

Nexus formatted alignment of group I introns

#NEXUS

begin data;
dimensions ntax=32 nchar=131;
format datatype=dna interleave=yes gap=- missing=?;
matrix
PteS516 AGCGCCAGATGTCTCGCATA CACTTCTAAATGCGGGGAG TGCAGGGTTGATCCGACAGG AAGCGCGAAGGGCAACCTTC AGAGACTCTAATGGAGTGGG
AspS516 ACAGCCAGATGTCAATGTA CAACCGTAAATGCGGGGAA CCATGGGCAATCCGACAGC AAGCTCCTGGGAGAAAGTTTC AGAGACTATATCGGATGGGC
PflS516 AAAGCCGGGATAAGCTGTA CACGCGTAAATGCGGGGAA CACTCAGTCAATCCGACAGC AAGCCCGTGGGGAAGTTTC AGAGACTATAATCGTGTGGG
BgrS516 ATAGCCAGATGTCCGCTGTA CAGTGTAAATGCGGGGAA CACTTGGTCAATCCGACAGC AAGCCCGTGC - GGAAGTTTC AGAGACTATAATCGACTGGG
FseS516 AT - GCCAGATGTCCGATGTA CAGCTGTAAATGCGGGGAA CGCTGGCCAATCCGACAGC AAGTCGCTGC - GGAAGTTTC AGAGACTATAATCGTGTGGG
GviS943 TCTGCGCTAACAGCAGCGA CACGCTCAAATGCGGGAAA CTCTGGGCAATCCGACAGC AAGCCCGACGGGACGGTTTC AGAGACTAGATGGGTGTGGG
BbaS943 TCTGCTCCCAAAGCAGCAA CACCCTCAAATGCGGGAAA CTCTGGGCAATCCGACAGC AAGCCCTCACGGGGAAGTTTC AGAGACTTAATGGGGTGGG
PchS943 TCCGCGCTGCG - GCAGCAA CACACTAACTGCGGGAAC GTTATGAACTATCCGACAGC AAGCCCTAGGGTGCAGTTTC AGAGACTTAATGGGTGTGGG
AleS956 TTAGGCCCTAAAGA - GTAA CACCATCAAATGCGGGAAC CACTCAAGTTGACCGGACAGC AAGCCCTAGGGTGCAGTTTC AGAGACTTAATGGGTGTGGG
AspS956 TTAGGCCCTAAAGAAGTAA CACCATCAAATGCGGGAAC CACTCAAGTTGACCGGACAGC AAGTTCCTGGGGCGCAGTTTC ACAGACTAAACGGTGTGGG
AroS956 TTAGGCCCTAAAGCGGTA CACTTCAAATGCGGGAAC CTCTGGGATAATGAGCAGCA AATCCTATGGATGCAGTTTC ACAGACTTAACGGTGTGGG
DirS9562 CTAGGCCAGTGC - CAGCGA CACTATCAAATGCGGGGAG TCCTTGGGTTATCCGACAGC AAGTCCCTGGGGCGCAGTTTC ACAGACTAAATGGTGTGGG
FseS956 - TAGGCCCGCAAAGCAGCA CATCATCAAATGCGGGGAG TCCTTGGATCAATCCGACAGC AAGTCCCAGGGATGCAGTTTC ACAGACTAAATGGTGTGGG
TthL1925 TTTGAGGGAAGTGTGCAA GACCGTCAAATGCGGGAAA GGGGTCCTAACCCGACAGC AAGTCCCTATGGATGCAGTTTC ACAGACTAAATGTCGGTGGG
PpoL1925 GTAGGAGAGAAAGGTGCGA GACCGTCAAATGCGGGAAA GGGGTCCTAACCCGACAGC AAGTCCCAGGGATGCAGTTTC ACAGACTAAATGTCGGTGGG
PflL1949 GAGCTGAGCAGCAAAACCGT GGA - TCTAACCGGGAAG GGTGGCTGACTGGCGACC AAGCCGACGTGCGAAGGCCA AGAGACTATACGGTTCCTTA
PpoL1949 GGCGTGCAGTTGAAACTAT GGA - TCTAACCGGGAAG GGTGGCTGACTGGCGACC AAGCCGAGGGAGGAAGGCCA AGAGACTATACGGTTCCTTA
FseL1949 CCATGACAAGGGGCGCTTTA GGT - T - AAAATGCAAGAA GGTGGCTGACTGGCGACC AAGCCCTAGGGCGAAGTTTC AGAGACTAGAAAGCCGGAAG
DirL1949 GAAGTGACCAAGAGGGGCGA TGAA - GCCTCGTGAATTTG C - GGGCCTGACCGGACAGC AAGCCGCAAAAGGGAGTTTC AGAGACTATAGGCACTATG
PflL2449 AAGGTCCAAAACAAAGCGA CGCA - TCTGGATGCTGAGAC GGGCGGCTTATCAGCAAGC GACGAGGGGTTAAGCGCTCC AACGACTGGAAGTGTGGG
PpoL2449 GAGGTCCCAAGTAAAGGGCA CGCA - TCTGGATGCTGAGAC GGGCGGCTTATCAGCAAGC GACGAGGGGTTAAGCGCTCC AACGACTGGAAGTGTGGG
FseL2449 GAGGTTCCTATGCGCTGCGA CGCG - CTGGATGCTGGGAA GGTGGGACCCATCAGCAGC AACGAGGTTTAAAGTGTGGG AACGACTGGAAGTGTGGG
DirL2449 GGAGTCCCAAGGGAAGCTAA CGTA - CTAGATGCTGAGCAG CTTATGGATAATCCGACAGC AACGAGGGGCTACGTTGTTTC AGAGCGCGAAAGGTCACGGG
BhyS516E GGTGCCAGAGATAGTGCGA GACAACCTGGATCGGGGAC CTTATGGATAATCCGACAGC GAGCCGC - AAGGCTGT - GT AGAGCGCGCAAGGTCCTGG
BoxS516E GGCGCCAGAGATAGTGCGA GACAACCTGGATCGGGGAC CTTATGGATAATCCGACAGC GAAATT - - - ACTTCTGT - GT AACGCGCGAAAGGTCACGGG
GviS516E GAGCCAGAGATAGTGCGA GGTGACCTGGTACGGGGAAG CTTATGGATAATCCGCTGCG GAGCCCGCCGGGCGCTGCT AGACCGCGCAAGGTCACGGG
PobS516E GAGCCAGAGATAGTGCGA GGT - ACTGGTACGGGGAAG CTTATGGATAATCCGCTGCG GAGCCCGCCGGGCGCTGCT AGACCGCGCAAGGTCACGGG
SknS956E CTAGGCCCGATAGTGCGA GACAACCTGGAACGGGAAA GCTATGGGTCCTCCGCGAG CTTATGGATAATCCGCTGCG CATCCCTCAAAGGTCCTGCT AACGCGCGCAAGGTCCTGG
DirS9561E CTAGGCCCGATAGTGCGA GACAACCTGGAACGGGAAA GCTATGGATAATCCGCTGCG GAGCTACAGATAGCGCTGCT AACGCGCGCAAGGTCCTGG
DspS956E TTAGGCCCAAGAGTGCGA GACAACCTGGTACGGGGGAA CACTGGCTCAATCCGCTGCT GAGCTTGAATAAGCCATCGT AACGCGCGCAAGGTCCTGG
RcuS1199E CAGCCCTGTTCAACCTGAA CACCG - CTGGTACAGGAGAC CTATGGATCGATTCTGTGCG GAGCTTCCGAAAGCCGCGCA AACGACTGCAAAAGTGTGGG
EmyS1199E CAGCCCTGTAGCAACGCTA CACCA - CTGGTATAGA - GGA TGCATGGCTGATTCTATGCG GAGTC - - AAAGGACCGTCTC AACGCGCGCAAGTGTGGG

PteS516 CAGCTTAAGGGAGAGTCCAA TCC - ACAATGG
AspS516 GACGTTAAGGTATAGTCCAC TCCCAACAAGG
PflS516 GCTCTTAAGGTATAGTCCAC TCCCCAAGGG
BgrS516 CAGCTTAAGGAATAGTCCAC TCCCACTAGGG
FseS516 CAGCTTAAGGGATAGTCCAC TCCCACTAGGG
GviS943 TAGCTTAAGATATAGTCCGG TATTGGATTTG
BbaS943 TAGCTTAAGATAAAGTCCGT CTATACAAAACG
PchS943 TTACTTGAGCTTTAATCCGT TCGCAATTACG
AleS956 TCACTTAAGATATAGTCCGG CCCATCAACCG
AspS956 CCGCTTAAGATATAGTCCGG CCCATCAACTG
AroS956 CCGCTTAAGATATATTCGGT CCCAACAATG
DirS9562 CTGCTTAAGATATAGTCCGG CCCAACAACCG
FseS956 TTGCTTAAGATATAGTCCGG CCCAACAACCG
TthL1925 GCTCATAAGATATAGTCCGA CCTCTCACTCG
PpoL1925 GCTCATAAGATATAGTCCGG CCTTCCCAGCG
PflL1949 AAAACGAAGAGATAGTCCAC CCACCAGGGTG
PpoL1949 AATATAAAGAGATAGTCCAC CCAACGGGGTG
FseL1949 GAAAAAAAAGGATAGTCCAC CAGACTCGGGT
DirL1949 GCCTACTAGTAAATAGTCCAC CGACAGAGGTT
PflL2449 CAGTTCAAGGAACAGTCTCT CTCGACATTTGG
PpoL2449 TCGTTCAAGGTACAGTCTCT CCTGGTATCGG
FseL2449 CGGCTTAAGGTACAGTCTCA AGCGGACTCGG
DirL2449 TCATCCAAGGAACAGTCTAC CCCTGCTAGGG
BhyS516E TGGCTTGAAGTACGTCCGA TCCCAAAATCG
BoxS516E TGGCTCGAGGACGTGCCAA CCACCTTAGTG
GviS516E TGACTCAAGGACGTGCTGT CCCATCAACG
PobS516E CGGCTCAAGGACGTGCTGT CCCATCAACG
SknS956E GGCCTTAAGGTACGTGCTAC TCCCTCAAAGG
DirS9561E GATCTTAAGGGACGTGCTAC TCCCCCAAAGG
DspS956E GCCCTTAAGGTACGTGCTAC TCCCATCAAGG
RcuS1199E CGGCTTAAAGTACAGTCTGT CCACCCACTCG
EmyS1199E CGTCTTAAGGTACGTGCTGG CACTCT?????

; end;

Nexus formatted alignment of homing endonucleases

#NEXUS

begin data;
dimensions ntax=29 nchar=117;
format datatype=protein interleave=yes missing=? gap=-;
matrix
CallL1923 ?????????? - V - LQIGFT KTRENKYQY - LHRSF - - - S SVKKDFGLE - KGFYDNTMTV SHLCHNSWCYNW - NHHILES LAVNKS RN - GCPG - - - - -
PpoL1925 LHCEYIIPAP - - - - - VGFA KNGPTRWQY - - KRTIN - - - Q VV - HRWGEPEDNINGKTCTA SHLCHNTRCHNP - LHLCWES LDDNKGRN - WCPGPN - - - - -
NaeL1926 NGCWELRRRTK - T - VKPGF - KTKG - KLQY - - KIEGR - - - Y VVAHRVWFR - EVLGRPNATV SHRCHNEHCLNP - DHLVVEP LQVNSQRN - TCT - - - - -
MgaS1199 ARTRSRGRPHK - P - KL - - - - - GHR - - - - L - AVIAWHP - EE SRR - - - - - RLLDDGWAH - SHLCGNPLCLEP - AHIHVEP KTANEAR - GCRFVSSS - -
EmyI1919 ERKRAKAKKV - K - DK - - - - - AHRVVVLL - HHRYDNP - LT SLR?Y? - - - - - IQRLLHQGGHT - SHLCHQPTCINS - SHLVVED KAANEAR - ICKGKVCFL - -
DirS9561 GMCLMSPPRH - Q - MR - SKG RRFKIQASA - LQAVLVN - N PSSHDELVLGLDRETTFFHS SHLCKDGCME - L - KHTLRVP AQTNLADHELCPAVVVIYGV
DirS9562 IQCLVLPSSD - T - YPRCDG KASLNSLAA - - QFGEDQ - RS FDQTSPLLEENDWEMIDDDA IHRCHNRCFTK - GHVYVFG KDQNRST - FCPAYVINGV
SdiS943 LACWTSPEPG - - - - - KKLNW RNRPSKEIG - HVFYAHH - IA VAAG - - - - - ISGAH - - - - - AAYNV SHLCGNGEFCNP - DHLVIET AGLNQRK - TCQAQVRLPN
CpiS943 GSCWAA - ESHG - R - NGRNTV G - PTGQKIG - VSPWAHQ - LA YLHNSI IAACKFKVLQEV SHLCHNAGCINP - DHVIEP AWINEFR - FCAWVRFPDL
MteS943 LGLNLQIREGN - A - AKKVTL TNKKY - KLG - KV - YMHV - VA IYLDNSI IAACKFKVLQEV SHLCHGKNCIRD - GHHNVES PEMNQSRD - AHY - - - - - YDP - - - S
PchS943 LGCWESNRPN - S - G - YQRNTV I - - - - - NGRKA - VQIYLHQ - LT LVIADRN - AAVLESELAME SHLCHDGRCFR - SHLVVEP AALNARN - SCHGSVIK - Y
BbaS943 LGCWFPKRPV - M - NYRNTV WDSYVLEAA - AQSYYHH - LC VVAKREGAGRKDSNGIAYEV SHLCHNGQCFNP - DHCVVEP RHNHRN - ECQNRVILNLV
EmyS943 TCCWLSGAPAH - N - NGRNTN KPCTNHKFD - CNPFFHQ - LA VAAG - - - - - QRLRTDGTGYH SHLCHNPQCFNP - EHLVIEL SHLCHNACVNP - LHMAPES GDVNSRL - YCRLFKNHP - E
CmoS943 RNCWPKTPKTH - R - NGRHTL LDASDKKIG - RCILYLHH - LT VAANKA - L - NAGKK - RLQ - V SHLCHRGWCFNP - EHLVIES QRLNARN - NCKGNFI - I
Bats1506 NHCITTTPAKD?S?AG - - - - - YVQVRW?G?VKYVCHL?IA VAA - - - - - MQRHPMPGEEA SHLCKDKRCVNP - SHMTLES GDLNKSRI - YCRLMRNHP - E
PteS1506 QGCHITNAAN - - - - - G - - - - - YVQVRW?G?VKYVCHL?IA VAA - - - - - MQRHPMPGEEA SHLCKDKRCVNP - SHMTLES GDLNKSRI - YCRLMRNHP - E
Pps1506 NQWPG - PVFR - P - YG - DGE YSQRVFCG - TKYAVHR - VG VGYLDDTM - - - - - - - - - - - SLSGRTTININP - YHLHTVR NALNKRH - ACHFLFLEYR - N
AsuS1506 IT - WPLDVPVYS - P - YAKDG - YSQRVSAALNKYVTHR - VG VDPAEKD? - - - - - SLIP - - - - - SLSGRTTININP - VFLAAET NALNQRST - ACFLFPDELLT
TorS1506 DPQWPN - PVF - P - LLKSEP TITADLLAAKVYVYHR - IA VNHPEDEA?MIDFTLQPNAN - - - - - SNGSININP - LQLIWD NSTNQTRK - GCFIHMRSK - A
PteS1506 RTCLLQ - - - - - E - T - PR - - - - - GRYTHQOR - KHVY - Y - QA V - AYARFGTTSKIGTDL - TI SHVCGTHCLAA - HHLMLPEP RAENKERO - TCNGHPIIK - Y
AspS516 RSCLL - - - - - LPA - - - - - GRFSY - QR - RKVRFY - QI V - AFKFPGBASK - DRSSKLL SHLCGTR?CCNV - DHLVIET KGINDERT - HCH??????L
NtS516 FGCLIG - AANK - T - KS - - - - - VRFLNNKG - SDSYVHH - VS AFANSTNESSSKDPEARTI SHLCGNGGCARP - GHLRIEK KSVNDERT - HCHFLRRSR
NcL516 FGCLIG - AANK - T - KA - - - - - VRFLNNKG - SDSYVHH - VS AFANSTNESSSKDPEARTI SHLCGNGGCARP - GHLRIEK KSVNDERT - HCHFLRRSR
NcaS516 FGCLIG - AANK - T - KS - - - - - VRFMNNKG - SDSYVHH - AS SRLLIRIVSSSKRDHARTI SHLCGNGGCARP - GHLRIEK KSVNDERC - HCHFLRRSR

```

NjaS516 FGCLIG-AASK-T-KS---- -VRFMNNKG-SDSYVHH-VS VFANSTCESSKDPDARTI SHLCGNGGCARP-GHLRIEK KSVNDERT-HCHFLLRRSQR
NloS516 FGCLIG-AANK-T-KS---- -VRFMNNKG-SDSYVHH-VS VLANSKAESARKTEDDARTI SHLCGNGGCCRP-GHLRIES KKTNDERT-HCHFVLRNSKR
NgrS516 FGCLIG-AATK-T-KS---- -VRFMNNKG-SDSYVHH-VS AFANSTNESSKDEARTI SHLCGNGGCSR-P-GHLRIEK KSVNDERT-HCHFLLRRSQR
NanS516 FGCLIG-AANK-T-KS---- -VRFMNDKG-SDSYVHH-VS VFANSTCESSKDPDARTI SHLCGNGGCARP-GHLRIEK KTVNDERT-HCHFLLRRSQR

CalIL1923 NSCRH-KVKCIIPGEYS
PpoL1925 GGCVH-AVVCRLRQGPLY
NaeL1926 GVCLH-RPKCIRDGNQA
NgaiS1199 -PCPC-TPACVTTVSTA
EmyIIS119 SPCPH-TPSCLFASREE
DirS9561 NLCTC-SRQCLVPGRRF
DirS9562 HVCTH-TPSCLVPGKRA
SdiS943 HPCQH-GKLCILPLRRA
CpiS943 DPCCH-VR-CLIPTWYK
MteS943 SECDH-HPQCIAP-RPA
PchS943 NPCRH-GRDCLLPVKV
BbaS943 DLCPH-?QSCLLPVRTV
EmyS943 HPCTH-WLQCVLPIREI
CmoS943 NPCPH-GPSCILPIKHI
BatS1506 FKCLH-VPECIHTQT--
PteS1506 LVCLH-DPACL-----
PspS1506 FKCLH-TPQCV-----
PpaS1506 VTCHR-SKRCKVHWGLW
AsuS1506 NGCLH-KLCCV-CWDEW
TorS1506 YVCFH-PKPKAYQNW
PteS516 LACPH-LPRCYSV----
AspS516 GGCPH-TPRCGSITLLE
NitS516 LACPH-SPKCFVNIYKI
NclS516 LACPH-TPRCFVNIYKI
NcaS516 MACPH-TPRCFVNIYKI
NjaS516 LACPH-TPRCFVNIYKI
NloS516 MACPH-EPKCFVNIYKI
NgrS516 LACPH-TPKCFVNIYKI
NanS516 LACPH-TPRCFVNIYKI

```

```

;
end;

```