AAAGCAGT | TAAG | cee | CAAC | seci
 | AAG | GCA | NGT C
 | AGA | CCT | сст | CCG | AAG
 | AAG | SCAC | GAG. | AGC
 | TCT | GAG | tcc | GAG
 | TCT | GAC | TCA | ACC
 | TCA | CAGGAT | 592 |
| | | | |
 | | | 190
 | • | | | |
 | | | |
 | 200 | | |
 | | | |
 | | 210 | |
| EAP | 0 | T | 0 | ĸ
 | P | ĸ | ٨
 | ٨ | ٨ | т | à | ٨
 | ĸ | ٨ | Ρ | т
 | к | ٨ | 0 | 7
 | к | А | Р | A
 | ж | PG | |
| GAAGCACC | 1010 | aco. | CACI | ACC
 | rea. | |
 | 000 | CCT | ACC | | CCT
 | | 0000 | - | 1.CT
 | | con | | 107
 | | | ~~~ | ~~~~
 | | - | 60. |
| | | | |
 | | | 720
 | | | ALU | uun |
 | | | | ne i
 | 220 | | | nc 3
 | ~~~ | occ | CCA | occ
 | ~~~ | 240 | 00. |
| | | | | -
 | | | 220
 | · | | | |
 | | | |
 | 230 | | |
 | | | |
 | | 290 | |
| 6 6 ¥ | ĸ | | â | P
 | ĸ | ٨ | ۸
 | N | G | ĸ | A | G
 | s | s | s | s
 | s | \$ | s | \$
 | s | s | s | D
 | D | SE | |
| CCACCAGO | GAAA | GC A | CAGO | сти
 | *** | GCA | AGCC
 | TAA: | GGC | AAA | GCA | GGC.
 | AGC | AGC/ | ACC. | AGC
 | AGT | AGCI | AGC | AGC
 | AGT | AGC | AGT | GAT
 | GAC | TCAGAG | 223 |
| | | | |
 | | | 250
 | | | | |
 | | | |
 | 260 | | |
 | | | |
 | | 270 | |
| | ĸ | | | ۵
 | Þ | ١. |
 | × | т | | D | ĸ
 | | 0 | v |
 | | | |
 | v | | ., | Ŧ
 | | | |
| | | <u>.</u> | <u> </u> | .c
 | · · | | ^
 | | | <u>.</u> | | . n.
 | <u>^</u> | Υ. | ۷. | •
 | . . . | | • | r
 | ٠ | | v | T
 | ^ | A P | |
| GAAGAGAA | GAAG | nc.A | ac re | ac ne
 | CT | сте | AAG
 | AAC | ACT | CCN | сст. | ***
 | AAGO | CARO | stee | CTC
 | ceci | AAG | CAC | CCA
 | GTA. | *** | CTA | ACT
 | CCT | secce | 863 |
| | | | |
 | | | 280
 | | | | |
 | | | |
 | 290 | | |
 | | | |
 | | 300 | |
| TQK | s | s | s | s
 | E | D | s
 | s | s | Е | Е | Е
 | E | Е | 0 | к
 | к | P | м | к
 | ĸ | ĸ | A | G
 | P | YS | |
| ACCCARAR | GAGT | TCT. | AGCA | GTO
 | SAGO | GAC | TOT
 | TCC | ACT | | GAG | GAN
 | - | 1880 | - |
 | | rcci | TG | 110
 | | | CC N | 001
 | cer | LACACT | 051 |
| | | | |
 | | | 210
 | | | | 0.10 |
 | | | |
 | 224 | | |
 | | | 0.0 |
 | | | ,,,, |
| | - | - | |
 | | | 310
 | | | | |
 | | | |
 | 320 | | |
 | | | |
 | | 330 | |
| SVP | P | Р | \$ | v
 | s | L | . s
 | ĸ | ĸ | s | v | G
 | ۸. | Q | s | ₽
 | ĸ | ĸ | A | A
 | ×. | Q | 7 | Q
 | ₽ | A D | |
| TCAGTTCC. | ACCA | CCT | 1010 | 5773
 | ICT: | TTA | ATCC
 | AAA | AAG | TCC | GTG | GGA
 | SCCO | CAGT | CTO | CCA
 | AAG | AAA | CGG | SCC
 | GCGI | CAA | ACA | CAG
 | ссто | SCAGAC | 1042 |
| | | | |
 | | | 340
 | | | | |
 | | | |
 | 350 | | |
 | | | |
 | | 360 | |
| | п | • | • | e.
 | F | • | D
 | c | e | e | P |
 | F | • | • | T
 | n . | | ~ | *
 | | | | ~
 | T | | | |
| | | | |
 | | |
 | | | | |
 | | · · | <u>.</u> |
 | | <u> </u> | | . ÷
 | | • | 3 | ~
 | | <u>د</u> | |
| ACCAGIGC | AGACI | NGC. | AGCG | w.or
 | 3401 | rer | GAT
 | TCA | AGT | rer | GAG | GAA
 | SAGA | MGA | ~~~ | ACT
 | CCA | 5C.1/ | AG. | ACA
 | GIC | CLC. | TCC. | AAG.
 | ACAG | CCCCCC | 1132 |
| | | | |
 | | | 370
 | | | | |
 | | | |
 | ~ | | |
 | | | |
 | | 200 | |
| | | | |
 | | |
 | | | | |
 | | | |
 | 380 | | |
 | | | |
 | | 390 | |
| кра | p | v | ĸ | ĸ
 | ĸ | A | E
 | s | s | s | D | s
 | 2 | D | s | Ъ
 | 380
S | s | E | D
 | Е | ٨ | р |
 | к | r v | |
| K P A | P
TCCA | otg. | K
NAGA | ж
(АА)
 | K | A | E
 | s
AGC | s
TCT | S
TCA | D
GAC | s
AGC*
 | 2
7627 | D | s
ICTO | D
 | 380
\$
AGT1 | s
tere | E | D
 | E | A
CCT | е | A
 | K | P V
CAGTO | 1 2 2 2 |
| AAACCAGC | P | V
STG | K
NAGA | K
NAJ
 | K
NAGO | A
SCC | E
GAG
 | S
AGC | s
TCT | S
TCA | D
GAC | s
AGC"
 | e
ICAO | D
ATT | s
CTC | DAC
 | 380
S
AGT:
A10 | s
ICTO | e
Sago | D
 | E
GAA | A
GCT | е | A
GCC
 | K
NAGO | CAGTO | 1222 |
| AAACCAGC | P
TCCA | STG. | NAGA |
 | K
AAGO | A
GCC | E
GAG
400
 | s
AGC | s
тст | s
TCA | D
GAC | s
AGC"
 | S
TCAC | DATT | s
cto | DAC
 | 380
S
AGT
410 | s
ICTO | E
SAGO | D
SAT
 | GAN | A
GCT | ест | A
GCC
 | K
NAGO | CAGTC
420 | 1222 |
| AAACCAGC
S A T | r
K | STG. | K
NAGA
P | к
(Лл)
[.
 | K
AAGO
S | sec
K | E
GAG
400
P
 | s
AGC | тст
v | тсл
т | D
GACI
P | s
AGC'
K
 | s
rcac
P | D
DATT
P | s
CTC | D
DAC
 | 380
S
AGT
410
K | s
icto
A | E
SAGO
V | D
SAT
 | GAAG | A
GCT
P | е
сст
к | a
GCC.
 | K
NAGO
P | CAGTC
420 | 1222 |
| K P A
AAACCAGC
S A T
AGTGCCAC | P
TCCAI
K
CAAGI | STG. | R
AAGA
P
CCCT | K
L
L
 | K
AAGO
S
AGCJ | A
GCC
K | E
GAG
400
P
 | S
AGC
A
GCT | s
TCT
v
GTC | S
TCA
T
ACT | D
GACI
P
CCTI | S
AGC"
K
AAG
 | S
TCAC
P
CCGC | P
CTG | S
CTO
A | A
B
B
B
C
A
 | 380
S
AGT
410
K
AAG | S
TCTO
A
SCAO | E
SAGO
V
STGO | D
SAT
A
SCA
 | E
T
ACTO | A
GCT
P
CCT. | R
R
R | A
GCC.
 | R
NAGO
P
CCTO | ECAGTC
420
A G | 1222 |
| K P A
AAACCAGC
S A T
Agtgccac | P
TCCAI
K
CAAGI | STG. | R
AAGA
P
CCCT | K
L
L
 | K
AAGO
S
AGCJ | A
GCC
K | E
GAG
400
P
ICCA
430
 | S
AGC
A
GCT | s
TCT
V
GTC | S
TCA
T
ACT | D
GACI
P
CCTI | S
AGC"
K
AAG
 | S
FCAC
P
CCGC | P
CTG | s
ICTO
A | D
SACI
A
SCAI
 | 380
S
AGT:
410
K
AAG(
440 | S
ICTO
A
SCAO | E
SAGO
V
STGO | D
SAT
A
SCA
 | E
GAAG
T
ACTO | A
GCT
P
CCT. | P
CCT
K
AAG | A
GCC.
Q
CAG
 | K
NAGO
P
CCTO | CAGTC
420
A G
CGCGC
450 | 1222 |
| K P A
AAACCAGC
S A T
AGTGCCAC | P
TCCAI
K
CAAGI
K | STG.
S
NGTO
P | P
CCCT | K
L
TAJ
 | K
AAGO
S
AGC <i>I</i>
R | A
SCC
K
NAG | E
GAG
400
P
GCCA
430
A
 | A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A | s
tct
v
GTC
s | S
TCA
T
ACT
S | D
GAC
P
CCT | s
AGC"
K
AAGI
 | P
CCGC
E | D
DATT
P
CTG | s
icto
a
icto
e | D
DACI
A
SCAI
 | 380
S
AGT
410
K
AAG
440
S | S
A
SCAC | E
SAGO
V
STGO | D
SATI
A
SCA
 | E
GAAG
T
ACTO | A
CCT
P
CCT. | R
K
AAG | GCC.
CAG
 | R
AAGO
P
CCTO
X | A G
CAGTC
420
A G
CGGGC
450
K S | 1222 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCA | P
TCCAI
K
CAAGI
K | STG. | AAGA
P
CCCT | L
TAN
 | K
AAGO
S
AGC/
R | A
SCC
K
AAG | E
GAG
400
P
GCCA
430
A
 | AGC
AGCT | S
TCT
V
GTC | S
TCA
T
ACT
S | D
GAC | S
AGC"
K
AAGG
S
 | P
CCGC
E | P
CTG | s
icto
A
icto
E | D
SACI
A
SCAI
 | 380
S
AGT:
410
K
AAGC
440
S | SCAC | E
SAGO
V
STGO | D
SATI
A
SCA
 | E
GAAG
T
ACTO | A
P
CCT.
E | K
AAG |
 | R
NAGO
P
CCTO
X | A G
CGGGGC
420
A G
CGGGGC
450
K S | 1222 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCA | P
TCCAI
K
CAAGI
K
GAAA | STG.
S
NGTI
P
CCTI | AAGA
P
CCCT | L
TAJ
 | K
AAGO
S
AGCJ
R
AGAJ | A
GCC
K
AAG
K | E
GAG
400
P
SCCA
430
A
SGCT
 | AGC
A
GCT
D
GAC | S
GTC
S
AGC | TCA
T
ACT
S
AGC | BAC | S
AGC"
K
AAGC
S
AGCO
 | P
CCGC
E
CAGC | P
CTG
E
SAGG | S
A
CTO
E
AG | A
BCAL
S
AGCI
 | 380
S
AGT
410
K
AAG(
440
S
AGC1 | S
CAC | E
SAGO
STGO
STGO | D
SATI
A
SCA
E
SAGI
 | E
GAAG
T
ACTO
E
GAAG | A
CCT
CCT.
E
SAG | R
AAG | ACCI
 | K
P
CCTO
X
AAGA | AAAGT | 1222
1312
1402 |
| K P A
AAACCAGC
S A T
Agtgccac
S G Q
Agtggccad | P
TCCAI
K
CAAGI
K
SAAA | STG.
S
NGTI
P
CCTI | P
CCCT | L
TAJ
 | K
AAGO
S
AGCJ
R
AGAJ | A
GCC
K
AAG
K | E
GAG
400
P
5CCA
430
A
5GCT
460
 | A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A | S
TCT
V
GTC
S
AGC | S
TCA
T
ACT
S
AGC | D
GACI
P
CCTI
S
TCCI | S
AGC"
K
AAGG
S
AGCG
 | S
P
CCGC
E
SAGC | P
CCTG
E
GAGG | S
A
ICTO
E
AGJ | A
BCAL
S
NGCI
 | 380
S
AGT
410
K
AAGC
440
S
AGC
470 | S
CAC | E
SAGO
STGO
STGO | D
SATI
A
SCA
E
SAG
 | Е
Т
АСТО
Е
БАЛО | A
CCT
CCT.
E
SAG | R
AAG
AAG | A
CAGI
T
ACCI
 | R
AAGO
P
CCTO
X
AAGO | P V
CCAGTC
420
A G
CCGGGC
450
K S
AAAGT
480 | 1222
1312
1402 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T | P
TCCAI
K
CAAGI
K
SAAA | STG.
S
AGTI
P
CCTI
K | AAGA
P
CCCT
Q
CAGA | K
L
TAJ
S
IGCJ
R
 | K
AGCJ
R
AGAJ
V | A
K
K
K
K
K
K | E
GAG
400
P
5CCA
430
A
5GCT
460
A
 | A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A | S
TCT
V
GTC
S
AGC | S
TCA
T
ACT
S
AGC
A | D
GAC
P
CCT
S
TCC
P | S
AGC"
K
AAGG
S
AGCG
S
 | S
P
CCGC
E
SAGC | P
CTG
E
SAGG | S
A
ICTO
E
IAGJ | D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
 | 380
8
AGT
410
K
AAGC
440
S
AGC
470
Q | S
CAC
S
CAC
S
CAC | E
SAGO
STGO
STGO
S
GTO
P | D
SATI
A
SCA
E
SAGI
R
 | E
GAAG
T
ACTO
E
GAAG | A
CCT.
CCT.
E
SAG | K
AAG
GCCi
G | A
CAGI
T
ACCI
D
 | K
P
CCTO
X
AAG/ | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S | 1222
1312
1402 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC | P
ICCA
K
CAAGI
K
SAAA
P
CCCTI | NGTG | AAGA
P
CCCT
Q
CAGA | K
L
TAJ
S
GCJ
R
 | K
AGCJ
R
AGAJ
V
JTGJ | A
K
AAG
K
AAG
7 | E
GAG
400
P
GCCA
430
A
GCT
460
A
SGCC
 | A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A | S
GTC
S
AGC | TCA
T
ACT
S
AGC
AGC | P
CCTI
S
TCCI
P
CCCI | S
AGC ⁴
K
AAGC
S
AGCC
S
TCTO
 | P
CCGC
E
EAGC
L | P
CTG
E
SAGG | A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A | A
BCAL
S
A
A
CAL
S
A
A
CAL
S
A
A
CAL
 | 380
AGT:
410
K
AAG0
440
S
AGC1
470
Q
CAG0 | S
CAC
S
CAC
S
CAC | E
SAGO
STGO
STGO
STGO
STGO | A
SCA
E
SAG
 | E
ACTO
E
SAAC | A
P
CCT.
E
SAG | K
AAG
ACCI
GCCI
GGA | A
CAGI
T
ACCI
D
GACI
 | R
P
CCTO
X
AGO
S | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S | 1222
1312
1402
1492 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC | P
R
CAAGI
K
SAAA
P
CCCTI | STG.
S
NGTI
P
CCTI
K | R
AAGA
P
CCCT
Q
CAGA
A
SCCA | L
TAJ
S
IGCJ
R
 | K
AGCJ
R
AGAJ
V
STGJ | A
SCC
K
AAG
K
AAG
7
ACC | E
GAG
400
P
5CCA
430
A
5GCT
460
A
5GCC
490
 | A
GCT
D
GAC
K
AAA | S
GTC
S
AGC
AGC | TCA
T
ACT
S
AGC
AGC | P
CCTI
S
TCCI
P
CCCI | S
AGC ⁴
K
AAGC
S
AGCC
S
TCTO
 | P
CCGC
E
SAGC
L | P
CTG
E
SAGG | A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A
A | A
BCAL
S
A
A
BCAL
S
A
A
C
A
A
C
A
 | 380
S
AGT:
410
K
AAGC
440
S
AGC1
470
Q
CAGC
500 | A
BCAG
S
CCJ | E
SAGO
STGO
STGO
STGO
STGO
STGO | A
BCA
E
SAGI
R
CGGI
 | E
ACTO
E
GAAO
ACTO | P
CCT.
E
SAG | GCTI
K
AAGI
GCCI
GGAI | A
CAG
T
ACCI
D
GACI
 | R
P
CCTO
X
AGC/
S
AGC/ | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S
GCTCC
510 | 1222
1312
1402
1492 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC | P
TCCAI
K
CAAGI
K
SAAAI
P
CCCTI | STG.
S
NGTI
P
CCTI
K
NAGO | AAGA
P
CCCT
Q
CAGA
A
SCCA | K
L
TAJ
S
IGCJ
R
IGGO
 | K
AGCJ
R
AGCJ
R
AGAJ
V
STGJ
E | A
3000
K
MAG
K
MAG
7
MCC
E | E
GAG
400
P
5CCA
430
A
5GCT
460
A
5GCC
490
F
 | A
GCT
D
GAC
K
AAA | S
TCT
V
GTC
S
AGC
AGC | TCA
T
ACT
S
AGC
AGC
A
GCA | P
CCTA
S
TCCA
P
CCCC | S
AGC ⁴
AAGC
S
AGCC
S
TCTO
 | E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGCC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGCC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
CCGC
E
C
C
C
C | P
CTG
E
SAGG
P
CTG | S
ICTO
E
BAGJ
A
ICCJ | A SCAL
 | 380
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
8
410
8
8
410
8
8
410
8
8
410
8
8
410
8
8
8
8
8
8
8
8
8
8
8
8
8
8
8
8
8
8
8 | S
CAC | E
SAGO
STGO
S
S
GCTO
K | D
SATI
A
SCA
E
SAGI
R
CGGI
 | E
ACTO
E
SAAC | A
CCT.
E
CCT.
E
AG | GCCT
AAG
AAG
GCCT
GCCT | A
CAGI
T
ACCI
D
GACI
 | K
P
CTC
X
AGC/
V | 200
200
200
200
200
200
200
200
200
200 | 1222
1312
1402
1492 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E | P
CAAGI
K
CAAGI
K
CAAGI
F
CCCTI | S
S
AGTI
P
CCTI
K
AAGG | AAGA
P
CCCT
Q
CAGA
A
SCCA | K
L
TAN
S
NGC/
R
NGGO
S
 | K
AGCJ
R
AGAJ
V
STGJ
E | A
SCC
K
AAG
K
AAG
7
CC
E | E
GAG
400
P
ICCA
430
A
IGCT
460
A
GCC
490
E
 | S
A
GCT
D
GAC
K
AAA
K | S
GTC
S
AGC
AGC
K | S
TCA
T
ACT
S
AGC
A
GCA | P
CCTA
S
TCCA
P
CCCC | S
AGC ⁴
K
AAGC
S
AGCC
S
TCTC
 | E
E
E
E
A
G
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | P
CTG
E
SAGG
P
CTG | S
ICTO
A
ICTO
E
IAGJ
A
ICCJ
P | D
SACI
SCAI
SCAI
SCAI
SCAI
SCAI
SCAI
SCAI
 | 380
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
40
8
4 | S
SCAC
SCAC
SCAC
SCAC | E
SAGO
STGO
STGO
STGO
STGO
STGO
K | D
SATI
SCA
E
SAGI
R
CGGI
A
 | E
ACTO
E
SAAC | A
SCT
P
SCT.
E
SAG | GCCT
K
AAG
GCCC
GGA
A | ACCI
D
CAGI
 | K
P
CCTO
X
AGC/
S
AGC/
V | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S
NGCTCC
510
P K | 1222
1312
1402
1492 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCA
V T T
GTGACAAC
D S E
GACTCAGA | P
R
CAAGI
K
SAAAA
P
CCCTI
SAGT | STG.
SAGTI
P
CCTI
K
AAGG | AAGA
P
CCCT
CAGA
A
SCCA
S
AGCA | K
L
TAJ
S
GCJ
R
GCGC
S
GTC
 | K
AGCI
R
AGCI
R
AGAJ
V
TTGJ
E
SAGC | A
SCC
K
AAG
K
AAG
7
ACC
E
SAG | E
GAG
400
P
GCCA
430
A
GCC
490
E
GCC
490
E
 | S
AGCT
D
GAC
K
AAA
K | S
GTC
S
AGC
AGC
AGC
A
AGC
A | S
TCA
T
ACT
S
AGC
A
GCA
T
ACG | P
GACI
P
CCTI
S
TCCI
P
CCCI
P
CCCI | S
AGC
K
AAGC
S
AGCC
S
TCTO
P
CCT/
 | E
E
E
E
A
G
A
G
A
A
A
A
A | P
CTG
E
SAGG
P
CTG | S
ICTO
A
ICTO
E
IAGJ
A
ICCJ
P
ICCO | D
SACJ
SCAJ
SCAJ
K
AGCI
A
A
SCTJ
 | 380
8
AGT
410
K
AAGC
410
S
AGC
410
S
C
AGC
500
K
AGC
500
K
AGJ | S
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC | E
SAGO
STGO
STGO
STGO
STGO
STGO
KAGO | D
SATI
SCA
E
SAGI
R
CGGI
A
SCAI
 | E
GAAG
T
ACTO
E
SAAG
SCTO
A
SCAG | A
SCT
P
CCT.
E
SAG
G
SGT
G
SGT | P
K
AAG
GCCi
GGA
GGA | A
GCCC
CAGI
T
ACCC
GACI
A
GCCC
 | K
AAGO
K
AAGJ
S
AGC/
V
STTC | 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 | 1222
1312
1402
1492
1582 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCA
V T T
GTGACAAC
D S E
GACTCAGA | P
R
CAAGI
K
SAAA
P
CCCTI
SAGT | V
STG
S
MGT
P
CCTI
K
MAG | AAGA
P
CCCT
CAGA
A
SCCA
S
SCCA | K
L
S
S
GCJ
R
GGG
S
 | K
AAGG
S
AGCJ
R
AGAJ
V
UTTGJ
E
SAGG | A
SCC
K
AAG
K
AAG
7
ACC
E
SAG | E
GAG
400
P
iccA
430
A
iGCT
460
A
GCC
490
E
iGAG
520
 | S
AGC
A
GCT
GAC
K
AAA
K
AAG | S
TCT
V
GTC
S
AGC
A
GCA
K
AAG | S
TCA
T
ACT
S
AGC
A
GCA
T
ACG | P
GACI
P
CCCTI
S
TCCI
P
CCCC
P
CCCC | S
AGC
K
AAGC
S
AGCC
S
TCTO
P
CCTO
 | S
TCAC
P
CCGC
E
SAGC
L
CTAC
K
K
AAAAC | P
P
CCTG
E
GAGG
P
CCTG
P
CCCC | s
icto
E
BAGJ
A
CCC | D
SACJ
SCAJ
SCAJ
SCAJ
K
AAAC
A
SCTJ
 | 330
S
AGT
410
K
AAGC
410
S
AGC
500
K
AGC
500
K
AGJ
530 | S
TCTC
A
SCAC
S
TCCJ
A
SCTC
K
A
GCTC | E
SAGO
STGO
S
GTC
P
CCTC
K
KAGO | D
SATI
SCA
SAGI
R
SGGI
A
SCAI
 | E
GAAG
T
ACTO
E
GAAG
A
GCTO
A
GCTO | A
SCT
P
CCT.
E
SAG
SGT
G
SGT | R
K
AAG
GCC
G
GGA
GGA | A
GCC.
CAGI
T
ACCI
D
GACI
A
SCCC
 | R
AAGO
P
CCTO
X
AAGJ
S
S
AGC/
V
STTO | CAGTC
420
A G
300GGC
450
K S
480
S S
480
S S
510
P K
200AA | 1222
1312
1402
1492
1582 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGA
P T P | P
CAAGI
K
SAAAA
P
CCCTI
SAGT:
V | V
STG.
S
NGTI
P
CCTI
K
NAG
S
TCCI
X | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
S
S
AGCA
K | к
L
ТАЛ
S
IGC/
R
IGC/
S
IGT(
A
 | K
AAGG
S
AGCJ
R
AGAJ
V
TTGJ
E
SAGG
A | A
SCC
K
AAG
K
AAG
7
ACC
E
SAG
A | E
CGAG
400
P
CCA
430
A
GCCT
460
A
CGCC
490
E
CGAG
520
E
 | S
AGCT
D
GAC
K
AAAA
K
AAAG
S | S
TCT
V
GTC
S
AGC
A
AGC
K
AAG.
S | S
TCA
T
ACT
S
AGC
A
GCA
T
ACG | P
GACI
P
CCCTI
S
TCCI
P
CCCC
S | S
AGC
K
AAGC
S
AGCC
S
TCTO
P
CCTO
S
 | S
TCAC
P
CCCGC
E
SAGC
L
CTAC
K
K
ANAAC
S | D
DATT
P
CCTG
E
GAGG
P
CCTG
P
CCCC
S | s
icto
E
SAGJ
A
CCC
S | D
SACI
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
 | 380
S
AGT:
410
K
AGC
440
S
AGC
470
Q
CAGC
500
K
NAGJ
530
D | S
TCTC
A
SCAC
S
TCCT
A
SCTC
K
A
S
C
TCCT
S | E
SAGO
STGO
STGO
STGO
K
CCTO
K
K
AGO
S | D
SATI
SCA
E
SAGI
R
CGGI
A
CAI
E
 | E
GAAG
T
ACTO
E
GAAG
A
GCTO
A
GCTO
A
C
C
A | A
SCT
P
CCT.
E
SAG
G
SGT
G
G
G
G
G
G
G
G
G
C
C
T
C
C
C
T
C
C
C
T | R
K
AAG
GCC
G
GGA
SCA
K | A
GCC.
CAGI
T
ACCI
D
GACI
A
SCCC
 | K
AAGO
Y
AAGO
S
S
S
S
T
C
K | P K
CAGTC
420
A G
CCGGGC
450
K S
AAAAGT
480
S S
MGCTCC
510
P K
CCCAAA
F K | 1222
1312
1402
1492
1582 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAC
P T P
CCCACCCC | R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R | V
STG.
S
NGTI
P
CCTI
K
NAG
S
TCCI | K
NAGA
P
CCCT
Q
CAGA
A
S
CAGA
S
S
CAGA
K
K
AAAG | K
L
TAN
S
IGCJ
R
IGCG
S
IGTC
A
 | K
AAGCJ
R
AGCJ
V
TTGJ
E
SAGC
A
CAC | A
SCC
K
AAG
K
AAG
7
ACC
E
SAG
A | E
CGAG
400
P
CCCA
430
A
CGCC
490
E
CGAG
520
E
CGAG
 | S
AGCT
D
GAC
K
AAAA
K
AAAG
S | S
TCT
V
GTC
S
AGC
K
AAG
S | S
TCA
T
ACT
S
AGC
AGCA
T
ACG | P
CCTI
S
TCCI
P
CCCC
S | S
AGC
K
AAGC
S
AGCC
S
TCTO
P
CCT/
S
 | S
TCAC
P
CCCGC
E
SAGC
L
CTAC
K
K
NAAC
S | D
DATT
P
CCTG
E
DAGG
P
CCTG
P
CCCC
S | s
icto
E
icto
E
icto
S
icto
S
icto
S | D
SACI
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
 | 380
S
AGT:
410
K
AAGC
440
S
AGC
470
Q
CAGC
500
K
AAGJ
530
D
547
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
50 | S
SCAC
S
SCAC
S
CCC
K
K
S
S
CCC
K
S
S
S
S
S
S
S
S
S | E
SAGO
STGC
S
S
GTC
R
K
K
AGO
S | D
SATI
SCA
E
SAGI
R
CGG
A
CAI
E
 | E
T
ACTO
E
SAAC
A
SCTO
A
SCAC | A
SCT
P
CCT.
E
SAG
G
SGT
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | R
K
AAG
GCCi
GGA
GGA
X
X
X | A
GCC.
Q
CAGI
T
ACCI
GACI
A
CCCC
A
CCCC
A
CCCC
A
CCCC
C
A
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | R
AAGO
P
CCTC
X
AAGO
S
S
AGC/
V
STTC
R
R | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S
NGCTCC
510
P K
CCAAA
240
S S | 1222
1312
1402
1492
1582 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAC
P T P
CCCACCCC | R
CAAGI
SAAAG
SAAAG
SAAAG
SAGT
SAGT | V
STG.
S
AGTI
P
CCTI
K
AAGI
S
FCCI
K | K
AAGA
P
CCCT
Q
CAGA
A
S
CAGA
S
S
AGCA
K
AAAG | K
L
TAJ
S
S
GCJ
R
S
GCGC
S
S
GCTC
A
 | K
AAGCJ
R
AGCJ
R
TTGJ
E
SAGC
A
SCAC | A
SCC
K
AAG
K
AAG
7
CC
E
SAG
A
SCO | E
CGAG
400
P
CCCA
430
A
GCCC
430
A
CGCC
490
E
CGAG
520
E
CGAG
 | S
A
GCT
D
GAC
K
AAA
K
AAG
S | S
TCT
V
GTC
S
AGC
K
AAGC
S
AGC | S
TCA
T
ACT
S
ACC
S
ACC
S
ACC
S
ACC | P
CCTI
S
TCCI
P
CCCC
S
AGCI | S
AGC ⁴
K
AAGC
S
TCT(
P
CCT)
S
AGCI
 | S
TCAC
P
CCCGC
E
CAGCO
L
CTAC
K
ANAAC
S
S | P
CCTG
P
CCTG
P
CCTG
P
CCTG
S
CCTG | s
icto
E
icto
E
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
icto
S
i
i
i
i
i
i
i
i
i
i
i
i
i
i
i
i
i
i | A
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCA
 | 380
S
AGT:
410
K
AAGC
440
S
AGC
500
K
AGC
500
K
AGC
500
K
AGG
500
K
AGG
500
K
AGG
500
K
AGC
500
K
AGC
500
K | S
SCAC
S
SCAC
S
SCCC
A
S
S
C
C
C
S
S
C
C
C | E
SAGO
STGC
S
GTC
K
K
AGO
S
GTC | D
SATI
A
SCAI
E
SAGI
A
SCAI
E
SAAI
 | E
T
ACTO
E
SAAO
A
SCAO
A
SCAO
E
DAAO | A
SCT
P
CCT.
E
SAGI
G
SGT
C
C
SGT
C
C
SAGI | R
K
AAG
GCCi
GGA
GGA
X
X
K |
 | R
AAGO
P
CCTO
X
AAGO
S
S
S
S
S
S
T
T
C
R
C
T
T
C
T
T
C
T
T
C
T
T
C
T
T
C
T
T
C
T
T
C
T
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
T
C
C
T
C
C
T
C
C
T
C
C
T
C
C
T
C
C
C
T
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | P V
CCAGTC
420
A G
CCGGGC
450
K S
KAAAGT
480
S S
S
GCCCA
P K
CCCAAG
540
P K
CCCAAG | 1222
1312
1402
1492
1582
1672 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAACC
D S E
GACTCAGAM
P T P
CCCCACCCC | P
TCCAI
K
CAAGI
K
SAAA
P
TCCTI
SAGT?
V
TGTG/ | V
STG.
S
AGTI
P
CCTI
K
AAGI
S
FCCI
K | K
AAGA
P
CCCT
CAGA
A
CCAGA
A
S
CAGCA
K
K
AAAAG | K
L
TAJ
S
MGCJ
R
MGGC
S
MGTC
S
MGTC
A
 | K
AAGCJ
R
AGCJ
V
TTGJ
E
SAGC
A
SCAC | A
SCC
K
AAG
K
AAG
T
C
E
SAG
A
SCO | E
CGAG
400
P
F
CCCA
430
A
460
A
CGCC
490
E
CGAG
550
E
CGAG
550
 | S
A
GCT
D
GAC
K
AAA
K
AAG
S
AGC | S
TCT
V
GTC
S
AGC
K
AAGC
S
AGC | S
TCA
T
ACT
S
ACC
S
CA
T
CCA
T
ACC
S
ACC | P
GACI
P
CCTI
S
TCCI
P
CCCC
S
S
AGCI | S
AGC"
K
AAGG
S
TCTO
P
CCTJ
S
AGCJ
 | S
TCAC
P
CCGC
E
CAGC
L
CTAC
K
AAAAC
S
S
MGCT | P
CCTG
P
CCTG
P
CCTG
P
CCTG
S
CCTG | s
icro
E
BAGJ
A
CCC
S
CCC | A
SCAL
SCAL
SCAL
SCAL
K
MACCI
SCAL
E
STAL
 | 380
SAGT:
410
K
AAGC:
410
S
AGC:
410
S
AGC:
500
K
AGC:
530
C
AGC:
530
D
SATI
550 | S
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC | E
SAGO
STGO
STGO
R
CCTO
K
NAGO
S
GTO | D
SATI
A
SCAI
E
SAGI
A
SCAI
E
SAAI
 | E
GAAC
T
ACTO
E
SAAC
A
SCAC
E
CAAC | A
SCT
P
CCT.
E
SAGI
G
SGT
C
G
SGT
E
SAGI | R
AAG
GCCC
GGGA
GGGA
K |
 | R
AAGO
X
AAGO
S
S
S
S
S
T
T
C
R
S
S
T
T
C
S
S
T
T
C
S
S
S
S
S
S
S
S
S
S | P V
CCAGTC
420
A G
CCGGGC
450
K S
MAAAGT
480
S S
MGCTCC
510
P K
CCCAAG
510
P K
CCCAAG
570
570 | 1222
1312
1402
1492
1582
1672 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAA
P T P
CCCACCCC
S K A | P
TCCAG
K
SAAAA
P
CCCTJ
S
SAGT
V
T
T | V
STG.
S
AGTI
P
CCTI
K
AAGO
S
CCCI
K
AAGO
P | K
AAAGA
P
CCCCT
Q
CAGA
A
A
S
CCAGA
S
S
A
AGCCA
K
K
K | K
L
TAN
S
GCJ
R
GCGC
S
GCGC
A
CAC
P
 | K
AAGC
S
AGCJ
R
AGAJ
V
TTGJ
E
SAGC
Q | A
SCC
K
AAG
T
C
SAG
A
SCC
A | E
GAG
400
P
GCCA
430
A
460
A
560
C
490
E
GAG
520
E
GAG
550
G
 | S
GCT
D
GAC
K
AAA
K
AAG
K | S
TCT
V
GTC
S
AGC
A
AGC
A
AGC
A | S
TCA
T
ACT
S
ACC
ACC
ACC
ACC
N | P
GACI
P
CCTI
S
TCCI
P
CCCC
S
CCCC
S
AGCI
G | S
AGC"
K
AAGC
S
TCTO
P
CCTJ
S
AGCJ
V
 | S
FCAC
P
CCGC
E
E
GAGC
L
CTAC
K
K
AAAAC
S
S
AGCT
P | P
P
E
CCTG
P
CCTG
P
CCTG
S
CCTG
S
CCTG | s
icro
E
BAGJ
A
CCC
S
CCC
S | D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
 | 380
S
AGT:
410
K
AAGC
440
S
AGC
470
Q
CAGC
500
K
MAGJ
530
D
SATT
560
N | S
SCAC
S
SCAC
S
SCAC
S
S
S
C
C
C
S
S
C
C
C
S
S
C
C
C
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | E
SAGO
STGC
S
GTGC
K
K
AGO
S
K
GTG
K
L | D
SATI
BCA
E
SAGI
R
CGOI
A
CAI
E
SAAI
A
 | E
GAAC
T
ACTO
E
SAAC
A
SCAC
E
CAAC
G | A
SCT
P
CCT.
E
SAG
SGT
C
SAG
SAG | R
AAG
GGCC
GGA
SCAC
K
AAAA |
 | R
AAGO
X
AAGO
S
S
S
S
S
T
T
C
S
T
T
C
S
S
T
T
C
S
S
S
S | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S
NGCTCC
510
P K
CCCAAA
540
P K
CCCAAA
540
P K
S70
S70
E E | 1222
1312
1402
1492
1582
1672 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAA
P T P
CCCACCCC
S K A
AGCAAAGC | p
TCCAN
K
CCAAGU
K
SGAAAA
P
CCCCTI
S
SGAGT'
V
TGTGJ
T
TACTO | V
STG.
S
AGTI
P
CCTI
K
AAGG
FCCJ
K
AAGG | K
AAAGA
P
CCCCT
Q
CAGA
A
A
AGCCA
K
K
AAAAG
K
K
AAAAC | K
L
TAN
S
GCJ
R
GCGC
S
GCGC
A
CAC
 | K
AAGC
S
AGCJ
R
AGAJ
V
TTGJ
E
SAGC
A
CAGO | A
SCC
K
AAG
K
AG
SAG
ACC
A
SCC
A
SCC | E
CGAG
400
P
5CCA
400
A
6GCT
460
A
5GCC
490
E
520
E
550
G
GAG
550
G
GGAG
 | S
A
GCT
D
GAC
K
AAA
S
AAG
K
AAG | S
TCT
V
GTC
S
AGCA
K
AAGCA
K
AAGCA
AGCCA | S
TCA
T
T
ACT
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
S
ACC
ACC | P
GACI
P
CCTI
S
TCCI
P
CCCC
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
S
CCCC | S
AGC
K
AAGC
S
AGCC
S
TCTC
P
CCT/
S
CCT/
S
TCTC
 | S
FCAG
P
ECCGG
EAGC
L
CTAG
S
AGCT
P
CCAG | P
CCTG
E
SAGG
P
CCTG
P
CCCTG
S
CCTG | s
icto
E
iAGJ
A
icco
S
icco
S
icco
S | D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
 | 380
S
AGT:
410
K
AAGC
410
K
AAGC
410
S
AGC
470
Q
CAGC
500
K
AAGC
530
D
SATT
560
N
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAG | S
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SC | E
SAGO
STGC
S
CTC
K
AGO
S
S
GTC
K
LAGO | D
SATI
A
SCAL
E
SAGI
R
SCAI
E
SAAI
A
SCAI
 | E
GAAG
T
ACTO
E
GAAG
A
GCTO
A
GCAG
G
GCO | A
SCT
P
CCT.
E
SAG
SGT
SAG
SAG | R
AAG
GCC
G
GCA
K
CCA
K |
 | R
AAGO
P
CCTC
X
AAGO
S
S
S
S
S
S
S
S
S
S
AGC | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S
NGCTCC
510
P K
CCCAAA
510
P K
CCCAAA
510
CCCAAG
S S
CCCAAAG
S S
S
CCCAAG
S S
CCCAGC
S S
S
S
S
CCCGGC
S S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | 1222
1312
1402
1492
1592
1672 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAN
P T P
CCCACCCC
S K A
AGCAAAGC | P
TCCAN
K
CAAGJ
K
SAAAA
P
CCCTJ
S
SAGT
V
T
T
TACTO | S
S
AGTI
P
CCTI
K
AAGG
S
FCCJ
X
AGG | K
AAAGA
P
CCCCT
Q
CAGA
A
SCCA
A
SCCA
S
AGCA
K
AAAAG
K
AAAAC | K
L
MAAJ
S
S
GCJ
R
GCGC
S
GCGC
A
S
GCAC
 | K
AAGCI
R
R
AGAJ
B
B
G
AGCA
C
AGC | A
SCC
K
MAG
K
AAG
T
C
E
SAG
A
SCO
A
SCA | E
GGAG
400
P
GCCA
430
A
GGCT
460
A
CGCC
490
E
GGAG
550
G
GGA
550
G
GGA
550
G
 | S
A
GCT
D
GAC
K
AAAG
S
AAGC
K
AAG | S
TCT
V
GTC
S
AGC
A
AGC
A
AGC
A
AGC | S
TCA
T
ACT
S
ACC
ACC
ACC
ACC
ACC
ACC
ACC
ACC
ACC
A | P
GACI
P
CCTI
S
TCCI
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
S
CCCC
S
S
S
S
CCCC
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S
AGC
K
AAGC
S
AGCC
S
TCTC
P
CCT/
S
CCT/
S
CCT/
S
TCTC
 | S
FCAG
P
E
EAGGO
L
L
CTAG
K
AAAAC
S
AGCT
P
CCAG | P
P
E
CCTG
BAGG
P
CCTG
P
CCCTG
S
CCCT
A | s
icto
E
iAGJ
A
icco
S
icco
S
icco
S | A
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCA
 | 380
S
AGT
410
K
AAGC
410
S
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
K
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
500
AGC
AGC
500
AGC
AGC
500
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC | S
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SC | E
SAGC
STGC
S
S
CCTC
K
K
AGC
S
GTC
K
LAAC | D
SATI
SCAL
E
SAGI
R
CGGI
A
SCAI
E
SAAI
A
CAI
 | E
GAAG
T
ACTO
E
GAAG
A
GCTO
A
GCAG
GC/ | A
SCT
P
CCT.
E
SAGI
G
SGT
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | R
AAG
GCC:
G
GCA
K
CAAA |
 | R
AAGO
P
CCTC
X
AAGJ
S
S
S
S
T
C
R
C
S
AAGC | E V
CCAGTC
420
A G
CCGGCC
450
K S
AAAAGT
480
S S
MGCTCC
510
P K
CCCAAG
510
P K
CCCAAG
570
E E
SAGGAA
600 | 1222
1312
1402
1492
1592
1672
1762 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCAC
V T T
GTGACAAC
D S E
GACTCAGAN
P P P
P T P
CCCACCCC
C
CCCACCCC
S K A
AGCCAAGC
C | P
TCCAN
K
CAAGI
K
GAAAM
P
CCCTI
S
SAGT
V
TGTGJ
T
TACTO | V
STG.
S
AGTI
P
CCTI
K
S
CCCI
K
AAGI
P
CCCI | K
AAAGA
P
CCCCT
Q
CAGA
A
AGCCA
K
AAAAG
X
AAAAC | K
L
TAJ
S
GCJ
R
GCGC
S
GCAC
P
CAC
 | K
AAGC
S
AGCJ
R
R
AGAJ
E
SAGC
A
SCAGC
K | A
SCC
K
MAG
K
AGC
E
SAG
A
SCO
A
SCA | E
GAG
400
P
GCCA
430
A
GGCT
430
A
GGCT
430
A
GGCT
430
C
A
GGCT
520
E
GAG
6
GGA
6
GGA
6
GGA
6
GGA
6
C
A
A
430
A
430
A
430
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
A
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
520
C
50
C
5
 | S
AGCT
D
GAC
K
AAAA
K
AAGC
K
AAGC | S
TCT
V
GTC
S
AGC
A
AGC
A
AGC
A
AGC
C | S
TCA
T
ACT
S
AGC
A
GCA
T
ACG
S
AGC
N
AATI | P
GACI
P
CCTI
S
TCCI
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCCC | S
AGC
S
AGC
S
TCTO
P
CCT/
S
AGC/
V
STTO
 | S
TCAC
P
CCGC
E
SAGC
K
K
K
K
K
K
K
K
K
K
K
K
K
K
K
CAGC
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | P
ECTG
E
EGAGG
P
CCTG
P
CCCTG
S
CCCT
A | A
A
A
A
A
A
A
A
A
A
A
A
A
A | D
GAC
S
GCA
S
GCA
S
A
A
C
A
GCA
S
A
A
C
A
GCA
S
A
A
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
S
C
 | 380
S
AGT:
410
K
AAGC
410
S
AGC:
410
S
AGC:
500
S
AGC:
500
S
AGC:
500
NAGC
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
S
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC:
500
AGC: | S
SCAC
S
SCAC
S
SCAC
S
S
S
S
S
S
S
S
S
S | E
SAGC
STGC
S
S
GTC
K
S
GTC
K
S
GTC
K
LAAC | D
SATI
SCA
E
SCA
R
CGGI
A
SCA
CA
CA
CA
CA
CA
CA
CA
CA
CA
 | E
GAAG
T
ACTO
E
GGCTO
A
GCAG
G
GGC/ | A
SCT
P
CCT.
E
SAGI
G
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
S
SGT
C
SGT
S
SGT
S
SGT
C
SGT
S
SGT
S
SGT
C
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
SGT
S
S
SGT
S
S
S
S | R
K
AAG
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G |
 | R
AAGO
P
CCTC
X
AAGO
STTC
R
STTC
R
STTC
R
STTC
R
S
SAGC | P V
CCAGTC
420
A G
CCGGCC
450
K S
KAAAGT
480
S S
S
CCCAAG
510
P K
CCCAAG
510
P K
CCCAAG
510
C S
70
E E
S
AGGAA
60 | 1222
1312
1402
1492
1582
1672
1762 |
| K P A
AAACACACC
S A T
AGTGCCAC
AGTGGCCAC
S G Q
AGTGGCCAACC
D S E
GACTCAGAA
P T P
CCCCACCCC
S K A
AGCAAACC
B E D | P
TCCAN
K
CAAGJ
K
SGAAA
P
CCCTJ
S
SGAGT
V
TGTGJ
T
TACTO
T | P
CCTI
K
AGTI
P
CCTI
K
AAGI
P
CCCI
E | K
AAGA
P
CCCT
Q
CAGA
A
S
CCAGA
S
S
S
CCA
S
K
AAAAG
K
K
AAAAG | K
L
TAJ
S
GGCJ
R
GGCG
S
GGTC
A
CAC
N
 | K
AGCJ
R
R
AGAJ
V
TTGJ
E
SAGC
Q
C
AGC
K | A
SCC
K
AG
K
AG
T
C
SAG
A
SCA
K
K
K
K
K
K
K
K
K
K
K
K
K | E
GAG
400
P
GCCA
430
A
GCCC
490
E
GAG
520
E
GAG
6
550
G
GAG
6
550
G
GAG
6
550
G
A
GAG
490
A
E
GCC
490
A
A
A
GCCA
400
A
A
A
GCCA
400
A
A
A
A
GCCA
400
A
A
A
GCCA
400
A
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
A
A
GCCA
A
A
GCCA
A
A
GCCA
A
A
A
 | S
AGC
A
GCT
D
GAC
K
AAA
S
AAG
K
AAG
K
AAG | S
TCT
V
GTC
S
AGC
A
AGC
A
AGC
A
AGC
A
A
GCC
C | S
TCA
T
ACT
S
ACC
A
CCA
T
A
CCA
S
A
A
CCA
T
A
A
CCA
T
T
T | GACI
P
CCTI
S
CCCT
P
CCCC
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCTI | S
AGC
K
AAGC
S
TCTC
P
CCT/
S
CCT/
S
TCTC
P
P
CCT/
P
CCT/
P
CCT/
P
 | S
TCAC
P
CCGC
E
CTAC
K
AAAAC
S
S
S
CAGCT
P
CCAG | P
E
E
BAGG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S | S
CTC
A
CC
A
CC
A
CC
C
C
C
C
C
C
C
C
C
C | D
SACA
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
 | 380
S
AGT:
410
K
AAGC
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
5
K
5
K
5
K
5
K
5
K
5
K
5
K
5
K
5
K
5
K
5
K
5 | S
SCAC
S
SCAC
S
SCAC
S
SCAC
S
SCAC
S
S
C
C
C
S
S
C
C
C
S
S
S
C
C
C
S
S
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
C
S
S
C
S
C
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | E
SAGO
STGC
S
S
CCTC
K
AAGO
S
GTC
K
K | A
SCA
E
SAG
R
CGG
A
CA
CA
CA
CA
CA
CA
CA
CA
 | E
ACTO
E
SAAC
SCTO
A
SCTO
A
SCAO
E
SAAC
C
C
SCAO
C
N | A
SCT
P
CCT.
E
SAG
SGT
G
SGT
G
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
S
S
SGT
S
S
S
S
S
S
S
S
S
S
S
S
S
S | K
AAG
GCCC
GGA
CGGA
CGGA
CGGA
CGGA
CGGA |
 | R
AAGO
P
CCTO
X
AAGO
STTO
K
STTO
E
CAAGO
D | P K
CCAGTC
420
A G
CCGGCC
450
K S
AAAAGT
480
S S
S
AGCTCC
510
P K
CCCAAG
510
P K
CCCAAG
570
E E
SAGGAA
600
E A | 1222
1312
1402
1492
1582
1672
1762 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAN
P T P
CCCACCCC
S K A
AGCAAAGC
<u>E E D</u>
GAGCAACM | P
TCCAN
K
CAAGJ
K
SAAAA
P
CCCTJ
S
SAGT
S
V
TGTGJ
T
TACTO
T
CACAO | STG.
SAGTI
PCCTI
KAAGO
SCC/
KAAGO
PCCC/
ECC/ | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | K
L
TAJ
S
GCJ
R
GGC
S
GGC
S
GGC
S
GGC
S
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
 | K
AAGO
S
AGCJ
R
R
AGAJ
V
V
TTGJ
E
SAGO
A
AGAJ
V
V
S
CAGO
C
AGO
K
AAGAJ | A GCC K AAG T ACC E GAG A GCC | E
CGAG
400
P
CCA
430
A
430
A
500
E
CGAG
520
E
CGAG
550
G
G
GAG
550
G
A
580
A
580
A | S
AGCT
D
GAC
K
AAA
S
AAGC
K
AAG
AGCC | S
TCT
V
GTC
S
AGC
A
AGC
A
AGC
A
AGC
A
AGC
A
AGC
A
A
GCC
A
A
GCC
A | S
TCA
T
ACT
S
ACC
A
ACC
A
ACC
A
A
A
A
A
A
A
A
A
A
A
 | D
GACI
P
CCCTI
S
TCCI
P
CCCCI
P
CCCCI
S
CCCCI
S
CCCCI
C
S
S
S
S
S
S
S
S | S
AGC
S
AGC
S
TCTO
P
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
CCT/
S
S
S
CCT/
S
S
S
CCT/
S
S
CCT/
S
S
S
C
CCT/
S
S
S
S
C
C
S
S
S
S
C
S
S
S
S
S
S
S
S | S
TCAC
P
CCCCC
E
E
SAGC
C
CCAC
S
SAGCT
P
CCAC
G
G
G
G
G
G
G
G
 | D
DATT
P
CCTG
E
DAGG
P
CCTG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG | S
CTC
E
BAGJ
A
CCC
S
CCC
S
CCC
G | A SCAL
 | 380
S
AGT:
410
K
AAGC
410
K
AAGC
500
K
AGC
500
K
AGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
K
AAGC
500
AAGC
500
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC
AAGC | S
A
S
C
C
C
C
C
C
C
C
C
C
C
C
C | E
SAGO
STGC
S
S
CCTC
K
AGCO
S
S
GTC
K
K
AGCO | D
SATI
SCA
E
SAGI
R
CGGI
A
CAGI
A
CAGI
H
CAGI | E
GAAG
T
SACTO
E
SAAG
A
SCAG
C
SAAG
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
 | A
SCT
P
CCT.
E
SAGI
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | R
AAG
GCCC
GGA
GGA
SCA
K
E
GAAJ
T
ACAC |
 | R
AAGO
P
CCTC
X
AAGO
STTC
STTC
E
DAGO
D | P V
CCAGTC
420
A G
CCGGGC
450
K S
NAAAGT
480
S S
NGCTCC
510
P K
CCCAAA
540
P K
CCCAAA
540
P K
CCCAAA
540
P K
S
CCCAAAG
540
P K
S
CCCAAG
540
P K
S
CCCAGTC
450
S S
S
CCGGGC
510
P K
S
CCCAGTC
510
P K
S
CCGGGC
510
P K
S
CCCAAGTC
510
P K
S
CCCAGTC
510
P K
S
CCCAGTC
S
S
S
CCGGGC
S
S
S
S
CCGGCC
S
S
S
S
CCCGGC
S
S
S
S | 1222
1312
1402
1492
1582
1672
1762 |
| K P A
AAACCAGC
S A T
AGTGCCAC
S G Q
AGTGGCCAC
D S E
GACTCAGAN
P T P
CCCACCCC
S K A
AGCAAAGC
<u>E E D</u>
GAGCAACAM | P
TCCAI
K
CAAGJ
K
SAAAA
P
CCCTI
S
SAGT'
V
TGTG/
T
TACTO
T
CACAO | V
STG.
S
AGTI
P
CCTI
K
AAGU
S
CCCI
K
AAGU
P
CCCI
E
CCCI | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
S
SCCA
S
S
SCCA
K
AAAAG
X
X
AAAAC
Q
CAGA
 | K
L
TAJ
S
GCJ
R
GGC
S
GGC
S
GGC
S
GGC
S
C
ACAC
N
-
 | K
AAGCJ
R
AGCJ
R
AGAJ
V
TTGJ
E
SAGC
Q
CAGC
X
AGC | A GCC K A A G K A A G C C A A G C C A A G C C A A G C C A A G C A A A G C A A A G C A A A G C A A A G C A A A A | E
GAG
400
P
GCCA
430
A
430
A
GCC
490
E
GAG
520
C
GAG
GAG
GAG
GAG
GAG
GAG
A
GAG
A
100
P
P
400
P
CCA
430
A
430
A
430
A
430
A
430
A
430
A
430
A
430
A
430
A
430
A
430
A
430
A
430
A
400
A
430
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
400
A
40
A
400
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
40
A
4
A
4
 | S
AGCT
D
GAC
K
AAAG
S
AAGC
K
AAAG
K
AAGC | S
TCT
GTC
S
AGC
A
GCA
K
AAG
S
AGC
A
AGC
A
AGC
A
GCC
C
GGC | S
TCA
T
ACT
S
ACT
S
CCA
T
ACC
S
ACC
N
T
ACC | P
CCCTI
S
TCCI
P
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCTI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCCI
S
TCCC | S
AGC
K
AAGC
S
AGC
S
TCTC
P
CCT/
S
AGC/
V
STTC
P
CCAC
 | S
TCAC
P
CCGC
E
DAGC
L
CTAC
S
MGCT
P
CCAG
G
G
OTT | P
CCTG
E
BAGG
P
CCTG
P
CCCC
S
CCCTG
S
CCCTG
S
CCCC
S | A
CTC
A
CTC
C
C
C
C
C
C
C
C
C
C
C
C
C | D
SACA
SCAL
SCAL
SCAL
E
SAAC
SCAL
E
SAAC
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
 | 380
S AGT:
410
K
AAGC
540
S AGC
500
K
AGC
530
S ATT
560
N
AGC
530
S ATT
560
N
AAGC
530
S ATT
560
N
AAGC
500
S K
AAGC
500
S K
AGC
500
S K
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC | S
A
S
S
C
C
C
C
C
C
C
C
C
C
C
C
C | E
SAGO
STGC
S
GTC
K
AGO
S
GTC
K
AGO
K
L
AGO | D
SATI
SCA
E
SAGI
R
CGOI
A
CCAI
E
SAAI
A
CCAI
H
CCAI | E
ACTO
E
SAAC
A
SCTO
A
SCAC
B
SCAC
C
C
SCAC
C
N
N
 | A
SCT
P
CCT.
E
SAGI
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | R
AAG
GCCC
GGGA
SCAC
K
CAAJ
T
ACAC |
 | R
AAGO
X
AAGO
STTC
X
AGC/
V
STTC
R
STTC
R
STTC
D
SAGC | P V
CCAGTC
420
A G
CCGGCC
450
K S
480
S S
GCTCC
510
P K
CCAAA
S
510
P K
CCAAA
S
510
P K
CCAAA
600
E A
630
C
CAGTC
420
420
420
450
K S
510
510
510
510
600
E A
600
E A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
600
A
800
A
800
A
800
A
800
A
800
A
800
A
800
A
800
A
800
A
800
A
800
A
800
A
800
A
800 | 1222
1312
1402
1492
1592
1672
1762 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAN
P T P
CCCACCCC
S K A
AGCAAAGC
E E D
GAGCAACAC | p
rccai
k
caagi
k
gaaaa
p
cccti
s
sagt:
v
rotgi
t
racto
<u>t</u>
cacao | STG.
STG.
P
CCTI
K
AAGU
S
CCCI
E
CCCI
S
S | к
мада
р
р
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
а
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост | K
L
TAJ
S
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
NGCJ
R
N
NGCJ
R
N
NGCJ
R
N
N
N
N
N
N
N
N
N
N
N
N
N
N
N
N
N
N
 | K
AGCJ
R
AGCJ
R
AGAJ
V
TTGJ
E
SAGC
Q
CAGC
Q
CAGC
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X | A GCC KAAG KAAG TACC E SAG A GCC A G | E
GAG
400
P
GCCA
430
A
GCCT
400
E
GCCC
490
E
GAG
520
E
GAG
550
GAG
550
GAG
610
L | S
AGCT
D
GAC
K
AAAA
AAG
K
AAG
K
AAG
C
C | S
TCT
V
GTC
S
AGC
K
AGC
K
AAGC
A
AGCC
GGGJ
T | S
TCA
T
ACT
S
ACC
S
ACC
S
ACC
N
AATI
T
ACC
P
 | D
GACI
P
CCTI
S
TCCI
P
CCCC
P
CCCC
S
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | S
AGC
K
AAGC
S
TCTC
P
CCT/
S
TCTC
P
CCT/
S
TCTC
T
C
CCAC | S
TCAC
P
CCGC
E
E
CTAC
K
K
S
S
CCAC
G
G
G
TT
F
 | D
DATT
P
CCTG
E
SAGG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCAG
P | A
CTC
A
CTC
E
A
CTC
S
CTC
S
CTC
C
S
CTC
C
S
CTC
C
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTC
S
CTCC
S
CTCC
S
CTC
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
S
CTCC
CTCC
S
CTCC
S
CTCC
C
CTCC
C
CTCC
CTCC
C
CTCC
C
C
C
C
C
C
C
C
C
C
C
C | D DACI
D DACI
SCAL
SCAL
SCAL
E DACI
E | 380
8
AGT:
410
K
AAGK
440
S
AGC
440
S
AGC
500
K
AGC
500
S
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
500
N
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
 | S
A
S
S
C
C
C
C
C
C
C
C
C
C
C
C
C | E
SAGO
STGO
STGO
SCTO
KAGO
K_GTO
K_GTO
K_GTO
G | D
SATI
A
SCAI
E
SAAI
E
SAAI
E
CAI
E | E
ACTO
E
SAAC
A
SCTO
A
SCAC
B
SAAC
B
SAAC
B
SAAC
K
 | A
SCT
P
SCT.
E
SAGI
G
SGT
G
SGT
C
E
SAGI
R
A
A
G
SGT
R | R
AAGI
A
GCCCi
GGGAI
A
GGGAI
A
CACA
A | | R
AAGO
X
AAGO
STTC
R
AAGO
STTC
BAGO
D
JATO
 | P V
CCAGTC
420
A G
CCGGGC
450
K S
S
NAAAGT
510
P K
CCCAAAG
510
P K
CCCAAAG
570
E E
S70
E A
GOAG
600
E A
GOA
630
P F | 1222
1312
1402
1492
1582
1672
1762
1852 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGGCCAC
V T T
GTGACAAC
D S E
GACTGAGA
P T P
CCCACCCC
S K A
AGCAAAGC
<u>E E D</u>
GACCAAAGC | P
TCCAI
K
CAAGI
K
SAAA
P
CCCTI
S
SAGT
V
TCTGJ
T
TCACTO
T
CAAPA | V
STG.
STG.
P
CCTI
K
S
CCCI
E
S
I
CCTI
S
I
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCTI
CCT | к
мада
р
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
о
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост
сост | K
L
S
MGCJ
R
MGCJ
R
MGCGC
S
MGTC
A
C
A
C
A
C
A
 | K
AAGC
S
AGCJ
R
AGCJ
E
SAGC
S
AGC
K
AGC
K
AGC
C
AGC
K
AGC
C
AGC
C
AGC
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
AG | A GCC K AAG K AAG T CC E GAG A GC A GC A GC A GC A GC A GC | E
GAG
400
P
GCCA
430
A
GCCT
400
A
GCCT
400
A
GCC
520
E
GAG
550
G
S50
G
S50
G
A
M
610
L
 | S
AGC
A
GCT
D
GAC
K
AAAA
S
AGC
K
AAG
CAC | S
TCT
V
GTC
S
AGC
K
AAG
S
AGCA
K
AAG
C
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | S
TCA
T
ACT
S
ACC
S
ACC
S
ACC
N
ACC
S
ACC
N
ACC
N
ACC
N
ACC
S
ACC
S
ACC
S
ACC
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
S
S
C
ACC
T
S
ACC
S
S
C
ACC
S
S
C
ACC
S
S
C
ACC
S
S
C
ACC
S
S
C
ACC
S
S
C
ACC
S
S
C
ACC
S
S
C
ACC
S
S
S
C
ACC
S
S
S
C
ACC
S
S
S
C
ACC
S
S
S
S | D
GACI
P
CCTI
S
TCCI
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCTI
S
TCCI
S
CCCTI
S
TCCI
S
TCCI
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
TCCI
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S
AGC
K
AAGC
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
TCTO
 | S
TCAC
P
CCGC
E
SAGC
L
CTAC
S
S
S
S
CTAC
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | D
DATT
P
CCTG
E
DAGG
P
CCTG
P
CCCC
S
CCTG
S
CCTG
S
CCTG
P
CCTG
P
CCTG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCG
P
CCCCCG
P
CCCCG
P
CCCCG
P
CCCCCCG
P
CCCCG
P
CCCCCCCC | | D D D D D D D D D D D D D D D D D D D
 | 3300 s
AGT:
410 k
AAGC
410 k
AAGC
500 k
AGC
500 k
AAGC
500 k
AAGC
500 k
AAGC
500 k
AAGC
500 k
AAGC
500 k
AAGC | SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC | E
SAGO
STGO
S
CTGO
K
S
GTGO
K
LAGO
G
G
G
G
G
G
G
G
G | A
SCA
E
SAGI
A
SCA
E
SAGI
E
SAGI
E
SAGI
E
 | E
GAAG
T
T
GAAG
E
G
G
G
C
C
C
C
C
C
C
C
C
C
C
C
C | A
SCT
P
SCT.
E
SAG
SGT
G
SGT
E
SAG
SGT
E
SAG
SGT
R
R | R
K
AAG
GCC:
GCGA
GCGA
K
K
CAAA
CAAA
K
CAAA
T
T
CAAA |
 | R
AAGO
X
AAGJ
S
AGC/
V
STTC
R
S
AAGO
S
S
AGC/
S
S
AGC/
S
S
AGC/
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | 530
4200
4200
450
450
450
450
450
510
P K
510
P K
510
F F F K
510
F F F K
510
F F F F K
510
F F F F F F F F F F F F F F F F F F F | 1222
1312
1402
1492
1582
1672
1762
1852 |
| K P A
AAACCAGC
S A T
AGTGGCCAG
S G Q
AGTGGCCAG
V T T
GTGACAACC
D S E
GACTCAGAM
P T P
CCCACCCC
S K A
AGCAAAGC
E E D
GAGGAACAG
A
T P
GGAACTCC | P
R
R
CAAGI
K
SAAAA
P
COCTI
SAGT
V
IGTGJ
T
TACTO
T
CAAAI | STG.
STG.
STG.
P
CCTI
K
S
CCCI
E
CCCI
S
S
I
CCTI | R
AAAGA
P
CCCT
CAGA
A
SCCA
S
SCCA
R
S
SCCA
R
S
SCCA
R
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | K
L
L
L
L
L
L
L
L
L
L
L
L
L
 | K
AAGC
S
AGCJ
R
AGCA
V
JTGJ
E
E
SAGC
A
C
AGC
X
AGC
X
TTA | A SCC KAG KAG TCC ESAG ACC KAG | E
GAG
400
P
GCCA
430
A
GGCT
400
E
GAG
520
E
GAG
550
C
GAG
550
C
GAG
610
L
C
CCA
 | S
A
GCT
D
GAC
K
AAA
S
AGC
K
AAG
K
AAG
CAG | S
TCT
V
GTC
S
AGC
K
AAGC
K
AAGC
S
GGC
G
GGG
T
T
C | S
TCA
T
ACT
S
AGC
A
GCA
T
A
CCA
S
AGC
N
T
A
CCA
T
T
A
CCA | D
GACI
P
CCTI
S
TCCI
P
CCCI
P
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
N
A
A
A | S
AGC
²
K
AAGC
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
T | S
TCAC
P
CCGC
E
EAGC
L
CTAC
K
K
S
S
CTAC
S
CAG
G
G
TTC | P
CCTG
E
BAGG
P
CCTG
P
CCCC
S
CCCC
S
CCCC
A
CCCC
CCCC
CCCC
S
CCCC
CCCC
S
CCCC
CCCC
S
CCCCC
S
CCCCC
S
CCCCCC
 | S
CTC
A
CTC
E
A
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCCC
S
CCCCCC
S
CCCCCCCC | A SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL | 3380
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
8
410
8
8
410
8
8
8
8
8
8
8
8
8
8
8
8
8 | SCAC
 | E
SAGTO
STGC
S
CTCC
K
AGTO
S
G
G
G
G
G
G
G | A
SCA
B
SAGI
R
SGGI
A
SCA
B
SAAS
A
SCA
B
B
SAAS
A
SCA
C
B
SAAS | E
ACTO
E
SAAC
SCAC
E
SCAC
E
SCAC
E
SCAC
E
SCAC
C
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
C
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC | A
SCT
E
SAGI
G
SGT
G
SGT
E
SAGI
E
SAGI
R
AGG
 | R
AAG
GGCG
GGA
K
SCA
K
CAAA
T
ACAC
A
ACAC | A GCC.
Q CAGI
T ACC:
D GACI
A GCC.
A GCC. | R
AAGO
X
AGO
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | 3 50 420 A G 3 G
 | 1222
1312
1402
1492
1582
1672
1762
1952 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGGCCAC
V T T
GTGACAAC
D S E
GACTGAGAA
P T P
CCCACCCC
S K A
AGCAAAGC
E E D
CACGAACAC | P
TCCAI
K
CAAGJ
K
SAAAA
P
CCCTI
S
SAGT'
V
TGTGJ
T
TCACTO
T
CAACA | V
STG.
S
NGTI
P
CCTI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
K
NAGU
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI
TCCI | K
AAGA
P
CCCT
Q
CAGA
X
SCCA
S
SCCA
X
SCCA
X
AAAG
X
X
AAAAG
X
X
AAAAG
X
X
AAAAG | K
L
MAN
S
S
G
C
S
G
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
 | K
AAGC
S
AGCJ
R
AGCJ
E
SAGC
Q
CAGC
V
AGC
V
AGC
V
AGC
V
AGC
V
AGCJ
V
TTA | A GCC KAG KAG TCC EAG A GCC KAG KAG TCC EAG A GCC KAG | E
GAG
400
P
GCCA
400
E
GCC
400
E
GCC
400
E
GCC
520
GAG
550
GAG
550
GAG
610
L
CCCG
640
 | S
AGCT
D
GAC
K
AAA
K
AAG
S
CAG
CAG | S
TCT
V
GTC
S
AGC
A
GCA
K
AAG
S
AGC
A
AGC
GGG
T
T
ACCC | S
TCA
T
ACT
S
ACT
S
ACC
T
C
CA
S
ACC
S
ACC
S
ACC
T
ACC
T
ACC
T
C
C
A
C
C
A
C
T
C
A
C
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
S
ACC
S
S
S
S | D
GACI
P
CCTI
S
TCCI
P
CCCC
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
N
S
CCCC
N
S
S
CCCC
N
S
S
S
S | S
AGC ²
K
AAGC
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
CCT/
S
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
TCAC
P
CCGC
E
GAGC
L
CTAC
S
S
CTAC
S
S
CTAC
S
S
CTAC
S
S
CTAC
S
S
CTAC
S
S
CTAC
S
S
S
S
S
S
S
S
S
S
S
S
S | D
DATT
P
CCTG
E
SAGG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
P
CCTG
S
CCTG
S
S
CCTG
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S CTC
A CTC
E AGJ
A CCC
S CCC
S CCC
S CCC
C C C
C C C C | A AGO
 | 380
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
410
8
40
8
40
8
40
8
40
8
40
8
40
8
8
40
8
8
40
8
8
40
8
8
40
8
8
8
8
8
8
8
8
8
8
8
8
8 | S
A
S
C
C
C
C
C
C
C
C
C
C
C
C
C | E
SAGO
STGC
S
SGTC
KAGO
K_
KAGO
G
G
G
G
G
G
G
G | A
SATI
A
SCAL
E
SAGI
E
SAAI
E
SAAI
E
SAAI
 | E GAAG
T T ACTO
E SAAG
SCTO
A SCTO
A SCAG
E SAAG
E SAAG
C A SCAG
E SAAG | A
SCT
P
CCT.
E
SAGI
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | P
CCT
K
AAG
GCC
GGA
GGA
K
GCA
K
GCA
CAAAA
CAAAAAAAAAA | A GCCI
CAGI
T ACCI
D
GACI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI
ACCI | R
AAGO
X
AAGJ
S
AGC/
V
V
STTC
R
S
AGC/
V
S
S
TC
C
S
AGC/
C
S
AGC/
C
S
S
TC
C
S
AGC/
C
S
S
TC
C
S
S
S
S
S
S
S
S
S
S
S
S
S
S | J. Solution J. Solution 420 A. G. 420 A. G. 450 K. S. S. MARAGT 480 S. S. S. MGCTCC 510 P. K. -STOCKARG S70 E. P. K. A. G. G. A. | 1222
1312
1402
1492
1582
1672
1762
1852
1942 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAACC
D S E
GACTCAGAA
P T P
CCCCACCCC
S K A
AGCAAAGC
E E D
GCACGACCAC
A T P
CCCACCCCC
R R V | P
TCCAI
K
CAAGI
K
SAAAA
P
CCCTI
S
SAGT'
V
TCCAA
T
TCAAA
R | V
STG.
S
MGTI
P
CCTI
K
MAGI
S
CCCI
E
CCCI
E
CCCI
E | K
AAGA
P
CCCT
Q
CAGA
X
SCCA
X
SCCA
X
AAAG
X
AAAAG
X
AAAAG
X
AAAAG
X
AAAAG
X
AAAAG
X
X
AAAAG
X
X
X
X | K
L
MAN
S
MGCJ
R
MGCG
S
MGTC
A
CAC
N
K
AAG
E
 | K
AAGC
S
AGCJ
E
SAGC
AGC
K
AAGC
V
TTTA | A CC KAG TAC E SAG A CC KAG KIG | E
GAG
400
P
GCCA
430
A
GCC
490
E
GCC
490
E
GCA
550
G
GAG
550
G
GAG
550
G
GAG
610
L
C
C
C
GAG
400
P
C
C
A
430
A
A
GCC
400
A
A
GCC
400
A
C
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
C
GCCCA
400
A
C
GCCCA
400
A
C
GCCCA
400
A
C
GCCCA
400
A
C
GCCCA
400
A
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
AGCT
AGCT
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | S
TCT
S
GCC
S
AGC
K
AAGC
A
AGC
G
G
G
G
G
G
G
G
G
G
G
G
G
G | STCA
TCA
ACT
SAGC
SAGC
TACGI
SAGC
NATI | D
GACI
P
CCTI
S
TCCI
P
CCCI
P
CCCI
S
G
CCCI
S
G
CCCI
N
AATA
V | S
AGC
K
AAGC
S
TCTC
P
CCT/
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
TCTC
S
S
C
C
TCTC
S
S
C
C
S
S
C
C
S
S
C
C
S
S
C
C
S
S
C
C
S
S
C
S
S
C
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
 | S
P
CCGC
E
CTAC
K
K
AAAC
S
G
CTAC
S
G
CTAC
C
CTAC
S
C
C
C
C
C
C
C
C
C
C
C
C
C | P
CCTG
E
SAGG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
S
CCTG
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S CTC
A CTC
A CTC
A CTC
A CTC
A CTC
S CTC
C CTC
C CTC
A A A A
S CTC
S CTC
C CTC
A A A A
A A A A
A A A A A A A A A A | A AGCAL
SCAL
SCAL
SCAL
SCAL
SCAL
E
SAAC
Q
Q
CAGL
F
 | 330
s
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
400
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
40
K
10
K
10
K
10
K
10 | STOCA
SCAC
SCAC
SCCC
SCCC
K
SCCC
SCAC
SCCC
K
SCCC
SCAC
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCCC
A
SCCCC
A
SCCCC
A
SCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCC
A
SCCCCCC
A
SCCCCCC
A
SCCCCCC
A
SCCCCCCCC | E
SAGC
STGC
STGC
P
CCTC
K
AGC
S
SGTC
S
SGTC
S
S
G
G
G
G
G
G
G |
D
SATI
A
SCA
E
SAG
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
E
SAA
S
SAA
E
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SAA
S
SA
S
SAA
S
SAA
S
SAA
S
SA
S
SAA
S
SAA
S
SAA
S
SA
S
S
SA
S
S
SA
S
SA
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | E
ACTO
E
SAAC
SCTO
A
SCTO
A
SCAC
E
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SC | A
SCT
P
CCT.
E
SAGI
G
GGT
G
GGT
E
SAGI
E
SAGI
R
AGGC
A | R
AAGI
GGGA
SGGA
SGGA
SGGA
SGGA
SGGA
SGGA
 | | R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R
R | J SO A G 420 A G CCGGGC 450 SCGGGC 480 S S S SGCAAG 510 P K 570 SCCAAG 570 S S 560 S S 570 S S S S S 570 S S S S S S 570 S S S
 | 1222
1312
1402
1492
1582
1672
1762
1852
1942 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAM
P T P
CCCACCCCC
S K A
AGCAAGCCC
E E D
CAGCAAGAC
A T P
GCAACCCCC
R R V
CCCAACCCCCC | P
TCCAI
K
CAAGU
K
SAAA
P
CCCTI
S
SAAA
P
CCCTI
S
SAAA
V
TCTGJ
T
TACTO
Q
TCAAC
R
CAGGO | V
STG.
S
AGTI
P
CCTI
K
AGG
S
CCCI
K
AGG
S
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
CCCI
E
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
C | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
A
SCCA
K
AAAC
Q
CAGA
K
AAAC
Q
CAGA
K
CAGA
K
CAGA | K
L
MAN
S
S
S
S
S
S
S
S
S
S
S
S
S
 | K
AGGO
R
AGGO
E
SAGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC | A CC KAG KAG TACC EGAG ACC AG | E
GAG
400
P
SGCA
430
A
SGCT
460
A
SGC
490
E
SGAG
550
GA
550
GA
550
GA
550
GA
550
C
GA
610
L
C
CCC
640
V
C
GA
GA
GA
GA
GA
GA
GA
GA
GA
GA
GA
GA
GA
 | S
AGCT
D
GAC
KAAA
S
AAG
KAAG
CAG
GAC
CAG | S
TCT
V
GTC
S
AGC
K
AGC
K
AAGC
GGC
GGG
T
T
C
C
GGG
C
C
C
GGG
C
C
C
C | S
TCA
T
ACT
S
ACT
S
ACC
T
ACC
S
ACC
N
ACC
N
T
C
C
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
T
S
ACC
S
S
ACC
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
ACC
S
S
A
S
S
S
ACC
S
S
ACC
S
S
S
S | D
GACI
P
CCTI
S
TCCI
P
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
S
CCCI
S
S
S
S | S
AGC ²
K
AAGC
S
AGCO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TC
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
TCTO
 | S
P
CCGC
E
CAGC
L
CTAC
S
S
CAGC
C
CAGC
C
CAGC
C
CAGC
C
C
CAGC
C
C
C
C
C
C
C
C
C
C
C
C
C | D
DATT
P
CCTG
BAGG
P
CCTG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
S
CCTG
S
S
S
CCTG
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S
CTC
E
A
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCCC
S
CCCCCCCC | A SCAL
SACIONAL SCAL
SACAN
SCAL
SAAC
SAAC
SAAC
SAAC
SAAC
SAAC
SAAC
 | 3300 s
AGT::
410 k
AAGC
410 k
AAGC
410 k
AAGC
410 k
AAGC
50 k
AAGC
50 k
AAGC
50 k
AAGC
50 k
AAGC
50 k
AAGC
50 k
50 k
50 k
50 k
50 k
50 k
50 k
50 k | STCC | E
SAGTO
STGC
S
S
CTC
K
S
G
CTC
K
S
G
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | D
SATI
A
SCA
E
SAGI
R
SCA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
E
SAA
CA
CA
E
SAA
CA
E
SAA
CA
CA
E
SAA
CA
CA
CA
CA
CA
CA
CA
CA
CA
 | E GAAG
T ACTO
E GAAG
SCTO
A GOGO
K A GOGO
K A A A A A A A A A A A A A A A A A A A | A
SCT
P
CCT.
E
SAG
SGT
E
SAG
SGT
E
SAG
SGT
R
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
S
SGT
A
SGT
A
SGT
A
SGT
A
S
SGT
A
S
S
SGT
A
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | K
AAG
GGGA
GGGA
K
CAA
SCAA
AAAA
T
AAAAA
A
CAA
A
SCAA | |
 | R
AAGO
X
AAGO
V
STTC
R
S
AGC
V
S
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
S
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC | J SO J SO 420 A G 420 A G 430 SCGGGC 450 K S SMGCTCE S10 510 F SCCAAA S00 SCCAAA S00 E F SCCAAA S00 E A SCCAAA S00 SCCAAA S00 E A SCCAAA S00 E A GCCAAA S00 E A SCCAAA S00 E A SCCCAAA S00 SCCCAAA | 1222
1312
1402
1492
1582
1672
1852
1942
2032 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGGCCAC
V T T
GTGACAAC
D S E
GAGTCAGAC
P T P
CCCCACCCC
CCCCCCCCC
CCCACCCC
E E D
GAGCAAACGC
A T P
GAGCAAACGC
R R V
CGAAGGGTC | P
TCCAI
K
CAAGU
K
SAAAA
P
CCCTI
S
SAGT'
T
TACTO
T
TACTO
T
CAAGO | V
STG.
S
AGTI
P
CCTI
K
AAGU
S
S
CCCI
E
S
AGC
S
I
CTI
E
S
AGC | к
мада
Р
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
О
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
ССССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССТ
СССС |
K
L
L
S
GCJ
R
GCG
S
GCA
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A | K
AGG
R
R
GAJ
V
TTGJ
E
GAGC
C
AGC
C
AGC
C
AGC
C
AGC
I
TTTA | A GO K AAG TAC E GA A GO A GA K AAG TAC E GA GA CA K AG TAC E GA GA CA K AG CA K AG CA |
E
GAG
400
P
FCCA
430
A
GCCC
430
A
GCCC
450
E
GAG
520
E
GAG
550
GAG
550
GAG
550
C
A
GCA
400
P
A
A
GCCCA
400
A
A
GCCCA
400
P
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
A
A
GCCCA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
400
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
60
C
CA
C
CA | S
A
GCT
D
GAC
K
AAA
K
AAG
S
CAG
CAG
GAC | S
TCT
V
GTC
S
AGC
A
AGC
A
AGC
A
AGC
A
AGC
C
G
G
G
G
G | S
TCA
T
ACT
S
ACC
S
C
ACC
S
C
ACC
S
C
C
C
C
C
C
C
C | D
GACI
P
CCTI
S
TCCI
P
CCCC
P
CCCC
S
S
S
S
S
S
S
S
S
S
S
S
S | S
AGC
K
AGC
S
TCTO
P
CCT/
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
TCAC
P
CCGC
E
EAGC
L
CTAC
K
K
NAAC
S
S
CAG
C
T
T
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | D
DATT
P
CCTG
E
DAGG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
A
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
C
C
C
C
C
C
C
S
C
C
S
C
C
S
C
C
S
C
C
S
C
C
S
C
S
C
S
C
S
S
C
S
S
C
S
S
C
S
S
C
S
S
C
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S |
S
CTC
E
A
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCCC
S
CCCCCCC
S
CCCCCCCC | D SACI
SACI
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL | 3300 s
AGT:
410 k
AGG
410 k
AGG
410 k
AGG
410 k
AGG
500 k
500 k
50 | S
S
S
S
S
S
S
S
S
S
S
S
S
S | E
SAGTO
STGC
S
S
CTC
K
S
S
CTC
K
S
S
CTC
K
S
S
CTC
K
S
S
CTC
C
K
S
S
CTC
C
K
S
S
CTC
C
K
S
S
CTC
C
S
S
S
CTC
C
S
S
S
S
S
S
S
S
S |
D
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SATI
SAT | E
ACTO
E
SAAC
ACTO
E
SAAC
A
SCAC
E
SAAC
E
SAAC
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
C
A
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SC | A
SCT
P
CCT.
E
SAGI
G
G
G
G
G
G
G
G
G
G
G
G
G | R
AAG
GGGA
A
GGGA
A
AAAJ
ACAC
A
ACAC
A
ACAC
 | | R
AAGO
X
AAGO
S
S
TTTC
R
R
S
S
TTTC
R
S
S
TTTC
S
S
TTTC
S
S
TTTC
S
S
TTTC
S
S
TTTC
S
S
TTTC
S
S
S
S | 5 30
CCAGTCC
420
420
420
420
420
420
420
420
 | 1222
1312
1402
1492
1592
1672
1962
1942
2032 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GGACACCCC
D S E
GACTCAGAM
P T P
CCCACCCCC
S K A
AGCAAAGCC
<u>E E D</u>
GGAACTCCC
R R V
CGAAGGCTCC
R R V
CGAAGGCTCC | P
TCCAI
K
CAAGJ
K
GAAAA
P
CCCTJ
SAGAC
V
V
TGTGJ
T
TACTC
T
TCAAT
R
CAGGC | V
STG.
S
AGTI
P
CCTI
K
AAGI
S
CCCI
E
CCCI
E
CCCI
E
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
A
SCCA
K
AGCA
K
AAAC
Q
CAGA
E
SAGG | K
L
MAN
S
GGCJ
R
GGGC
S
GGC
S
GGC
S
GGC
S
C
AC
A
S
C
AC
A
S
C
AC
A
S
C
AC
A
S
 | K
AGCJ
R
R
AGCJ
E
E
AGCA
C
Q
C
AGC
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X | A CC K AG K AG E CG A CC A CC K AG E CG E CG A CC A CC A CC A CC A CC A | E
GAG
400
P
GCAA
430
A
GCCT
460
A
CAG
520
E
GAG
550
G
GA
550
C
GGA
550
C
GGA
610
C
C
GGA
610
C
C
A
GCCT
610
C
A
C
C
A
C
A
C
C
A
C
A
C
A
C
A
C
A
C
 | S
AGCT
D
GAC
KAAA
S
AGC
KAAA
G
G
CAG
G
AC
K | S
TCT
V
GTC
S
AGCA
AGCA
AAGC
S
AGCA
AAGC
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | S
TCA
T
ACT
S
AGC
A
GCA
T
ACG
S
AGC
N
ATI
T
ACCI
P
C
C
A | D
GACI
P
CCCTI
S
TCCI
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCCI
S
CCCCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S | S
AGC'
K
AAGK
S
AGCO
S
TCTO
P
C
C
T
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
TCAC
P
CCGC
E
CAGC
L
CTAC
S
CAGC
T
T
T
C
CAGC
D
C
CAC
D
C
CAC
D
C
CAC
D
C
CAC
C
C
C | D
DATT
P
CCTG
E
SAGG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
S
CCTG
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S
CTC
E
A
CTC
E
A
CTC
C
C
C
C
C
C
C
C
C
C
C
C
C | A SCAL
 | 380
8
AGT:
4
4
4
4
4
10
8
4
4
10
8
4
10
8
4
10
8
4
10
8
4
10
8
4
10
8
10
10
10
10
10
10
10
10
10
10 | STOCK | E SAGO
V STGO
STGO
S STGO
K AGO
S SO
S S S S S S S S S S S S S S S S S |
D
SATI
A
SCAL
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
E
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI
SAGI | E GAAG
T ACTO
E E GAAG
SCTO
A GCAG
E GAAG
E GAAG
C A C A C A C A C A C A C A C A C A C A | A SCT
P
CCT.
E
SAGI
G
G
G
G
G
G
G
G
G
G
G
G
G | R
K
AAG
GCCi
GGA
A
GCCi
GGA
K
C
GGA
K
C
C
A
A
C
C
C
C
C
C
C
C
C
C
C
C
C
 | A GCC.
Q CAGI
T ACCI
GACI
A GO
A GO | R
AAGO
X
AAGJ
S
AGC/
V
SJITC
E
DAGC
D
DATC
S
CCC
D
D
ACT | 5 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | 1222
1312
1402
1492
1582
1672
1852
1942
2032 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGA
P T P
CCCCCCCCC
S K A
AGCAAAGC
CCCACCCCC
R R V
CCGAAGGGTC
E R R V | P
TCCAI
K
CAAGU
K
SAAAA
P
CCCTI
S
SAGT'
V
TCTGJ
T
TACTO
T
TACTO
T
CAAAT
R
CAAGO | V
STG.
S
MGTI
P
CCTI
K
MAGI
S
FCCI
E
S
AGI
S
S
I
CTI
E
S
AGI | K
AAGA
P
CCCT
Q
CAGA
AGCCA
AGCCA
K
AAAG
CAGA
K
AAAG
X
CAGA
K
AAAG
X
CAGA
K
AAAG
X
CAGA
K
AAAG
X
CAGA
A
AGCCA
K
AAAG
X
AGCCA
K
AAAG
X
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AGCCA
K
AAAG
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
AAAG
A
A
A
A
A
A
A
A
A
A
A
A
A |
K
L
L
S
GCJ
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGC
R
GCGCC
R
GCGC
R
GCGC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
R
GCAC
C
C
C
C
C
C
C
C
C
C
C
C
C | K
AGCJ
R
R
AGAJ
V
TTGJ
E
SAGC
Q
CAGC
V
TTJ
I
I
TTJ
I
K | A GC K AG K AG E GAG F | E
GAG
400
P
iccA
430
A
iGCT
460
A
iGCT
520
E
GAG
550
E
GAG
550
E
GAG
610
L
iCCA
610
L
CTCI
640
C
T
 | S
GCT
GAC
K
AAA
GCC
CAG
CAG
CAG
CAG
CAG
CAG | S
TCT
V
GTC
S
AGC
A
AGC
A
AGC
A
AGC
A
AGC
A
A
C
C
G
G
G
G | STCA
TCA
TACT
SAGC
SCA
SGCA
SGCA
NATI
TCCCI
RCCA
RCCA
RCCA | D
GACI
P
CCTI
S
TCCI
P
CCCC
P
CCCC
S
G
CCCC
S
G
CCCC
S
G
CCCC
S
S
G
CCCC
S
S
G
CCCC
S
S
G
CCCC
S
S
G
CCTI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCI
S
TCCCI | S
AGC
K
AAGC
S
AGCO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
TCTO
 | S
TCAC
P
CCGC
E
CAGC
L
CTAC
K
K
K
K
K
K
CTAC
S
CAGC
P
CCAG
G
C
TTC
D
CACA
R
R | D
DATT
P
CCTG
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D | A CTC
A CTC
E AGJ
A CCC
S CCC
S CCC
S CCC
C CCC
C CCC
S CCCC
S CCC
S CCCCC
S CCC
S CCC
S CCC
S CCC
S CCCC
S CCC
S CCC
S CCCC
S CCCCC
S CCCCC
S CCCCC
S CCCCCCCCCC | A GGI
 | 3300 s
410 s
410 k
410 s
410 s
40 s | S TCCC
A SCAC
S TCCC
A SCTC
K AGU
S TCCC
C A SCTC
K AGU
C A SCCC
K AGU
A SCCC
K AGU | E SAGC
V STGC
STGC
STGC
CTC
K AGO
SGAG
GAG
K AGO | D D D D D D D D D D D D D D D D D D D | E GAAG
T ACTO
E GAAG
SCTO
A SCTO
A SC | A
SCT
P
CCT.
E
SGT
G
SGT
G
SGT
C
C
C
C
C
C
C
C
C
C
C
C
C
 | R
AAG
GCCG
GGA
GGGA
K
GGGA
K
GGGA
K
GGAA
A
CACAC
A
CACAC
A
CACAC
A
CACAC
A
CACAC
A
CACAC
A
CACAC
A
CACAC
A
CACAC
A
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC
CACAC | A GCC.
Q CAGI
T ACCC
GACI
GACI
K AGJ
A GCC
A GCCC
A GCCC
A GCCC
A GCC
A GCCC
A GCCC
A GCCC
A GCCC
A GCCC | R
AAGO
X
AAGO
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
AGC/
X
X
AGC/
X
AGC/
X
AGC/
X
X
AGC/
X
X
AGC/
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
AGC/
X
X
X
AGC/
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X
X | 5 9 V
CCAGTC
420
A G
CCCGCCC
450
K S
480
S S
S S
S S
S S
S S
S S
S S
S
 | 1222
1312
1402
1492
1592
1672
1952
1952
1942
2032 |
| K P A
AAACCAGC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAACC
D S E
GACTAGAA
P T P
CCCAACCCA
CCCACCCC
E E D
CCACCAACC
E R V
CCCAACGCTC
E R A
GACCAACCC | P
TCCAI
K
CAAGI
K
SAAAA
P
CCCTI
SAGT
T
TACTO
T
TACTO
T
TACTO
R
R
CAGGO
N
SAATO | V
STG.
S
MGTI
P
CCTI
K
MAGI
S
FCCI
E
S
AGC
Q
CAGC | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
A
SCCA
K
AAAG
X
CAGA
K
AAAG
X
CAGA
K
AAAG
X
CAGA
K
AAAG
X
CAGA
K
AAAG
X
SCCA
K
AAGA
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
A
SCCA
S
S
S
A
SCCA
S
S
S
A
SCCA
S
S
S
S | K
L
MAN
S
MGCJ
R
MGCJ
R
R
MGCJ
R
R
MGCJ
R
R
MGCJ
R
R
MGCJ
R
R
MGCJ
R
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
MGCJ
R
M
MGCJ
R
M
M
M
M
M
M
M
M
M
M
M
M
M
M
M
M
M
M
 | K
AAGC
S
AGCJ
R
AGAJ
E
E
SAGC
R
AGAJ
E
E
SAGC
X
C
AGC
TTTA
I
I
TTTA
I
I
I
I
I
I
I
I
I
I
I
I
I | A GOC KAG KAG TOC EGA ACC A CAG KAG FIC | E
GAG
400
P
CCA
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
GCC
430
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
GCC
400
A
A
GCC
400
A
A
GCC
400
A
A
GCC
400
A
A
A
A
A
A
A
A
A
A
A
A
A
 | S
A
GCT
D
GAC
K
AAA
S
GCC
CAG
GAC
K
AAA | S
TCT
V
GTC
S
AGC
K
AAG
S
AGC
C
AAG
G
GGG
C
T
C
G
GGG
C
C
GGG
C
C
C
GGG
C | S
TCA
T
ACT
S
AGC
A
CGA
S
AGC
A
AGC
A
AGC
A
S
AGC
A
S
AGC
A
S
AGC
A
S
AGC
A
S
AGC
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
A
S
C
S
S
S
C
S
S
S
S | D
GACI
P
CCTI
S
TCCI
P
CCCC
S
CCCC
S
G
G
G
G
G
G
G
G
C
C
C
C
C | SAGC
KAAGC
SAGC
STCT
STCT
STCT
STCT
STCT
SCCT
STCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
SCCT
 | S
TCAC
P
CCGC
E
CTAC
S
CTAC
S
CCAC
S
CCAC
C
C
CCAC
C
C
C
C
C
C
C | D
CATT
P
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
CCTG
C
C
CCTG
C
C
C
C | S CCCC
S CCCCC
S CCCC
S CCCCC
S CCCC
S CCCCC
S CCCCC
S CCCCC
S CCCCC
S CCCCCCC
S CCCCCCCCCC | A AGE A AGE A A A A A A A A A A A A A A
 | 3300 s
AGT:
410 k
AAGC
410 k
AGC
410 k
AGC
410 k
AGC
500 k
AGC
50 | STOCK
AGAN
STOCK
AGAN
STOCK
AGAN
STOCK
AGAN
AGAN
AGAN | E SAGO
STGC STGC
STGC SCTC
KAGO
SGAO
GGAO
KAGO
KAGO | D
SATI
A SCAL
E SAGI
CAC
H
CAC
H
CAC
H
CAC
H
CAC
H
CAC
H
CAC | E GAAG
T T C E GAAG
E GAAG
E GAAG
E GAAG
R L G GAG
 | A GCT
FCT
E GGT
GGT
GGT
GGT
E GGT
GGT
R GGC
A GGC
GGC
GGC
A GGC | ECTI
KAAGI
GCCI
GGAI
AGCCI
KAAAI
ECAAI
ACACA
AGCA
SCTC
SCTC | A GCC.
Q CAGE
T ACCI
GACI
A ACCI
A ACCI | R AAGO
P COTO
X AAGJ
S AGCJ
V V
STTC
E CO
D SACT
R COGO | 5 9 V
CRAGTCC
420
A G
450
K S
480
K S
MGCTCC
510
P K
570
P K
570
F K
570
F K
630
P F
630
P F
630
G G
G G
G G
G G
G G
G G
G G
G
 | 12222
1312
1402
1492
1582
1672
1762
1942
2032
2122 |
| K P A
AAACCAOC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTAGAU
P T P
CCCACCCCC
S K A
AGCAAAGC
E E D
GAGGAAGAC
R R V
CCAAGGGCT
E R A
GGACGAGCC | P
TCCAI
K
CAAGI
K
GAAAG
P
CCCTI
S
SAGT
T
TCCTG
T
TACTO
T
TACTO
T
TACTO
T
TACTO
T
TACTO
T
TACTO
T
TACTO
T
TACTO
T
TACAG | V
STG.
S
MGTI
P
CCTI
K
MAGU
S
S
CCCI
K
MAGU
S
S
CCCI
S
AGC
S
AGC
S
AGC
S
AGC
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
S
S
AGCA
K
AAAG
X
AAAC
Q
CAGA
K
AAAC
Q
CAGA
K
AAAC
Q
CAGA
K
AAAG
X
S
CCT
T
CCT
T
C
CCT
T
C
C
C
C
C
C
C
C
C | K
L
S
S
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
 | K
AGCJ
R
R
AGCJ
E
SAGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGCJ
C
AGC
C
AGCJ
C
AGC
C
AGCJ
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
C
AGC
AG | A GO KAG KAG KAG KAG KAG FIT | E
GAG
400
P
FCCA
430
A
GCC
430
A
GCC
430
A
C
GCC
430
A
C
GCC
430
A
C
GCC
430
A
C
GCC
430
A
C
GCC
430
A
C
GCC
430
A
C
C
C
A
C
A
C
A
C
C
A
C
A
C
A
C
A
C
 | S
AGCT
D
GAC
KAAA
S
CAG
CAG
CAG
CAG
CAG
CAG
CAG
CAG
CAG | S
TCT
V
GTC:
S
GCA
K
AGC:
A
AGC:
A
AGC:
T
C
GGGA
C
GGGA | S
TCA
T
ACT
S
AGC
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
A
C
C
A
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
A
C
C
A
C
A
C
C
A
C
A
C
C
A
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
C
A
C
C
A
C
C
A
C
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
C
A
C
C
C
A
C
C
A
C
C
A
C
C
A
C
C
A
C
C
C
A
C
C
C
C
C
A
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | D
GACI
P
CCCTI
S
TCCI
P
CCCC
P
CCCC
S
CCCC
S
G
G
G
G
G
G
G
G
G
G
G
G
G | S
AGC
K
AAGC
S
AGC
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
TCTO
 | S
FCAC
P
CCGC
E
CCTAC
E
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
S
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC
CCTAC | D
DATT
P
CCCG
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D | A COLOR | A SCAL
 | 380
8 AGT:
410
8 AGC
410
8 AGC
410
8 AGC
410
8 AGC
410
8 AGC
410
8 AGC
410
8 AGC
410
8 AGC
8 AGC | SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO
SICCO | E
SAGC
V
STGC
S
CTC
K
S
GTC
K
S
GTC
K
S
GTC
K
S
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G
G | D TA A SCALE SAGE REGALE ACTION A SCALE SAGE REGALE ACTION | E GAAG
 | A GCT
P CCT.
E GGT
G GGT
G GGT
E GGT
G GGT
B A GG
A A GGC
G GGC
G GGC | R
CCTI
K
AAG
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCC
CCCCC
CCCCC
CCCCC
CCCCC
CCCCC
CCCCCC | | R AAGO
P COTO
X AAGJ
S AGC
V STTTC
R COTO
D AATO
S S COTO
D AATO
S S COTO
R COTO
R COTO
R COTO
R COTO
 | 5 9 V
CRASTC 420
A G
420
A G
450
K S
S
S
S
CCTCC
CCAAA
510
P K
- 510
P C
CCAA
510
P C
CCAA
600
8 S
S
CCCAAA
6 S
CCCAAA
6 S
CCCAA
6 S
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA
CCCAA | 12222
1312
1402
1492
1592
1672
1952
1952
1942
2032
2122 |
| K P A
AAACCACC
S A T
AGTGCCAC
S G Q
AGTGGCAL
V T T
GTGACAAAC
P T P
CCCACCCCC
S K A
AGCAAACC
E E D
GACGAACCC
CAACCAAC
CAACCAACC
E R V
CCGAACGCCT
E R A
GGGCGACCCC
S I S | P
TCCAI
K
CAAGI
K
SAAAA
P
CCCTI
SAAAA
V
TCTGJ
T
TACTO
TCAAA
CAGGO
N
SAATO
V | V
STG.
SMGTI
PCCTI
KAGG
STCCI
KAGG
STCCI
ECCI
ECCI
ECCI
ECCI
ECCI
ECCI
ECCI | K
AAGA
P
CCCT
Q
CAGA
A
BCCA
A
BCCA
S
A
GCCA
S
A
GCCA
S
A
GCCA
S
A
GCCA
A
A
BCCA
S
A
GCCA
A
A
BCCA
S
A
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
A
BCCCA
S
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
CAGA
A
C
CAGA
A
C
CAGA
A
C
CAGA
A
C
CAGA
A
C
CAGA
A
C
CAGA
A
C
C
CAGA
A
C
C
C
C | K
L
TAJ
S
GCJ
S
GCAC
N
AAAG
L
TGA
N
 | KAGC
SAGCI
RAGAJ
E SAGC
Q
CAGC
V
AGC
SAGC
V
AGC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC | A GC KAG KAG TCC EGA ACC ACA KAG EGA FTC. | E
GAG
400
P
CCA
430
A
GCC
430
C
CAG
520
E
GAG
550
C
CAG
550
C
CAG
550
C
CAG
550
C
CAG
610
C
C
CAG
610
C
C
CAG
610
C
C
CAG
610
C
C
C
CAG
C
CAG
C
C
CAG
C
C
CAG
C
C
CAG
C
C
CAG
C
C
CAG
C
C
C
C
 | S
AGC
AGC
A
AAA
AAA
AAA
AAA
AAA
AAA
AAA
A | S
TCT
V
GTC
S
AGC
K
AAGC
K
AAGC
AAGC
GGGJ
T
C
GGGJ
D | STCA
TCA
ACT
ACT
ACC
ACC
ACC
ACC
ACC
ACC | D
GACI
P
CCCTI
S
CCCT
P
CCCC
P
CCCC
S
S
CCCT
S
S
CCCT
S
S
CCCT
S
S
CCCT
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
S
CCCTI
S
S
S
CCCTI
S
S
S
C
CCCTI
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
C
C
C
S
S
S
S
C
C
C
S
S
S
S | S
AGC
K
AAGC
S
AGCO
S
TCTO
S
TCTO
S
TCTO
S
TCTO
T
TCTO
 | S
TCAC
P
CCGC
E
E
SAGC
L
CTAC
S
CTAC
S
G
CTAC
S
CAG
C
TTC
D
SACA
R
CGCC | D
CATT
P
CCTG
SAGG
P
CCTG
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTTG
P
CCTG
S
CCTG
S
CCTG
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
S
CCTG
S
CCTG
S
S
CCTG
S
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
CCTG
S
C
C
C
C
S
C
C
C
C
S
C
C
C
C
S
C
C
C
S
C
C
S
S
C
C
S
S
C
C
S
S
C
S
S
C
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S CCCC
S CCCCC
S CCCC
S CCCCCC
S CCCCC
S CCCC
S CCCCC
S CCCCC
S CCCCC
S CCCCC
S CCCCC
S CCCCCC
S CCCCCC
S CCCCC
S CCCCCCCC | A BCAL
BCAL
BCAL
BCAL
BCAL
BCAL
BCAL
BCAL
 | 380
SAGT:
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
410
K
5
K
5
K
5 | S TOCK
A SCAC
S S TOCK
A SCAC
S S TOCK
S S TOCK
S S S S S S S S S S S S S S S S S S S | E SAGC
V STGC
S SGTC
K AGC
SGTC
K AGC
C K AGC
K AGC | D SATI
 | E GAAG
T ACTO E GAAG
SGAAG
A GCCAG
E GAAG
C A GCCA
R GCCAG
R G | A GCT
P CCT.
E GGT
G GGT
C G GGT
E GGT
C G GGT
R GGC
A GGC
G GGC | ECTI
KAAGI
GCCI
GGGAI
AGCCI
AAAJ
ECAI
ACAI
ACAI
ACAI
ACAI
ACAI
ACAI
ACAI | | K
MAGO
P
CCTC
X
AGO
STTC
STTC
STTC
STTC
STTC
STTC
STTC
STT
 | B V CRASTC 420 A G 420 A G 420 A G A G Coccocc 450 K S S S S GUE S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S <td>1222
1312
1402
1492
1582
1672
1762
1942
2032
2122</td> | 1222
1312
1402
1492
1582
1672
1762
1942
2032
2122 |
| K P A
AAACCACC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
D S E
GACTCACAC
D S E
GACTCACACC
S K A
AGCAAAACC
E E D
GCAACTCCC
R R V
CGAACGCCC
S I S
GCACCACCCC
S I S
CCACACCCC | P
TCCAI
K
CAAGI
K
SAAAC
P
CCCTI
SAAAC
C
CAAC
C
CAAC
C
CAAC
C
C
CAAC
C
C
CAAC
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | V
STG,
STG,
P
CTI
KAG
STCCI
KAG
STCCI
KAG
STCCI
CO
STCCI
CO
STCCI
CO
STCCI
CO
STCCI
CO
STCCI
CO
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STCCI
STC | К
МАДА
Р
СССТ
О
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
А
САДА
САДА
А
САДА
А
САДА
А
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
САДА
С
САДА
С
САДА
С
С
С
С
С
С
С
С
С
С
С
С
С |
K
L
TAJ
S
GCJ
S
GCAC
S
GCAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
CAC
S
C
CAC
S
C
S | K
AAGC
S
AGC
E
SAGC
E
SAGC
A
C
AGC
C
AGC
C
AGC
C
AGC
S
C
AGC
S
C
AGC
S
C
AGC
S
S
C
AGC
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | A CC KAG KAG TAC EGG ACC A CA KAG FIT | E
GAG
400
P
CCA
430
A
GCC
490
CA
520
E
GAG
520
E
GAG
520
E
GAG
520
E
GAG
520
C
C
GAG
520
E
GAG
490
A
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
AGCT
AAAA
GAC
KAAA
AAG
GAC
CAG
GAC
KAAA
FTT | STOT
VOGTO
SAGO
KAAG
AGCA
AAGCA
AAGCA
GGGGA
TOTC
GGGA
ACCO
GGGA | S TCA
TCA
ACT
ACT
ACC
ACC
ACC
ACC
ACC
ACC | D
GACI
P
CCCTI
S
TCCCI
P
CCCCT
P
CCCCT
S
S
CCCCT
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
TCCCI
S
S
TCCCI
S
S
CCCTI
S
S
TCCCI
S
S
CCCTI
S
S
TCCCI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
CCCTI
S
S
S
CCCTI
S
S
S
CCCTI
S
S
S
CCCTI
S
S
S
CCCTI
S
S
S
CCCTI
S
S
S
S
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
C
CCCTI
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S
AGC ⁴
K
AAGC
S
AGCC
S
S
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
FCAC
P
CCGC
E
E
CCGC
E
CCAC
S
CCAC
C
CCAC
C
CCAC
C
CCAC
C
CCAC
C
C
CCGC
C
C
C
C
C
C
C
C
C
C
C
C
C | D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D
D | S CCC S CCCC S CCC S CCCC S CCC S CC | A BCAN
BCAN
BCAN
BCAN
BCAN
BCAN
BCAN
BCAN
 | 3400 SAGC
410 SAGC
410 AAGC
410 AAGC
410 AAGC
410 AAGC
410 AAGC
410 AAGC
500 AAGC
500 SATO
500 S | SICCO
A SCAR
SCAR
SCAR
SCAR
SCAR
SCAR
SCAR
SCAR | E SAGC
V STGC
SGTC
V STGC
K AGC
SGTC
K AGC
K AGC
K AGC | D SATI | E GAAG
T ACTO
E GAAG
SCAC
E GAAG
E GAAG
C C C C C C C C C C C C C C C C C C C
 | A
GCT
P
CCT.
E
GGT
GGT
GGT
C
GGT
C
GGT
C
GGT
C
GGT
C
GGT
C
GGT
C
C
GGT
C
C
GGT
C
C
C
C
C
C
C
C
C
C
C
C
C | R
CCT
K
AAG
CCC
CCC
CCC
CCC
CCC
CCC
CCC | | K AAGO
P ECTO
X AAGJ
S AGC/
V 57TTC
K C
S C
S C
S C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | 5 9 V
CCASTC 420
A G
420 A G
450 A
480 S
5 10 P K
480 S
5 10 P K
5 | 1222
1312
1402
1492
1592
1672
1762
1942
2032
2122 |
| K P A
AAACCACC
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAAC
P T P
CCCACCCCC
S K A
AGCAAAGC
CCACCCCC
S K A
AGCAAGCCCCC
C R R V
CCGAAGGGCC
E R A
GGACGAGGCC
S I S S
TCCATCTCCC | p
r
r
r
r
r
r
r
r
r
r
r
r
r | V
STG.
S
MGTH
P
CCTI
K
MAGK
S
CCC
K
MAGK
S
CCC
S
S
CCC
S
S
CCC
S
S
CCC
S
S
CCC
S
S
CCC
S
S
S
CCC
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
S
AGCA
K
AAAC
Q
CAGA
K
AAAC
Q
CAGA
V
CTTC
V
V
TTCA | K
L
MAN
S
MGGC
S
MGGC
S
MGGC
S
MGAC
S
MGAC
S
MACA
L
TGA
N
ATT
 | KAGC
SAGCI
RAGAJ
E
SAGC
QAGC
VITGJ
E
SAGC
QAGC
VITGJ
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC | A CC KAG KAG TAC EGG ACC KAG KAG FTO VITO | E
GAG
400
P
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
430
A
GCC
450
C
450
C
460
A
520
C
520
C
520
C
520
C
520
C
520
C
520
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
530
C
7
7
7
7
7
7
7
7
7
7
7
7
7
 | S AGC
AGCT
D CAA
AAA
AAA
AAA
C C C C C C C C C C C C | S
TCT
V
GTC.
S
GCC.
S
AGCA
K
AAGC
S
GGCA
T
C
GGGA
C
S
GGGA
C
S
GGA
C
S
ACC | S TCA
TCA
T ACT
S AGC
AGC
T ACG
S AGC
NATI
T CCI
R AGC
S CTI
R AGC
S CTI
R AGC | D
GACI
P
CCCTI
S
CCCTI
P
CCCCI
P
CCCCI
P
CCCCI
S
CCCCI
P
CCCCI
S
CCCCI
P
CCCCI
P
CCCCI
S
CCCCI
P
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
P
CCCTI
S
CCCI
P
CCCTI
S
CCCI
P
CCCTI
S
CCCI
P
CCCTI
S
CCCI
P
CCCI
S
CCCI
P
CCCCI
P
CCCCI
P
CCCCI
P
CCCCI
P
CCCCI
P
CCCCI
P
CCCCI
S
S
CCCI
P
CCCCI
S
S
CCCI
P
CCCCI
S
S
CCCI
P
CCCCI
P
CCCCI
S
S
CCCI
P
CCCCI
S
S
CCCI
S
S
CCCI
P
CCCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
CCCI
S
S
S
CCCI
S
S
S
CCCI
S
S
S
CCCI
S
S
S
CCCI
S
S
S
CCCI
S
S
S
CCCI
S
S
S
S | S
AGC ⁴
K
AAGC
S
AGC
S
TCTC
P
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
FCAC
P
CCGC
E
CCGC
E
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC
CCAC | P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCTG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
P
CCCG
CCCG
P
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCCG
CCC | SICTO
A
E
E
A
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | A BCAL
SACL
SACL
SACL
SACL
SACL
SACL
SACL
S
 | 3300 SAGT | S ICTO
A SCAC
SICCI
K AGI
SCAC
K AGI
K I
AGI
K AGI
K AGI | E SAGC
V STGC
S SGTC
K AGC
S GTC
K AGC
S GTC
K AGC
K AGC | D SATI
 | E GAAG | A
GCT
P
CCT.
E
GGT
GGT
GGT
GGT
GGT
GGT
GGT | R
CCT
K
AAG
GGGA
A
GGGA
A
GGGA
A
CCA
CC |
 | K
AAGC
P
CCTC
K
AAGC
S
S
AGC
V
S
S
TTC
S
S
AGC
D
S
AGC
D
S
AGC
C
D
S
AGC
C
C
S
S
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | F V CRASTC 420 A G 420 A G 32050502 450 K 5 S S S 5 S S S 5 F F K 5 F F K 5 CCAAG S F 5 F F K' 5 CCAAG S S 5 F F K' 5 CCAAG S S 5 TO K G 5 GCAAG S S 5 SCCAAG G G CCTTTC GGCGCT GGOG G G G G G G GCAAG GCAGCC GCAGCC GCAGCC | 12222
1312
1402
1492
1582
1672
1962
2032
2122
22122 |
| K P AC
AAACCACK
S A T
AGTGGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
P T P
CCCACCCCC
S K A
AGCGAACCC
S K A
AGCGAACCC
C
CCACCCCC
B R V
CCCACGGC
S I S
TCCATCTCT
CACTACCT | P
TCCAI
K
CAAGI
K
SAAAA
P
CCCTI
S
SAGTC
T
CAAAT
C
CAAAT
C
V
CCAAT | V DTG.
S MGTI
P DTTG.
K MAG
S TCCI
K MGTI
S TCCI
S | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
S
AGCA
K
AAAG
S
AGCA
K
AAAG
C
AGCA
V
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | K
L
MAN
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
MGC
S
M
M
S
MGC
S
MGC
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
S
M
M
S
M
M
S
M
M
S
M
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
M
S
S
M
S
S
M
S
M
S
M
S
S
M
S
S
M
S
S
M
S
S
M
S
S
M
S
S
M
S
S
M
S
S
M
S
S
S
M
S
S
M
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | K
AGCI
R
AGCI
E
SAGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
AGC
 | A GO K KAG TAC E GA ACO A CA KAG E GA FTO TOTO | E
GAG
400
P
CCA
430
A
GCT
490
A
CGCC
490
C
CAG
520
C
CAG
550
C
CAG
550
C
CAG
550
C
CAG
550
C
CAG
610
C
C
CAG
A
C
CAG
A
C
C
CAG
A
C
C
CAG
A
C
C
CAG
A
C
C
CAG
A
C
C
CAG
A
C
C
C
C | S
AGCT
D
GAC
KAAA
S
AGC
C
A
GAC
KAAA
S
GCC
C
A
GAC
KAAA
F
TTT
GTT. | S
TCT
V
GTC
S
AGC
K
AAG
S
AGCA
AAG
GGGA
ACC
GGGA
C
S
ACC
MTT | S
TCA
T
ACT
S
ACT
S
CCA
T
ACC
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
CCA
S
S
S
CCA
S
S
S
CCA
S
S
S
CCA
S
S
S
S |
D
GACI
P
CCCTI
S
CCCC
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCTI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
C
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
C
CCCI
S
C
CCCI
S
CCCI
S
CCCI
S
C
CCCI
S
CCCI
S
C
CCCI
S
C
CCCI
S
CCCI
S
C
CCCI
S
C
CCCI
S
CCCI
S
CCCI
S
C
CCCI
S
C
CCCI
S
CCCI
S
C
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCCI
S
CCCCI
S
CCCI
S
CCCI
S
CCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S
CCCCI
S | S
AGC ²
K
AGC
S
CT
CT
S
CT
S
CT
S
CT
S
CT
S
CT
S
CT | S
FCAC
P
CCGC
E
CCGC
E
CCGC
K
ANACC
S
G
CCGC
C
CCGC
C
CCGC
C
CCGC
C
C
CCGC
C
C
C
C
C
C
C
C
C
C
C
C
C
 | D
DATT
P
CCTG
E
DAGG
P
CCTG
P
CCCC
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTT
S
CCTG
CCTG | S CTC | A GCAL
SACA
SCAL
SCAL
SCAL
SCAL
SCAL
SCAL
S | 3300 SATC
410 K
410 K | S ICTO
A GCA
S ICCO
A ICCO
K ICCO
G GAN
K ICCO
K ICCO
K
ICCO
K ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
ICCO
K
IC | E
SAGC
SAGC
STG
K
STG
K
SG
SAGC
SAGC
K
SG
SAGC
K
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC | D DATA
 | E GAAG | A
SCT
P
CCT
E
SGT
G
SGT
E
SGT
E
SGT
E
SGT
R
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SGT
A
SCA
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
C
A
SCA
SCA
C
A
SCA
SCA
C
A
SCA
SCA
SCA
SCA
SCA
SCA
SCA
SCA
SCA
SCA | ECTI
KAAGI
GGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGGA
KGGA
KGGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGGA
KGA
K
 | A GCC.
Q CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CAGI
CA | K
AAGO
P
CCTC
K
AAGJ
S
S
CCCC
D
S
AGC
D
S
AGC
D
S
AGC
D
S
AGC
D
S
AGC
S
S
CCCC
C
S
AGC
S
S
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | F ν CASTC 420 A G 420 A G 3000 GGGGGC 450 S S S S GGGGCC 510 S P K S S S S GGGCTCC 510 S S S S GGGCTCC 510 S S S S CCAAG 630 P K S S S S AGGAA 600 C P F K S S GGGGGC S S GGGGGC G G G G G G G G G G GGGCTCC GGGCTC |
12222
1312
1402
1492
1582
1672
1762
1942
2032
2122
2122
2302 |
| K P A A
AACCACC
S A T
AGTOCCAC
S G Q
AGTOGCCAC
Q
T T
GTGACCACC
Q
T T
GTGACCACC
Q
D S E
GACTCAGAN
P T P
CCCACCCCC
S K A
AGCCAAGCC
CCACCCCC
R R V
CGAAGCGAC
CGAAGCGAC
S T S
T S
T S
T S
T S
T S
T S
T S
T S
T | P
TCCAI
K
CAAGJ
K
SSAAA
P
CCCTJ
SSACAC
SSACAC
T
TCCAT
CCAT
CCCAT | V
STG,
STG,
P
CCTI
K
AGC
S
CCCI
K
AGC
S
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
E
CCCI
CCCI
E
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CCCI
CC | K
AAGA
P
CCCT
Q
CAGA
A
SCCA
A
SCCA
S
AGCA
K
AAAG
S
AGCA
K
AAAG
S
AGCA
K
AAAG
S
CAGA
C
A
S
C
CAGA
A
S
C
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
C
A
G
A
G |
K
L
L
M
L
S
S
G
G
C
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A
C
A | KAGCI
SAGCI
RAGAJ
EGAG
ESAGC
AGA
CAGA
CAGA
TTTA
I
I
I
I
I
I
I
I
I
I
I
I
I
I
I
I
I | A GO K K K K K K K K K K K K K K K K K K | E
GAGG
4000
A
GCCA
430
A
GCCA
460
A
GCAG
520
GAGG
550
GAG
550
GAG
550
GAG
610
L
CTCI
6400
F
CAG
610
C
T
GCA
610
C
T
GCA
640
A
C
C
CAG
640
A
C
C
CAG
640
A
C
C
CAG
652
C
C
A
C
CAG
652
C
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
CAG
652
C
C
CAG
652
C
C
CAG
652
C
C
CAG
652
C
C
C
CAG
650
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
 | S
AGCT
A
GCT
D
GAC
K
AAA
S
GCC
CAG
GAC
K
AAA
F
TTT
SGT | S
TCT
V
GTC
S
GCC
S
AGC
K
AAG
GCC
AAG
GGC
T
C
GGGA
T
C
GGGA
T
C
GGGA
T
TCT
C
GGA
C
M
TTTTTTTTTTTTTT | STCA
TCA
TACT
SAGC
AGCA
TACG
SGCA
TACG
SGCA
NATI
TCCI
R
SGCA
SGCA
SGCA
SGCA
SGCA
SGCA
SGCA
SGCA | D
GACI
P
CCCTI
S
CCCC
P
CCCC
P
CCCC
S
G
CCCC
S
G
CCCC
S
G
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
S
CCCC
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
S
CCCC
S
S
S
CCCC
S
S
S
CCCC
S
S
S
CCCC
S
S
S
CCCC
S
S
S
CCCC
S
S
S
S
CCCC
S
S
S
S
CCCC
S
S
S
S
S
CCCC
S
S
S
S
CCCC
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S
AGC ⁴
K
AGC
S
AGC
S
TCTO
P
TCTO
S
AGC
TCTO
TCTO
ACG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
A
CG1
C
CG1
A
CG1
A
CG1
A
CG1
A
CG1
C
CG1
C
CG1
C
C
CG1
C
C
CG1
C
C
C
C
 | S
FCAC
P
CCGCC
E
CCGCC
E
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGCC
CCGC | D
DATT
P
P
CCTG
E
SAGG
P
CCTG
P
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
CCCC
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S |
S
ICTC
A
ICTC
E
IAGJ
A
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCC
S
CCCCCC
S
CCCCCCCC | A GCAL
BCAL
S SCAL
S SCAL
S SCAL
S SCAL
S S S S S S S S S S S S S S S S S S S | 380
AGT:
410
K AAGC
410
K AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
K
AAGC
410
KAAA
AAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC
410
KAAAGC | S ICTO
A SCAC
SICCI
K AGI
SICCI
GAA
K AGA
K AGA
 | E
SAGC
SAGTO
SAGTO
SAGTO
KAGO
KAGO
KAGO
KAGO
KAGO
KAGO
KAGO
KAG | D SATI
A SCAL E SAGU
E SAGU
E SAAU
E SAAU
E SAAU
E SAAU
E SAAU
E SAAU
E SAAU | E GAAG | A
SCT
P
CCT.
E
SAGI
G
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
SGT
SGT
SGT
SGT
SGT
SGT
SGT | RECTORES
 | A GCC.
Q CAGI
T ACCI
D GACI
A GCC.
A GCC. | K AAGO
P CCTC
K AAGO
S S AGC/
V STTC
E CCC
D SACC
D SACC
S S CCCC
D SACC
S S CCCC
D SACC
S S CCCC
C D S S CCCCC
S S CCCCC
S S CCCCC
S S CCCCCCC
S S CCCCCCCC | p υ 2CASTC 420 A g 2CCGGCC 450 A g CCGGCC 450 K S S <td>12222
1312
1402
1492
1582
1672
1982
1982
2032
2122
2392
2392</td> | 12222
1312
1402
1492
1582
1672
1982
1982
2032
2122
2392
2392 |
| K P A
AAACCAC
S A T
AATGCCAC
S G Q
AGTGGCCAC
V T T
GTGACAAC
D S E
GACTCAGAC
D S E
GACTCAGAC
CCCACCCC
S X A
AGCAAAAGC
E D
GACCACCC
E B D
GACCACCC
CCCACCCC
E R A
GAGCGAGGCC
S I S S
CCATCCCCC
C R R V
CCAAGGCTAC
E R A
GAGCGAGCC
S I S S
CCCATCCCC
TTTTTTACAC | P
TCCAH
K
CAAGJ
K
SAAAC
P
CCCTJ
SAGT
SAGT
V
TGTGJ
T
TACTC
T
CAAG
N
CAAG
N
CAAGJ
N
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGG
CAAGJ
CAAGG
CAAGJ
CAAGG
CAAGJ
CAAGG
CAAGJ
CAAGG
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAGJ
CAAG | V
STG,
STG,
STG,
P
CCTI
KAGG
SCCI
E
SAGC
SCCI
E
SAGC
SCCI
E
SAGC
SCCI
E
SAGC
SCCI
E
SAGC
SCCI
E
SAGC
SCCI
E
SAGC
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI
E
SCCI | К
МАДДА
Р
СССТ
О
САДДА
Х
А
СССТ
О
СССТ
О
СССТ
О
СССТ
О
САДДА
Х
А
А
А
А
А
А
А
А
А
А
А
А
А | KAAJ
LAAJ
SGGCJ
RGGC
SGCJ
RGGC
SGCJ
RGGC
RGGC
RGGC
RGGC
RGGC
RGGC
RGGC
R
 | KAGCI
SAGCI
RAGAJ
E ESAGC
AGCI
AGCI
KAGI
SCCC
AGCI
KAGI
SCCC
AGCI
AGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
CAGCI
C | A GO K KAG TAC E GA A CO A CA KAG FIT |
E
GAG
4000
A
GCCA
430
A
GCCA
400
A
GCCA
400
A
GCCA
400
A
GCCA
400
A
GCCA
400
A
GCCA
400
A
CGCC
400
A
CGCC
400
A
CGCCA
400
A
CGCCA
400
A
CGCCA
400
A
CGCCA
400
A
CGCCA
400
CGAG
5200
CGAG
5200
CGAG
5200
CGAG
5200
CGAG
5200
CGAG
5200
CGAG
5200
CGAG
530
CGAG
530
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
500
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
CGAG
C | S
AGC
A
GCT
D
GAC
AAA
A
A
GCC
CAG
GAC
CAG
CAG
CAG
CAG | S
TCT
V
GTC
S
AGC
A
GCC
A
GCC
A
AGC
A
AGC
C
GGG
C
C
GGG
C
C
GGG
C
C
C
GGC
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | STCA
TCA
T ACT
SAGC
T ACG
T ACG
SCA
SCA
SCA
SCA
SCA
SCA
SCA
SCA
SCA
SCA | D
GACI
P
CCCTI
S
TCCCI
P
CCCCI
P
CCCCI
P
CCCCI
S
CCCCI
S
CCCCI
C
CCCI
S
CCCCI
C
CCCI
S
CCCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
CCCI
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C
C | S
AGC
²
K
AGC
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
C
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
CCTJ
S
C
C
S
C
C
S
C
S
C
S
S
C
S
S
C
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S
S | S
TCAC
P
CCGC
E
SAGC
L
CTAC
S
S
CAG
C
CAG
C
CAG
C
CAG
C
C
C
C
C
C
C
C
C
C
C
C
C |
D
DATT
P
CCTG
E
SAGG
P
CCTG
S
CCTG
S
CCCT
S
CCCT
S
CCCT
S
CCCT
S
CCCT
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCCTG
S
CCCCTG
S
CCCTG
S
CCCTG
S
CCCTG
S
CCCCTG
S
CCCTG
S
CCCT | S CTC CTC CTC CTC CTC CTC CTC CTC CTC CT | A A A A A A A A A A A A A A A A A A A |
380
AGT:
410
K
AAGC
410
K
AGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
500
CAGC
CAGC
CAGC
CAGC
CAGC
CAGC
CAGC
CA | STOCA
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
SCAC
S | E
SAGC
V
STGC
S
S
GTC
K
S
GTC
K
S
GTC
K
S
GTC
K
S
GAG
G
A
G
G
A
G
G
A
G
G
T
G
T
G
T
G
T
G | D JATI A GOI | E GAAG
T T E GAAG
E GAAG
B G C A B C A C C C C C C C C C C C C C C C
 | A
SCT
P
CCT,
E
SAGI
G
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
C
SGT
SGT
SGT
SGT
SGT
SGT
SGT
SGT | RECTORES | A GCC.
Q CAGI
TACCI
D GACI
A GCC.
A GCC. | K
AAGO
P
CCTC
X
AGO
STTC
STTC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC
SAGC
TGG | F υ CASTC 420 A G 3000 GCGGCC 450 4800 SGCTCC 510 510 F F 510 F K 510 F K 510 F K 510 G S 510 G S 510 F K 510 F K 510 G S 510 G G 600 G G G G G G G G G G G G G G G G G G G G G G G G G G
 | 12222
1312
1402
1492
1592
1672
1952
1952
2032
2122
2392
2392
2392 |

TACAGATITITIGTIGTIGTGTGTGTGTGTGTGTGTGTGTGTG	2514
TTATGTGAGTGCGCAGACCCACATTTGATCCCACCCCTCAAAGCCCCGGTGGGCCCTGACATAAGTCTTGTGATGTTTGACTGCCTAAGCA	2662
TGCCCTGTGCTCATCTTCATCCATTGGGCCTGACACCGAAGCTTCCCGAAGCCGGGCGTGGATCTGCCAACTTTGGGGATAAAATTGCAGT	2752
TCTTGGTACAATTTCCTACTGAACTGACAGGCAGGATTCTCGATGTGAGTGCGATGCAACGGTTTTTGTTTTGTTCTCAGTAGCTATTAG	2842
TGCTACGTGTTTACAGTGTGTTCTAGTTTTAATTTCGAAGTAAGCTTTTCTGACACTGAGAGGCATTTGCAACAACTTGACTCTTACCGC	2932
TGTTGTATATAAGCTCATGAATATATTTGATTCTTGTTAACATCATCAAGAGCAGAATGGTAAACTCCTGCATGGTTGAGGCACTGGTCA	3022
AGGAAGTGGAGATGACCTACCTGAGCCTCTGGGTGAAGTAGGTACGGTGGATAGATCCTGGGCACCTGCAGTGACGAGCAGGCGAAGGAA	3112
AGTGAGGCTGGGAGAGGGTCTGGGCAGGACCGTTCTGTCTG	3202
TTGGTCCCCTGGATTGATTATTGGTACCTTAACTATATGATGTTACTGATACAGGCTAACCAGGGGGGGG	3292
TAGTGGCGCTTACCAACCATTCAAGGACAGAGTGTGATCTCCATCAAGGCAGGAAGGA	3382
ATTATAAAGCCGTAAGATTTGACCATTGGCAGAGGCTCAGCCAGAGTGAGGGAAGGGGAAGAGCACCCCTGGCTAACCCGTGAGAACAGA	3472
CACGGAGCCTCCCTGGTTTGGTTTCCATGGTCACCTGGTAACCTGCTAAAAGTGGGTGCCTGTGGGCAGCTCCTTGAGGAAGTCTGCATGG	3562
тсаладттететететтлетлеллалфаталматератестетеллаллаллалалала. 3624	

В ртм6

93236	
ACCCCCAGCTGGGACCCACTGTGACAAGAACTATGGGAGAACGAGGCTGGGGGAATTCCTTAGTGACGCGTTTAACCCGGAGT 804 .	•
D P Q L G P T V T R T H G E R G W R N S V V T R L T R S	
-17	
GACCTGTCTGCCCAGCTGGAGGTTGAGACCCAATCTGGAGCTGGGCATAGGGTGAGGTATTCCCAGCAGACAAACAGCTCATGTCCCCTG	720
M S P	
-30	
GCCGAGATCGCTGCAGGATGTCTCGGGGGTCGGAGGTGAGTGTGTGT	630
AAGGATTCCACCCCGGTCCCCATCAGACATCCCTCAGCAGAAGTACCCTCCCACTTTCCTTCAGAACTGAATTCTGCGTCCCTCCGGAGA	540
AGCATGCTGATCAAAGCACGGGGCAAAGGGGACAGCACCAGCCTCGGCTCTAAGCTCTGGCTCAGGGTGGTTCTGCCTACAACCTTCCCAG	450
ACATGGCCACTTGCACCCACGAGGACAATGCAGACTCATGAGCAGCCTGGTCTCCTCTCTCCCCAAGGAAGCCAGGCTGAAGACAAGCCA	360
ACAGGAGGCCCAGCTTCCAGTTTATTCCAGTTGGGTAGAGGAGGCCTCAGCTAGAGCACACTTCCAGGGTGAGGCTTGGGCTG	270
AT5CTCCC6T6CCTGCTCTCCAGCCCCGAGGGGCTGAACTCTACATCCACAGCTGAGTTCTCTGGATCTGGGGCACTAACCCCCAACAGCA	180
ATAGCTCGCTTTTAATAACATACAAAATCTAGGGAAAACCCCAAGTAGTGCCCACTCCCAGAGACCAAGGA <u>ATG</u> GAACCAG <u>ATG</u> GGAGCC	90
1	

Figure 1. Nucleotide Sequence of Nopp140

(A) DNA sequence of the coding strand of pTM17 and predicted amino acid sequence in single letter code above. The nucleotide sequence between the L and the second black dot denotes the identical 3' sequence of pTM6, except for the shaded CAG triplet encoding glutamine 150, which is missing in pTM6, as confirmed by sequencing of several independent clones from two different cDNA libraries. The two black dots enclose the sequence not shown for pTM6 in (B). The shaded boxes indicate the potential polyadenylation signals for pTM6 (first) and pTM17 (second), and the open box points out the divergent 5' end of pTM17. The underlined amino acid residues correspond to the sequenced peptides, and the amino acids boxed by a broken line point out minimal NLS consensus sequences. The circled P above serine 567 indicates that this residue was phosphorylated, as confirmed by amino acid sequencing (see Results). tio of the two mRNAs varied not only between rat and human (Figure 2B) but also between different tissues of the same species (data not shown). Moreover, the coding sequence seemed to be better conserved between rat and human than the 3'- and 5'-UTRs, which did not crosshybridize under the applied conditions (Figure 2B, probes I and III). Interestingly, the processing and polyadenylation of the 3' end of the two Nopp140 mRNAs appeared to be related to their distinct 5'-UTRs, because a long 5' end always coincided with a short 3' end in pTM6 and vice versa in pTM17. This is further underlined by the detection of only two mRNAs on Northern blots, unlike the comparable case of the int-2 oncogene, which expressed all four possible classes of mRNA (Mansour and Martin, 1988).

Southern blot analysis was performed to examine whether the two mRNAs originated from one or two genes. Hybridization of common probe II to rat genomic DNA resulted in single bands in the case of EcoRI- and BamHIdigested DNA, two for HindIII, and several for PstI and BgIII (Figure 2C). This pattern corresponded to the restriction map of probe II and suggested, therefore, that the two Nopp140 mRNAs were transcripts from a single gene that were generated by alternative usage of polyadenylation sites, by differential splicing, and/or by use of two alternate transcription start sites.

In the remainder of this article, only Nopp140 encoded by pTM17 shall be discussed, both because the proteins encoded by pTM17 and pTM6 are essentially identical and because of the following reason. mRNAs with unusually long 5'-UTRs containing many ATGs, like pTM6 (Figure 1B), may not be efficiently translated, whereas even minor transcripts with a short 5'-UTR, like pTM17 (Figure 1A), may represent the functional mRNA (Kozak, 1991 and references therein).

All sequences obtained from peptides of purified Nopp-140 were found in the cDNA-deduced amino acid sequence (Figure 1A, underlined residues). The open reading frame of pTM17 encoded a protein of 704 aa residues with a calculated molecular mass of 73.6 kd, nearly half of the apparent molecular mass of 140 kd estimated by SDSpolyacrylamide gel electrophoresis (PAGE; Meier and Blobel, 1990). To determine whether Nopp140 cDNA encoded a protein with the expected relative mobility on SDS-PAGE, pTM17 cDNA was used to program an in vitro transcription/translation reaction. Indeed, a protein of apparent molecular weight 140 kd was produced that comigrated with the authentic Nopp140 on SDS-PAGE (Figure 2D).

⁽B) DNA sequence of the coding strand of pTM6 and predicted amino acid sequence in single letter code above. Only the 5' end of pTM6 is actually shown, and the sequence that is overlapping with pTM17 is indicated between the two black dots (see A). The sequence upstream of the L represents the divergent 5' end of pTM6. Underlined are 12 potential translation initiation sites of which, however, only the two most 3' are in-frame with the long open reading frame without encountering a stop codon. These ATGs would extend the predicted amino acid sequence by 17 or 30 residues, as indicated by the negative numbers above the methionines.



Figure 2. Analysis of the Two Species of Nopp140 Clones

(A) Schematic representation of pTM17 and pTM6. Open boxes identify the open reading frames, and the shaded area between the two clones indicates identity, aside from a single amino acid that is absent in pTM6 (open box interrupted by shaded stripe). The small shaded boxes labeled I, II, and III specify the probes used for Northern and Southern blotting.

(B) Northern blots of rat and human $poly(A)^+$ RNA probed with Nopp140 cDNA-specific probes. The same blot was hybridized successively with the probes indicated above, demonstrating the identity of the 3.6 and 3.2 kb mRNAs with pTM17 and pTM6, respectively. The apparent size of marker RNAs is given on the left, and the estimated size of the two Nopp140 RNA species is indicated on the right.

(C) Southern blot of rat genomic DNA hybridized with a probe common to both species of Nopp140 clones. The geno blot containing rat genomic DNA digested with the five indicated restriction enzymes was hybridized with probe II (see [A]). Note that the very high molecular weight band in the EcoRI-digested lane probably stemmed from hybridization to undigested DNA.

(D) Autoradiogram of in vitro translated Nopp140 analyzed by SDS-PAGE. Lane 2: pTM17 DNA was in vitro transcribed and the RNA translated in reticulocyte lysate in the presence of [³⁵S]methionine. Lane 1 shows the products of in vitro translation in the absence of exogenous RNA. The bar indicates the position of the migration of authentic Nopp140. Note that Nopp140 contains only a single internal methionine, which explains why background labeling is relatively high.

Amino Acid Sequence Analysis

Even a superficial inspection of the primary sequence of Nopp140 discloses a preference for only a limited number of amino acids and their alignment in clusters. In fact, two-thirds of the Nopp140 sequence consists of only 5 aa, namely, serine (17.2%), lysine (16.1%), alanine (14.3%), glutamic acid (9.7%), and proline (9.2%). On the other hand, the sequence contains only a single internal methionine and no cysteine residue. Most of the serine residues are clustered in ten exclusively acidic repeats that are separated by basic amino acid stretches comprising the abundant lysine, alanine, and proline residues (Figure 3A). The acidic serine clusters are remarkably conserved among themselves, as highlighted by the underlined bold characters in Figure 3, which are identical in all ten repeats and in an eleventh one that only contains a single serine. Dot plot analysis of the protein sequence revealed these novel repeats more dramatically by yielding a dot for every 6 identical aa residues out of 10 (Figure 3B).

Interestingly, the identically conserved amino acids of the acidic serine repeats constitute a consensus site for casein kinase II (CK II) phosphorylation (Marin et al., 1986; Kuenzel et al., 1987). In fact, a computer search for protein patterns in the primary sequence of Nopp140 revealed that 72 out of 704 residues represent potential phosphorylation sites. Forty-five of the 82 serine residues in the acidic repeats form consensus sites for CK II phosphorylation. Ultimately, once the 45 residues have been phosphorylated, the remainder become acceptor sites for CK II phosphorylation (Meggio and Pinna, 1988; see Discussion). Thus, the serine clusters are potentially converted into stretches of 13-17 continuous negatively charged residues. Aminoterminal sequencing of a Nopp140 peptide yielded a dehydro serine in position 567 (Figure 1A), which strongly suggested that at least one of the identically repeated consensus sites is phosphorylated (Figure 3A, serine 567). This confirmed that CK II type phosphorylation indeed occurs in cellular Nopp140. Besides CK II sites, protein pattern analysis revealed 19 protein kinase C consensus sites, most of which also form CDC2/histone H1 kinase phosphorylation sites and lie within the basic part of the ten repeats.

As Nopp140 exhibits a distinct nucleolar localization, the primary sequence was scanned for the presence of targeting signals. Several nucleolar proteins have been shown to contain specific targeting signals that comprise a series of arginine residues and function independently from NLSs (Siomi et al., 1988; Dang and Lee, 1989; Kubota et al., 1989; Cochrane et al., 1990). The predicted amino acid sequence of Nopp140 does not contain any sequence resembling such nucleolar targeting signals. However, seven sequences corresponding to the 4 aa motif Lys-Arg/ Lvs-X-Arg/Lys could be identified (Figure 1A, boxed with broken lines). This sequence represents a minimal NLS consensus sequence that is shared by most NLSs and is homologous to the SV40 T antigen sequence (Chelsky et al., 1989). It remains to be shown whether any of these sequences are functional.



Figure 3. The Nopp140 Motif

(A) Alignment of the 10-fold repeated negative-positive motif of the predicted amino acid sequence in single letter code. The Nopp140 sequence was aligned by hand to obtain an optimal match between the acidic serine repeats. The box labeled "acidic" contains not a single positively charged residue, while the "basic" box does not contain any acidic amino acid. Underlined residues in bold point out the identically conserved amino acids in all ten acidic serine clusters as well as a flanking sequence. The serine in these identically conserved residues constitutes a CK II consensus phophorylation site, one of which was confirmed to be phosphorylated in vivo by amino acid sequencing (circled P; see Results).
(B) Dot plot of Nopp140 protein sequence. The amino acid sequence of Nopp140 was compared with itself, yielding a dot for every six identical residues out of ten. Note that the residue numbers of the repeats match the ones of the repeats in (A) exactly.

Phosphorylation

To determine whether the abundance of potential phosphorylation sites reflected the situation in vivo, buffalo rat liver (BRL) cells were metabolically labeled with ³²P orthophosphate. The nuclei were isolated and solubilized with SDS, and Nopp140 was immunoprecipitated with anti-Nopp140 peptide antibodies (Figure 4A). Nopp140 constituted the most heavily phosphorylated protein in this frac-



Figure 4. Nopp140 Is Massively Phosphorylated

(A) Immunoprecipitation of in vivo ³²P-labeled Nopp140. BRL cells were labeled with ³²P for 16 h²r, nuclei were prepared, and Nopp140 was immunoprecipitated in the presence (lane 1 and 2) and absence (lane 3 and 4) of free competing Nopp140 peptide, and the precipitates (P) and one-fifth of the supernatants (S) analyzed by SDS-PAGE and autoradiography.

(B) Phosphatase treatment of in vitro translated Nopp140 as analyzed by SDS-PAGE and autoradiography. Lane 1: pTM17 DNA was in vitro transcribed and translated in reticulocyte lysate for 45 min at 25°C. Lane 2: same as lane 1, but translation was arrested by 2.5-fold dilution into phosphatase buffer, and incubation was continued for 1 hr at 37°C. Lane 3: same as lane 2 but alkaline phosphatase was added during 37°C incubation. Lane 4: same as lane 3, but phosphatase inhibitors (see Experimental Procedures) were added during phosphatase incubation. tion (lane 1) and was not precipitated in the presence of competing peptide (lane 2). In the absence of competing peptide Nopp140 was efficiently removed from the SDS extract (lane 3) and immunoprecipitated (lane 4). Similar results were obtained when ³²P-labeled proteins of whole cells were analyzed (not shown). Considering that Nopp140 is not a major cellular protein, e.g., it cannot be distinguished among Coomassie blue-stained nuclear proteins when separated by SDS-PAGE, Nopp140 is one of the most phosphorylated proteins in the cell.

We investigated whether the massive phosphorylation of Nopp140 was responsible for its aberrant migration on SDS-PAGE. As described above (Figure 2D), in vitro transcription/translation of Nopp140 cDNA in reticulocyte lysate generated a band on SDS-PAGE of Mr 140 kd (Figure 4B, lane 1). In addition to Nopp140, we observed a faster migrating band of 100 kd (Figure 4B, lane 1). Upon treatment with alkaline phosphatase, Nopp140 was converted to the 100 kd protein (Figure 4B, Iane 3), demonstrating that the aberrantly large Mr of Nopp140 was at least in part due to phosphorylation. Identical treatment of ³²P-labeled Nopp140 resulted in removal of all ³²P label (data not shown). Since the protein moiety of Nopp140 remains intact during phosphatase treatment (see Figure 4B, lane 3), the 100 kd protein most likely represents completely dephosphorylated Nopp140. Conversely, further incubation of the translation products in reticulocyte lysate after translation arrest by dilution with buffer resulted in conversion of the 100 kd protein into Nopp140 (Figure 4B, compare lanes 1 and 2). Therefore, reticulocyte lysate contained the kinase(s) required to generate fully phosphorylated Nopp140. The persisting discrepancy between the calculated (73.6 kd) and relative molecular mass (100 kd) of dephospho Nopp140 is probably caused by the fact that one-third of all amino acids are strongly basic or acidic.



Figure 5. Phosphorylation Controls NLS Binding to Nopp140

(A) Autoradiogram of a Western blot of purified Nopp140 tested with anti-Nopp140 antibodies and developed with ¹²⁵I-labeled protein A. Nopp140 purified from rat liver nuclei was treated with alkaline phosphatase in the absence (lane 1) and presence (lane 2) of phosphatase inhibitors and analyzed by SDS-PAGE and Western blotting.

(B) Binding of NLS peptide conjugates to phosphorylated (lane 2) and dephosphorylated (lane 1) Nopp140 on Western blots. Purified Nopp140 was treated with phosphatase and transferred to nitrocellulose as in (A). Nitrocellulose strips were cut in half, incubated in parallel with wild-type (w) and mutant (m) synthetic NLS peptides coupled to human serum albumin, and the peptide conjugates visualized by autoradiography after incubation with anti-albumin antibodies and ¹²⁵I-labeled protein A.

Phosphorylation Regulates NLS Binding

The identification of Nopp140 as an NLS binding protein (Meier and Blobel, 1990), combined with the high density of negative charges generated by its massive phosphorylation, prompted us to test whether the phosphorylation might be responsible for its interaction with positively charged NLS peptides. To address this question, we employed the NLS ligand blot assay described previously (Meier and Blobel, 1990). Nopp140 purified from rat liver nuclei was treated with phosphatase, separated on SDS-PAGE, transferred to nitrocellulose filters, and the phospho and dephospho forms were detected with anti-Nopp140 antibodies (Figure 5A). Phosphatase treatment of purified Nopp140 resulted in the same mobility shift as described for in vitro translated Nopp140 (Figure 4B). The proteins were then also probed with wild-type and mutant NLS peptide conjugates that differ by a single amino acid. As shown in Figure 5B, phosphorylated Nopp140, but not dephospho Nopp140, bound the wild-type NLS peptide conjugate, demonstrating that binding of the NLS to Nopp140 is dependent on phosphorylation.

Nucleocytoplasmic Shuttling and Anti-Nopp140 Antibodies

We have previously suggested that Nopp140, being an NLS-binding protein and a primarily nucleolar protein, might be involved in nucleocytoplasmic transport by shuttling between the nucleolus and the cytoplasm. To test this hypothesis directly, we followed a protocol that took advantage of the fact that immunoglobulin G (IgG) molecules are too large to enter the nucleus simply by diffusion when injected into the cytoplasm of tissue culture cells



Figure 6. Characterization of Anti-Nopp140 Antibodies

(A) Western blot of whole rat liver nuclei tested with anti-Nopp140 antibodies. Nuclear proteins were separated by SDS–PAGE and transferred to nitrocellulose filters. Anti-Nopp140 antibodies were detected with ¹²⁶I-labeled protein A and autoradiography. Note that besides the single band recognized on whole nuclei, anti-Nopp140 antibodies also reacted with a very minor band below Nopp140 that comigrated with the dephospho form of Nopp140.

(B) Indirect immunofluorescence on fixed and permeabilized BRL cells with anti-Nopp140 antibodies. Cells were fixed with 2% paraformaldehyde and permeabilized with 1% Triton X-100. Unfractionated antiserum and fluorescein-labeled secondary antibodies were used to detect Nopp140. Aside from the bright nucleolar pattern, a faint staining was observed throughout the cell.

(C) Competition of indirect immunofluorescence by free Nopp140 peptide. Same as (B), but 50 μ M free Nopp140 peptide was included during the first antiserum incubation. Note that only the nuclear fluorescence was competed for, not the cytoplasmic. Bar = 10 μ m.

(D) Indirect immunofluorescence with affinity-purified anti-Nopp140 antibodies. Cells were treated as in (B) but incubated with affinity-purified antibodies instead of unfractionated serum. Note that next to the bright nucleolar staining there is also a faint signal detectable in the nucleoplasm.

(Madsen et al., 1986; Borer et al., 1989). However, antibodies to a shuttling nuclear protein may bind to the protein while it is in the cytoplasm and be piggybacked into the nucleus.

Toward this end we raised antibodies against a synthetic peptide derived from the Nopp140 amino acid sequence (Figure 1A, residues 292–309). Anti-Nopp140 antibodies recognized a major polypeptide on Western blots of whole rat liver nuclei (Figure 6A) and also reacted with purified Nopp140, the phospho as well as the dephospho form (Figure 5A, lanes 2 and 1, respectively). In immunoprecipitation experiments, the antibodies precipitated a single band, which corresponded to Nopp140, as confirmed by competition of the precipitation with free synthetic peptide (Figure 4A). On fixed and permeabilized cells anti-Nopp140 antibodies produced a strong punctate nucleolar pattern and a weak nucleoplasmic stain (Figure 6B). Co-incubation of anti-Nopp140 antibodies with free pep-



Figure 7. Nopp140 Shuttles between Nucleus and Cytoplasm As Demonstrated by Nucleolar Accumulation of Cytoplasmically Injected Anti-Nopp140 Antibodies

(A) Direct double fluorescence of rhodaminelabeled anti-Nopp140 antibodies (panel 1) and fluorescein-labeled control antibodies against RNA polymerase I (panel 2). BRL cells were fixed and permeabilized as described in Figure legend 6B and incubated concomitantly with both labeled antibodies.

(B) Cytoplasmic coinjection of both labeled antibodies into living cells and incubation for 12 hr at 37°C. BRL cells were microinjected into the cytoplasm with both antibodies concomitantly. After 12 hr incubation the cells were fixed with 4% formaldehyde, and the direct fluorescence was observed. Panels as in (A).

(C) Anti-Nopp140 antibodies were imported into the nucleolus in the absence of de novo protein synthesis. Essentially the same as (B), except that the cells were incubated with $10 \,\mu g/$ ml cycloheximide 1 hr prior to injection, and the concentration was maintained throughout the experiment. After the indicated periods of incubation, the cells were fixed, and the fluorescence was observed (upper panels: anti-Nopp140 antibodies; lower panels: anti-RNA polymerase I antibodies). Bars = 10 μm .

tide resulted in the complete loss of both nucleolar and nucleoplasmic fluorescence (Figure 6C). Finally, indirect immunofluorescence with affinity-purified anti-Nopp140 antibodies produced strong nucleolar and weak nucleoplasmic labeling (Figure 6D).

To test for nucleocytoplasmic shuttling of Nopp140, affinity-purified anti-Nopp140 antibodies were chemically labeled with the fluorescent dye rhodamine, while control antibodies against another nucleolar protein, RNA polymerase I, were labeled with fluorescein. Examination of permeabilized cells by fluorescence microscopy showed that both labeled antibodies retained their reactivity and produced their characteristic staining of the nucleolus (Figure 7A; see Scheer and Rose, 1984 and Meier and Blobel, 1990). However, microinjection of the two labeled antibodies intò the cytoplasm of living cells, incubation for 12 hr at 37°C, and subsequent examination by fluorescence microscopy (Figure 7B) clearly demonstrated that anti-Nopp140 antibodies had accumulated in the nucleolus (panel 1) whereas anti-RNA polymerase I antibodies remained excluded from the nucleus (panel 2). These data suggested that anti-Nopp140 antibodies were indeed carried into the nucleus, presumably by binding to cytoplasmic Nopp140. Interestingly, the imported antibodies bind to an epitope situated in the center of the Nopp140 repeats (Figure 3A, residues 292-309) where NLS binding is proposed to occur (see Discussion). The nucleolar accumulation of anti-Nopp140 antibodies was not caused by newly synthesized Nopp140, because it occurred even when de novo protein synthesis was inhibited by cycloheximide (Figure 7C). Moreover, nucleolar accumulation of anti-Nopp140 antibodies occurred rapidly, as early as 10 min after cytoplasmic injection (Figure 7C).

Immunoelectron Microscopy

Incubation of thin sections of L. R. White-embedded BRL cells with affinity-purified anti-Nopp140 antibodies and subsequent detection by immuno gold revealed strong labeling of the dense fibrillar component of the nucleolus (Figure 8A). An identical nucleolar labeling pattern (not shown) was obtained by parallel staining with antifibrillarin antibodies (Aris and Blobel, 1988). However, unlike the anti-fibrillarin antibody labeling, which remained restricted to the dense fibrillar component of the nucleolus (not shown), staining with anti-Nopp140 antibodies also occurred in the nucleoplasm (Figures 8A and 8B), consistent with the nucleoplasmic staining observed in immunofluorescence (see Figure 6D). Most strikingly, while nucleoplasmic staining by anti-Nopp140 antibodies was detected in all thin sections, a few sections revealed staining along curvilinear tracks across the nucleoplasm (Figure 8). Depending on the plane of the section, these tracks could be followed, in rare cases, for almost their entire length (see Figure 8B) or, more frequently, for short distances. The longest track observed extended for several microns from the dense fibrillar component of the nucleolus through the nucleoplasm close to the nuclear envelope (arrowheads in upper portion of Figure 8B). Shorter tracks, either emanating from the nucleolus (arrowheads in Figure 8A) or apparently unconnected to it (arrowheads in lower

portion of Figure 8B) could be seen more frequently. In some sections tracks appeared to cross the nuclear pore complex (see Figure 8B insert). In cross sections, tracks could be represented by the randomly scattered clusters of gold particles in the nucleoplasm (Figures 8A and 8B). Similar results were also obtained by staining cryosections of BRL cells with anti-Nopp140 antibodies (not shown). These data suggest that Nopp140 is located on a limited number of tracks that extend from the dense fibrillar component of the nucleolus across the nucleoplasm to a limited number of nuclear pore complexes.

Discussion

Shuttling and Tracks

Our results demonstrate that Nopp140 is a member of a family of proteins that shuttle between nucleus and cytoplasm. As demonstrated by immunoelectron microscopy, shuttling of Nopp140 may proceed along a limited number of curvilinear tracks that extend from the dense fibrillar component of the nucleolus across the nucleoplasm to certain nuclear pore complexes.

Although the precise function of Nopp140 remains to be elucidated, its shuttling between the nucleolus and the cytoplasm suggests that it functions in transport, either in import (e.g., ribosomal proteins) or in export (e.g., preribosomal subunits). In either process Nopp140 might serve as a chaperone. Being a protein with alternating acidic and basic domains, it could function to cover and neutralize highly charged domains of preribosomal particles (export) or of ribosomal proteins (import). This might facilitate navigating these components between the Scylla and Charybdis of the highly charged nuclear chromatin either out of or into the nucleolar ribosome assembly sites.

What is the nature of the tracks along which Nopp140 moves? Are these tracks filaments? If they are, could they be composed of actin, whose presence in the nucleus has been reported (Jockusch et al., 1974; Clark and Merriam, 1977; Fukui, 1978; LeStourgeon, 1978)? If they are nuclear actin filaments, are there corresponding nuclear myosin motors? Would these motors specifically recognize and attach to Nopp140 to unidirectionally move Nopp140-associated structures along these actin filaments either into or out of the nucleolus? These are only some of the questions that are evoked by the visualization of these tracks and that await further investigation.

Our immunoelectron microscopic data suggest that there are only a limited number of tracks connecting nucleolar ribosome assembly sites with the cytoplasm through a limited number of pore complexes. In fact, the low frequency of tracks within the same plane as the thin sections presently does not allow statistical analysis to estimate the number of tracks per nucleus. However, a limited number of tracks might be directly demonstrable either by threedimensional reconstruction of optical sections obtained by laser scanning confocal microscopy of cells injected with fluorescein-labeled anti-Nopp140 antibodies and/or by electron microscopic analysis of serial sections of cells injected with gold-labeled anti-Nopp140 antibodies. Such experiments are currently in progress. The tracks might originate from the transcription sites of the 200 or so ribosomal RNA genes and the nearby assembly sites for preribosomal particles, such that import of supplies (ribosomal proteins) and export of products (preribosomal particles) might proceed only through those pore complexes to which these tracks are connected (Blobel, 1985).

It is interesting to note that tracks of mRNA molecules have been detected in the nucleoplasm by in situ hybridization, both on the light and electron microscopic level (Lawrence et al., 1989; Huang and Spector, 1991). Therefore most, if not all, macromolecular traffic in and out of the nucleus may proceed on tracks. Such tracks might originate or terminate in nuclear pore complexes and in transcription/ribonucleoprotein assembly sites.

Phosphorylation

Over 10% of all amino acid residues present in Nopp140 constitute potential phosphorylation sites, and the experimental results indicate that Nopp140 is indeed phosphorylated to an unusually high degree. This made us wonder whether Nopp140 may not have been noticed previously. In fact, a literature survey revealed that, solely on the basis of massive phosphorylation, several groups had identified proteins with relative molecular masses of about 140 kd, which could correspond to Nopp140 (Wehner et al., 1977; Pfeifle et al., 1981; Wang et al., 1981; Banville and Simard, 1982; Pfeifle and Anderer, 1984; Ahn et al., 1985). One of these proteins, pp135, was determined to incorporate 75 phosphate groups per molecule (Pfaff and Anderer, 1988), which coincides well with the theoretical number of 82 serine residues in the Nopp140 motif. None of the cDNAs for these proteins, however, have been cloned to date, and their relationship to Nopp140 remains to be proven.

Surprisingly, under all conditions studied so far, Nopp-140 occurred only in a completely phosphorylated or dephosphorylated state and not in forms of intermediate degree of phosphorylation. This is particularly obvious in Figure 4B (lane 4), where even in the presence of phosphatase inhibitors a minor fraction of Nopp140 becomes effectively converted to the dephospho form without detectable intermediates. We propose that phosphorylation occurs in a cooperative manner by the following mechanism. The consensus site for CK II phosphorylation consists of the serine or threonine acceptor site and an acidic amino acid 3 residues away on the carboxy-terminal side, as exemplified by the identically conserved amino acids in the Nopp140 repeats (Figure 3A, underlined residues in bold). Every additional acidic residue (Meggio and Pinna, 1988) near the target serine (threonine) improves the acceptor site (Marin et al., 1986; Kuenzel et al., 1987). Thus, phosphorylation of the conserved serine within a single Nopp140 repeat renders the preceding serine a better substrate for CK II. That serine will then become phosphorylated in turn, rendering the preceding serine a better substrate and so on. The whole process could take place at all ten repeats simultaneously, thereby explaining the "all or none" phenomenon of Nopp140 phosphorylation. Such a sequential addition of phosphate residues by CK II has





Cell 134 in fact been shown to occur for a peptide derived from SV40 T antigen (Marshak and Carroll, 1991).

NLS Binding

Nopp140 was identified as an NLS-binding protein by a ligand blot assay (Meier and Blobel, 1990). It remains to be determined whether the NLS binding is of physiological relevance. Our results here show that only the phospho form of Nopp140 binds NLS, not the dephospho form. Therefore, NLS binding may occur in the phosphorylated acidic repeats of Nopp140, and each of the ten negatively charged domains could constitute a binding site for a positively charged NLS. Such a multivalent binding site for NLSs would provide an explanation for the millimolar amounts of free NLS peptides required to compete for the binding of a substrate carrying multiple NLS sequences (Goldfarb et al., 1986; Meier and Blobel, 1990). The enhanced nucleocytoplasmic transport rate of nuclear proteins with multiple NLSs (Lanford et al., 1986; Roberts et al., 1987; Dworetzky et al., 1988) also could be explained by their increased affinity for such a multivalent receptor(s). The rapid and complete phosphorylation/dephosphorylation of Nopp140 may represent a means of regulating its affinity for NLS-containing proteins and thereby its ability to function in nucleocytoplasmic transport.

Aside from Nopp140, another nucleolar protein, No38, has been shown to bind to immobilized NLS peptides (Goldfarb, 1988). Interestingly, No38 also shuttles between nucleus and cytoplasm (Borer et al., 1989). In addition, an NLS-binding protein in yeast, which contains an extended negatively charged stretch of amino acids with multiple serine residues, also seems to be located in the nucleolus (Lee et al., 1991). However, it is not known whether this protein shuttles. It will be interesting to determine whether these proteins colocalize to the same tracks that are decorated by Nopp140. Ultimately, elucidation of the nature of the Nopp140 tracks should open up new avenues for understanding nucleocytoplasmic transport.

Experimental Procedures

Purification, Cleavage, and Amino Acid Sequencing of Nopp140

Nopp140 was purified from rat liver nuclei essentially as described (Meier and Blobel, 1990), with the slight modification that prior to elution of Nopp140 the hydroxyapatite column was washed with 0.5 M potassium phosphate buffer containing 1 M KCI. Subsequent elution with 1 M potassium phosphate buffer yielded essentially pure Nopp-140. After SDS-PAGE, Nopp140 was cleaved while still in gel slices by cyanogen bromide, and the fragments were resolved by 12% SDS-PAGE (Laemmli, 1970) and transferred to polyvinylidine difluoride membrane (immobilon PVDF; Millipore Continental Water Systems, Bedford, MA) as described (Nikkodem and Fresco, 1979; Schnell et al., 1990). Alternatively, Nopp140 was electroeluted from gels and digested (Cleveland et al., 1977) by trypsin (40 µg/ml; Sigma Chemical Co., T-8642) for 2 hr in 10 mM Tris buffer (pH 8.1), 0.1% SDS, and the fragments were transferred to immobilon PVDF membrane and subjected to amino-terminal sequencing as described (Schnell et al., 1990).

Generation of a Nopp140-Specific Oligonucleotide Probe and Screening of cDNA Libraries

Standard techniques of molecular cloning were used as described (Maniatis et al., 1989) if not stated otherwise. Restriction enzymes were from Boehringer Mannheim Biochemicals (Indianapolis, IN) and New England BioLabs (Beverly, MA).

Two degenerate oligonucleotides (5'-GAGAATTCGCCGG[CG]A-C[ACGT]AA[AG]CC[ACGT]GG-3' and 5'-GAGTCGACGT[CT]TC[AG]-TT[AG]TG[CT]TT-3') corresponding to the ends of a tryptic peptide sequence of Nopp140 (Figure 1A, residues 581–601) were synthesized and used in a polymerase chain reaction under previously described conditions (Shelness and Blobel, 1990) with rat cDNA as template to generate a specific probe for Nopp140. The product was cloned into the EcoRI–Sall sites of pBluescript II (Stratagene, La Jolla, CA) and sequenced by the dideoxy method (Sanger et al., 1977).

Synthetic nucleotides containing 48 bp of the confirmed sequence (Figure 1A, bp 1792–1840) were end labeled with [³²P]ATP and used to screen a commercial lambda ZAP II rat cDNA library (Stratagene) following the supplier's protocol. Two positive clones out of 1×10^6 phages screened were directly in vivo excised in pBluescript (Stratagene) and sequenced. The longer one (1.9 kb) was labeled with [³²P]dCTP using a random primed labeling kit (Feinberg and Vogelstein, 1983). This probe served then to screen 300,000 phages of an unamplified λ gt10 library constructed from BRL cell RNA (generously provided by Jun Sukegawa). Sixty-five positive clones were identified and analyzed at least with regard to their length. This was achieved using vector- and gene-specific oligonucleotides in a polymerase chain reaction on phage mixtures as template that were obtained in the primary screen of the library.

cDNA Sequencing and Analysis

EcoRI and HindIII restriction fragments of the longest λ gt10 inserts representing the two Nopp140 mRNAs (see Figures 1A and 1B) were subcloned into pBluescript II for DNA sequencing. Both strands were sequenced at least once, employing synthetic nucleotides as primers in the dideoxy method (Sanger et al., 1977) as modified by Schuurman and Keulen (1991). Note that the final 13 nucleotides before the poly(A) tail of pTM17 (Figure 1A) were derived from sequencing of an overlapping independent clone. DNA analysis was performed with the DNASTAR sofware program (DNASTAR, Madison, WI) and homology searches in GenBank by FASTA (Pearson and Lipman, 1988).

Northern and Southern Blots

Poly(A)⁺ RNA was prepared from HeLa, BRL, and N1S1 (rat hepatoma) tissue culture cells using RNAgents and PolyATtract isolation systems (Promega, Madison, WI). RNA from various rat and human tissues was from Clontech Laboratories, Inc. (Palo Alto, CA). The RNA was electrophoresed in denaturing agarose gels (Lehrach et al., 1977) and transferred to nitrocellulose filters (Thomas, 1980). Fragments of the two clones were [³²P]dCTP labeled by the random primer method and used for hybridization (Maniatis et al., 1989). A commercial Southern blot (Southern, 1975; Clontech Laboratories, Inc.) containing rat geno-

Figure 8. Immunoelectron Microscopy with Affinity-Purified Anti-Nopp140 Antibodies

(A) Immunogold labeling on sections of L. R. White-embedded BRL cells. BRL cells were fixed with 2% paraformaldehyde, 0.05% glutaraldehyde and embedded in L. R. White for sectioning. The sections were incubated with affinity-purified anti-Nopp140 antibodies and subsequently with secondary antibodies coupled to 10 nm gold particles. Note the strong labeling of the dense fibrillar component of the nucleolus (No) and some scattered labeling in the nucleoplasm. Arrowheads mark a track extending from the nucleolus into the nucleoplasm. Bar = $0.5 \mu m$.

(B) Immunogold labeling of Nopp140 along tracks between the nucleolus (No) and the nuclear envelope (NE). The sample was prepared as in (A). Tracks decorated by gold particles are pointed out by arrowheads. A largely continous track is shown in the upper part, and a disrupted track(s) is shown in the lower part. Note labeling in the nucleoplasm (N) in the form of clusters or short tracks. The cytoplasm (C) shows very little, scattered labeling. Inset shows gold particles lining up at the nuclear envelope near a nuclear pore complex. Bars = $0.5 \mu m$.

mic DNA digested by five different restriction enzymes was probed analogously.

Cell-Free Transcription and Translation

pTM17 cDNA isolated in minipreps was in vitro transcribed using T7 RNA polymerase (Stratagene) after linearization with BamHI. The pTM17 RNA was then in vitro translated in the presence of [³⁶S]methionine in rabbit reticulocyte lysate (Promega) and the samples processed as described (Nicchitta et al., 1991), except that the gels were enhanced (Enlightning, DuPont New England Nuclear, Boston, MA) for autoradiography.

Phosphorylation

BRL cells were grown to subconfluency in 100 cm² petri dishes in modified essential medium (GIBCO Laboratories, Grand Island, NY). After incubation of the cells for 1 hr in minimal essential medium without phosphate, they were labeled by addition of 1 ml (1 mCi) of [32P]sodium dihydrogen phosphate (DuPont New England Nuclear, Wilmington, DE, 1 mCi/mmol) for 1, 4, or 16 hr at 37°C. The labeling pattern of proteins and nucleic acids, separated by SDS-PAGE and autoradiographed as described below, looked identical after all incubation times, except that the overall labeling was weaker after 1 hr compared with the other time points. The cells were rinsed with ice-cold phosphatebuffered saline, scraped into 10 ml of phosphate-buffered saline, pelleted at 1000 x g for 5 min, and the pellet was frozen in liquid nitrogen. The frozen cell pellet was thawed and resuspended in 0.5 ml of homooenization buffer (0.1 M sodium phosphate (pH 7.4), 40 mM NaF, 0.6 mM Na₃VO₄, 1 mM dithiothreitol, 2 mM phenylmethylsulfonyl fluoride, 2 µg/ml aprotinin, and 1 µg/ml each leupeptin, antipain, chymostatin, and pepstatin A), incubated for 15 min on ice, and homogenized by 20 strokes in a glass teflon homogenizer. The homogenate was fractionated into nuclei and cytosol by a 10 min centrifugation at 1000 × g and the nuclear pellet resuspended in 0.5 ml of homogenization buffer. Then 0.5 ml of buffer containing 0.1 M sodium phosphate (pH 7.4), 0.3 M NaCl, 10 mM MgCl₂, 5 µg/ml RNAase A, and 250 µg/ml DNAase I was added to the nuclei and cytosol. After incubation for 30 min on ice, 20 µl of 0.5 M EDTA and 40 µl of 10% SDS were added to the 1 ml fractions. After incubation for 5 min at 50°C and for 5 min in a sonicator waterbath, the samples were diluted with twice the volume of 2% Triton X-100 in 0.1 M sodium phosphate buffer (pH 7.4), 150 mM NaCl and used for immunoprecipitation.

Immunoprecipitation

Aliquots were either incubated directly with anti-Nopp140 antibodies and subsequently with protein A–Sepharose (Pharmacia Fine Chemicals, Piscataway, NJ) or with protein A–Sepharose to which anti-Nopp140 IgGs had previously been adsorbed. To demonstrate the specificity of Nopp140 immunoprecipitation, 0.5 mM free Nopp140 peptide was included in half the samples. After incubation for 2 hr at room temperature, the supernatant was removed for further analysis and the protein A–Sepharose beads were washed twice with 0.1 M Tris (pH 7.4), 1% Triton X-100, 0.2% SDS, 150 mM NaCI, twice with 0.1 M Tris (pH 7.4), 0.5% Tween 20, and once with 0.1 M Tris (pH 7.4) and were eluted with SDS–PAGE sample buffer. The eluates and one-fifth of the supernatants were analyzed by SDS–PAGE and autoradiography.

Phosphatase Treatment

Purified Nopp140 was treated with calf intestinal (20 U/ μ g Nopp140, Boehringer Mannheim Biochemicals) or bacterial alkaline phosphatase (3.5 U/ μ g Nopp140, Worthington Biochemical Corporation, Freehold, NJ) as described for nucleoplasmin (Cotten et al., 1986). ³²P-labeled Nopp140 was immunoprecipitated (see above) and analogously treated with phosphatase either directly on protein A–Sepharose beads or after elution with free Nopp140 peptide. In control samples a mixture of 15 mM Na₃MO₄, 0.3 mM Na₃VO₄, and 20 mM NaF was included as phosphatase inhibitor. The proteins were analyzed by SDS–PAGE, transferred to nitrocellulose, and probed with either anti-Nopp140 antibodies or NLS peptide conjugates as described (Meier and Blobel, 1990).

Reticulocyte lysate containing in vitro translated Nopp140 was diluted 2.5-fold with phosphatase buffer (0.1 M Tris [pH 8.8], 10 mM MgCl₂) and incubated in the presence or absence of phosphatase inhibitor mixture (see above) with calf intestinal alkaline phosphatase (0.48 U/ μ I, GIBCO Laboratories, Grand Island, NY) for 1 hr at 37°C. The samples were analyzed by SDS-PAGE and autoradiography.

Antibodies and Immunological Techniques

Polyclonal antibodies were raised in rabbits against a synthetic peptide of Nopp140 (Figure 1A, residues 292-309) coupled to keyhole limpet hemocyanine (Calbiochem Behring Co., La Jolla, CA) through an additional carboxy-terminal cysteine as described (Green et al., 1982; Meier and Blobel, 1990). The Nopp140 synthetic peptide was coupled to SulfoLink coupling gel (Pierce, Rockford, IL) according to the supplier. Anti-Nopp140 IgGs were then affinity purified by passing serum over the peptide resin, washing with 1 M NaCl, eluting with 0.1 M glycine (pH 2.5), and dialyzing against phosphate-buffered saline. Nopp140 IgGs were fluorescently labeled with TRITC or FITC (Sigma) as described for nucleoplasmin (Newmeyer et al., 1986). Immunofluorescence experiments were carried out on paraformaldehyde-fixed and Triton X-100-permeabilized BRL cells (Meier and Blobel, 1990) using the sera at a 1:100 dilution, the affinity-purified antibodies at 1.1 µg/ml, and competing free Nopp140 peptide at 50 µM concentration. Dilution of anti-Nopp140 antibodies was routinely 10-fold higher in incubations of Western blots (Meier and Blobel, 1990).

Shuttling Experiments

BRL cells were coinjected with TRITC-labeled anti-Nopp140 IgGs (~1.3 mg/ml) and FITC-labeled control IgGs against RNA polymerase I (~2.2 mg/ml; kind gift by K. M. Rose, University of Texas, Houston) as described for peptide conjugates (Meier and Blobel, 1990). Where specified, 10 µg/ml cycloheximide was added to the cell culture medium 1 hr prior to injection and the same concentration maintained during incubation after injection.

Immunoelectron Microscopy

After trypsin detachment and pelleting, BRL cells were fixed with 2% paraformaldehyde, 0.05% glutaraldehyde in 100 mM cacodylate (pH 7.4) for 30 min at 4°C. The cells were either prepared for cryosectioning (Tokuyasu, 1973) or embedded in L. R. White (Electron Microscopy Sciences, Fort Washington, PA) and sectioned. Sections were incubated on Formvar carbon coated nickel grids either with affinity-purified anti-Nopp140 antibodies at dilutions between 1 and 10 μ g/ml or with straight culture supernatant from hybridoma cells secreting anti-fibrillarin monoclonal antibodies D77 (Aris and Blobel, 1988; kindly provided by J. Aris). Secondary incubations were performed with antibodies to 0 nm gold particles (Amersham Life Sciences, Arlington Heights, IL). The grids were stained with uranyl acetate and viewed on a JOEL 100 CX electron microscope operated at 80 kV.

Acknowledgments

We are particularly indebted to Helen Shio of the Rockefeller University EM facility for performing the immunoelectron microscopy, to Jun Sukegawa for the generous gift of his unamplified cDNA library and rat cDNA, and to baboon trainer Danny Schnell for his advice in molecular biology. We thank the members of the Rockefeller University biopolymer facility for peptide sequencing and synthesis and for oligonucleotide synthesis; K. M. Rose for anti-RNA polymerase I antibodies; Einar Hallberg for poly(A)⁺ RNA; Giovanni Migliaccio and Debkumar Pain for advice on many occasions; and Susan Smith, Chris Nicchitta, and Mary Moore for critically reading the manuscript.

During initial stages of this work, U. T. Meier was supported by the Swiss National Foundation. This paper is dedicated to George E. Palade.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC Section 1734 solely to indicate this fact.

Received April 2, 1992; revised May 6, 1992.

References

Adam, S. A., and Gerace, L. (1991). Cytosolic proteins that specifically bind nuclear location signals are receptors for nuclear import. Cell 66, 837–847. Adam, S. A., Lobl, T. J., Mitchell, M. A., and Gerace, L. (1989). Identification of specific binding proteins for a nuclear location sequence. Nature 337, 276–279.

Ahn, Y. S., Choi, Y. C., Goldknopf, I. L., and Busch, H. (1985). Isolation and characterization of a 125-kilodalton rapidly labeled nucleolar phosphoprotein. Biochemistry 24, 7296–7302.

Aris, J., and Blobel, G. (1988). Identification and characterization of a yeast nucleolar protein that is similar to a rat liver nucleolar protein. J. Cell Biol. *107*, 17–31.

Bachmann, M., Pfeifer, K., Schröder, H. C., and Müller, W. E. G. (1989). The La antigen shuttles between the nucleus and the cytoplasm in CV-1 cells. Mol. Cell. Biochem. *85*, 103–114.

Banville, D., and Simard, R. (1982). Changes in phosphorylation on nucleolar proteins correlated with inhibition of ribosomal precursor RNA processing. Exp. Cell Res. *137*, 437–441.

Benditt, J. O., Meyer, C., Fasold, H., Barnard, F. C., and Riedel, N. (1989). Interaction of a nuclear location signal with isolated nuclear envelopes and identification of signal-binding proteins by photoaffinity labelling. Proc. Natl. Acad. Sci. USA *86*, 9327–9331.

Blobel, G. (1985). Gene gating: a hypothesis. Proc. Natl. Acad. Sci. USA 82, 8527-8529.

Borer, R. A., Lehner, C. F., Eppenberger, H. M., and Nigg, E. A. (1989). Major nucleolar proteins shuttle between nucleus and cytoplasm. Cell 56, 379–390.

Chelsky, D., Ralph, R., and Jonak, G. (1989). Sequence requirement for synthetic peptide mediated translocation to the nucleus. Mol. Cell Biol. 9, 2487–2492.

Clark, T. G., and Merriam, R. W. (1977). Diffusible and bound actin in nuclei of Xenopus laevis oocytes. Cell 12, 883-891.

Cleveland, D. W., Fischer, S. G., Kirschner, M. W., and Laemmli, U. K. (1977). Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. 252, 1102–1106.

Cochrane, A. W., Perkins, A., and Rosen, C. A. (1990). Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. J. Virol. *64*, 881–885.

Cotten, M., Sealy, L., and Chalkley, R. (1986). Massive phosphorylation distinguishes Xenopus laevis nucleoplasmin isolated from oocytes or unfertilized eggs. Biochemistry 25, 5063–5069.

Dang, C. V., and Lee, W. M. F. (1989). Nuclear and nucleolar targeting sequences of c-erb-A, c-myb, N-myc, p53, HSP70, and HIV tat proteins. J. Biol. Chem. *264*, 18019–18023.

Dworetzky, S. I., Lanford, R. E., and Feldherr, C. M. (1988). The effects of variations in the number and sequence of targeting signals on nuclear uptake. J. Cell Biol. *107*, 1279–1287.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabeling DNA endonuclease fragments to high specific activity. Anal. Biochem. *132*, 6–13.

Fukui, Y. (1978). Intranuclear actin bundles induced by dimethyl sulfoxide in interphase nucleus of Dictyostelium. J. Cell Biol. 86, 146–157.

Goldfarb, D. S. (1988). Karyophilic peptides: applications to the study of nuclear transport. Cell Biol. Int. Rep. 12, 809–832.

Goldfarb, D. S., Gariépy, J., Schoolnik, G., and Kornberg, R. D. (1986). Synthetic peptides as nuclear localization signals. Nature *322*, 641–644.

Goldstein, L., and Ko, C. (1981). Distribution of proteins between nucleus and cytoplasm of Amoeba proteus. J. Cell Biol. 88, 516-525.

Green, N., Alexander, H., Olson, A., Alexander, S., Shinnick, T. M., Sutcliffe, J. G., and Lerner, R. A. (1982). Immunogenic structure of the influenza virus hemagglutinin. Cell 28, 477–487.

Guiochon-Mantel, A., Lescop, P., Christin-Maitre, S., Loosfelt, H., Perrot-Applanat, M., and Milgrom, E. (1991). Nucleocytoplasmic shuttling of the progesterone receptor. EMBO J. *10*, 3851–3859.

Huang, S., and Spector, D. L. (1991). Nascent pre-mRNA transcripts are associated with nuclear regions enriched in splicing factors. Genes Dev. *5*, 2288–2302.

Hügle, B., Scheer, U., and Franke, W. W. (1985). Ribocharin: a nuclear

Mr 40,000 protein specific to precursor particles of the large ribosomal subunit. Cell 41, 615–627.

Imamoto-Sonobe, N., Matsuoka, Y., Semba, T., Okada, Y., Uchida, T., and Yoneda, Y. (1990). A protein recognized by antibodies to Asp-Asp-Glu-Asp shows specific binding activity to heterogeneous nuclear transport signals. J. Biol. Chem. 265, 16504–16508.

Jockusch, B. M., Becker, M., Hindennach, I., and Jockusch, H. (1974). Slime mold actin: homology to vertebrate actin and presence in the nucleus. Exp. Cell Res. 89, 241–246.

Kalderon, D., Richardson, W. D., Markham, A. F., and Smith, A. E. (1984a). Sequence requirements for nuclear location of simian virus-40 large-T antigen. Nature *311*, 33–38.

Kalderon, D., Roberts, B. L., Richardson, W. D., and Smith, A. E. (1984b). A short amino acid sequence able to specify nuclear location. Cell 39, 499–509.

Kozak, M. (1991). An analysis of vertebrate mRNA sequences: intimations of translational control. J. Cell Biol. 115, 887–903.

Kubota, S., Siomi, H., Satoh, T., Endo, S., Maki, M., and Hatanaka, M. (1989). Functional similarity of HIV-I rev and HTLV-I rex proteins: identification of a new nucleolar targeting signal in rev protein. Biochem. Biophys. Res. Commun. *162*, 963–970.

Kuenzel, E. A., Mulligan, J. A., Sommercorn, J., and Krebs, E. G. (1987). Substrate specificity determinants for casein kinase II as deduced from studies with synthetic petides. J. Biol. Chem. 262, 9136–9140.

Kumar, A., and Warner, J. R. (1972). Characterization of ribosomal precursor particles from HeLa cell nucleoli. J. Mol. Biol. 63, 233–246.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lanford, R. E., and Butel, J. S. (1984). Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. Cell 37, 801–813.

Lanford, R. E., Kanda, P., and Kennedy, R. C. (1986). Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. Cell *46*, 575–582.

Lawrence, J. B., Singer, R. H., and Marselle, L. M. (1989). Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. Cell 57, 493–502.

Lee, W.-C., and Mélèse, T. (1989). Identification and characterization of a nuclear localization sequence-binding protein in yeast. Proc. Natl. Acad. Sci. USA *86*, 8808–8812.

Lee, W.-C., Xue, Z., and Mélèse, T. (1991). The nsr1 gene encodes a protein that specifically binds nuclear localization sequences and has two RNA binding motifs. J. Cell Biol. *113*, 1–12.

Lehrach, H., Diamond, D., Wozney, J. M., and Boedtker, H. (1977). RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical reexaminaton. Biochemistry *16*, 4743– 4751.

LeStourgeon, W. M. (1978). The occurrence of contractile proteins in nuclei and their possible function. In The Cell Nucleus, VI, H. Busch, ed. (New York: Academic Press), pp. 305–326.

Li, R., and Thomas, J. O. (1989). Identification of a human protein that interacts with nuclear localization signals. J. Cell Biol. 109, 2623–2632.

Madsen, P., Nielsen, S., and Celis, J. E. (1986). Monoclonal antibody specific for human nuclear proteins IEF 8Z30 and 8Z31 accumulates in the nucleus a few hours after cytoplasmic microinjection of cells expressing these proteins. J. Cell Biol. *103*, 2083–2089.

Mandell, R. B., and Feldherr, C. M. (1990). Identification of two HSP70related Xenopus cocyte proteins that are capable of recycling across the nuclear envelope. J. Cell Biol. *111*, 1775–1783.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Mansour, S. L., and Martin, G. R. (1988). Four classes of mRNA are expressed from the mouse int-2 gene, a member of the FGF gene family. EMBO J. 7, 2035–2041.

Marin, O., Meggio, F., Marchiori, F., Borin, G., and Pinna, L. A. (1986). Site specificity of casein kinase-2 (TS) from rat liver cytosol: a study with model peptide substrates. Eur. J. Biochem. 160, 239-244.

Marshak, D. R., and Carroll, D. (1991). Synthetic peptide substrates for casein kinase II. Meth. Enzymol. 200, 134-156.

Meggio, F., and Pinna, L. A. (1988). Phosphorylation of phosvitin by casein kinase-2 provides the evidence that phosphoserines can replace carboxylic amino acids as specificity determinants. Biochim. Biophys. Acta 971, 227–231.

Meier, U. T., and Blobel, G. (1990). A nuclear localization signal binding protein in the nucleolus. J. Cell Biol. *111*, 2235–2245.

Newmeyer, D. D., Finlay, D. R., and Forbes, D. J. (1986). In vitro transport of a fluorescent nuclear protein and exclusion of non-nuclear proteins. J. Cell Biol. *113*, 2091–2102.

Nicchitta, C. V., Migliaccio, G., and Blobel, G. (1991). Biochemical fractionation and assembly of the membrane components that mediate nascent chain targeting and translocation. Cell 65, 587–598.

Nigg, E. A., Baeuerle, P. A., and Lührmann, R. (1991). Nuclear importexport: in search of signals and mechanisms. Cell 66, 15-22.

Nikkodem, V., and Fresco, J. R. (1979). Protein fingerprinting by SDSgel electrophoresis after partial fragmentation with CNBr. Anal. Biochem. 97, 382–386.

Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., and Sharp, P. A. (1986). Splicing of messenger RNA precursors. Annu. Rev. Biochem. 55, 1119–1150.

Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA 85, 2444–2448.

Pfaff, M., and Anderer, F. A. (1988). Casein kinase II accumulation in the nucleolus and its role in nucleolar phosphorylation. Biochim. Biophys. Acta 969, 100–109.

Pfeifle, J., and Anderer, F. A. (1984). Isolation and localization of phosphoprotein pp 135 in the nucleoli of various cell lines. Eur. J. Biochem. *139*, 417–424.

Pfeifle, J., Hagmann, W., and Anderer, F. A. (1981). Cell adhesiondependent differences in endogenous protein phosphorylation on the surface of various cell lines. Biochim. Biophys. Acta 670, 274–284.

Pinol-Roma, S., and Dreyfuss, G. (1992). Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature 355, 730-732.

Prestayko, A. W., Klomp, G. R., Schmoll, D. J., and Busch, H. (1974). Comparison of proteins of ribosomal subunits and nucleolar preribosomal particles from Novikoff hepatoma ascites cells by two-dimensional polyacrylamide electrophoresis. Biochemistry *13*, 1945–1951.

Rechsteiner, M., and Kuehl, L. (1979). Microinjection of the nonhistone chromosomal protein HMG1 into bovine fibroblasts and HeLa cells. Cell *16*, 901–908.

Roberts, B. L., Richardson, W. D., and Smith, A. E. (1987). The effect of protein context on nuclear location signal function. Cell *50*, 465–475.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain termination inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.

Scheer, U., and Rose, K. M. (1984). Localization of RNA polymerase I in interphase cells and mitotic chromosomes by light and electron microscopic immunocytochemistry. Proc. Natl. Acad. Sci. USA 81, 1431–1435.

Schnell, D. J., Blobel, G., and Pain, D. (1990). The chloroplast import receptor is an integral membrane protein of chloroplast contact sites. J. Cell Biol. *111*, 1825–1838.

Schuurman, R., and Keulen, W. (1991). Modified protocol for DNA sequence analysis using Sequenase 2.0. BioTechniques 10, 185.

Shelness, G. S., and Blobel, G. (1990). Two subunits of the canine signal peptidase complex are homologous to yeast SEC11 protein. J. Biol. Chem. 265, 9512–9519.

Silver, P., Sadler, I., and Osborne, M. A. (1989). Yeast proteins that recognize nuclear localization sequences. J. Cell Biol. 102, 983–989.

Siomi, H., Shida, H., Nam, S. H., Nosaka, T., Maki, M., and Hatanaka, M. (1988). Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. Cell 55, 197–209.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503–517. Stochaj, U., Osborne, M., Kurihara, T., and Silver, P. (1991). A yeast protein that binds nuclear localization signals: purification localization, and antibody inhibition of binding activity. J. Cell Biol. 113, 1243–1254.

Thomas, P. S. (1980). Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77, 5201–5205.

Tokuyasu, K. T. (1973). A technique for ultracryotomy of cell suspension and tissue. J. Cell Biol. 57, 551–565.

Wang, T., Foker, J. E., and Malkinson, A. M. (1981). Protein phosphorylation in intact lymphocytes stimulated by concanavalin A. Exp. Cell Res. *134*, 409–415.

Warner, J. R. (1989). Synthesis of ribosomes in Saccharomyces cerevisiae. Microbiol. Rev. 53, 256–271.

Wehner, J. M., Sheppard, J. R., and Malkinson, A. M. (1977). Densitydependent phosphorylation of a specific protein in cultured cells. Nature 266, 842–844.

Wickens, M., and Stephenson, P. (1984). Role of the considered AAU-AAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation. Science 226, 1045–1051.

Yamasaki, L. Y., Kanda, P., and Lanford, R. E. (1989). Identification of four nuclear transport signal-binding proteins that interact with diverse transport signals. Mol. Cell. Biol. 9, 3028–3036.

GenBank Accession Numbers

The accession numbers for the sequences reported in this article are M94287 (pTM17) and M94288 (pTM6).