Electronic Supplementary Material (ESI) for ChemComm.

Supplementary Information

Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation

Annabel C. Murphy, Hui Hong, Steve Vance, R. William Broadhurst and Peter F. Leadlay

General Analytical Procedures

HPLC-MS analysis was performed using an HPLC (Hewlett Packard, Agilent Technologies 1100 series) coupled to a Finnigan MAT LCQ mass spectrometer fitted with an electrospray ionization (ESI) source. The HPLC was fitted with a Jupiter 5µ C4 column (2.0 x 250 mm, Phenomenex). A solvent system of CH_3CN and H_2O, both containing 0.1% trifluoroacetic acid, was used. Samples were eluted at a flow rate 0.3 mL min^{-1} with a linear gradient of 5 to 45% CH_3CN over 5 min, followed by 45 to 55% CH_3CN over 13 min, then 55 to 95% CH_3CN over 2 min and finally 95% CH_3CN over 4 min. The mass spectrometer was run in positive ionization mode, scanning from m/z 200 to 2000, and the collision energy was set to 35%. HPLC-MS data were processed using Xcalibur (version 1.1, Thermo Finnigan) and deconvolution of the various charge states observed was carried out using Promass 2.8 (Novatia).
DNA sequencing was carried out by the DNA Sequencing Facility in the Department of Biochemistry, University of Cambridge. High resolution mass spectroscopy was carried out using positive ion electrospray ionization on a Thermo Fisher Orbitrap Mass Spectrometer.

Bacterial strains and culture conditions

E. coli strains (DH10B and BL21 DE3) were grown in Luria-Bertani (LB) broth (1% tryptone, 0.5% yeast extract, 1% NaCl) or on LB agar (1% tryptone, 0.5% yeast extract, 1% NaCl, 1.5% agar) at 37 °C with kanamycin selection (50 µg mL^{-1}).

Materials, DNA isolation and manipulation

Oligonucleotides used in this work are summarised in Table S1. Restriction endonucleases (Fast Digest) and Pfu DNA polymerase were purchased from Thermo Scientific, shrimp alkaline phosphatase was purchased from USB, T4 DNA ligase was purchased from New England Biolabs. Chemicals were from Sigma-Aldrich. Plasmid DNA was isolated from an overnight culture using the Plasmid Mini Kit I (Omega BioTek). Purification of DNA fragments from agarose gels was performed using the Anachem Gel Recovery Kit, then concentrated using the Zymo DNA Clean and Concentrator kit. PCR amplifications were carried out using a Mastercycler (Eppendorf) and the Phusion High-Fidelity PCR Master Mix (GC buffer) from New England Biolabs, with a final concentration of DMSO of 3-4%.

A C-terminal His-tag EryKS3AT3 was generated by amplification using the primers given in Table S1. The PCR product was digested with NdeI and EcoRI, purified by agarose gel, extracted and concentrated. This fragment was then ligated into pET29b that had been previously digested with the same restriction enzymes and shrimp alkaline phosphatase, purified by agarose gel, extracted and concentrated. The same protocol was used to clone EryACP3 into pET28a to generate EryACP3 with an N-terminal His-tag. The MlsAT9 gene was cloned into a pET28a(TEV) (N-terminal sequence: MGSSHHHHHHSSGENLYFQG) using the primers given in Table S1 by one step, isothermal Gibson Assembly.[1]

For general DNA manipulations, *E. coli* strain DH10B was used, while for protein expression *E. coli* strain BL21 DE3 was used. Chemically competent cells were generated using calcium chloride and then transformed with the plasmids described above, and selected on LB agar plates containing kanamycin (50 μg mL^{-1}).

Site-directed mutagenesis was carried out using the Quikchange method (Stratagene) with inclusion of DMSO at 6% and the primers given in Table S1.

Model generation

The amino acid sequence of the 16 MLS KS domains were aligned with the KS domain of DEBS module 3 and a sequence identity of 63% was found (excluding the DEBS N-terminal docking domain). An *in silico* model of the KS from MlsA2 was generated using Phyre2 and iTASSER.[2,3] Beyond minor differences in side-chain placement, the programs generated very similar models of the active site and the Phyre2 model was then compared with the EryKS3AT3 didomain crystal structure.

Protein expression and purification

Expression of EryKS3AT3 and all mutant variants was carried out in 150 mL LB in 500 mL Ultra Yield flasks (Thompson) with kanamycin selection and, after incubation at 37 °C until an OD_{500} of 0.6, the culture was cooled to 22 °C and induced with IPTG (0.2 mM) for 16 h. Purification was then carried out by incubating the lyzed cell supernatant (10 mM imidazole, 50 mM sodium phosphate pH 7.2, 300 mM NaCl) with Ni-NTA resin at 4 °C for 1 hour, washing with 10 column volumes of the above loading buffer, then eluting with 150 mM imidazole, 50 mM sodium phosphate pH 7.2, 300 mM NaCl. Preparative gel filtration using Sephadex S200 (column volume approximately 120 mL) and eluting with 100 mM sodium

phosphate buffer at pH 7.2 was carried out, followed by concentration of the resulting protein using Amicon Ultra-4 30 kDa cutoff spin concentrators (Millipore) to an approximate concentration of 30 µM.

The same protocol was used for apo-EryACP3 except that expression was carried out using a 500 mL LB culture in a 2L baffled flask, and elution was carried out using 500 mM imidazole, 50 mM sodium phosphate pH 7.2, 300 mM NaCl. After desalting into 100 mM sodium phosphate buffer (pH 7.2) using a PD10 cartridge (GE life sciences), the apo-EryACP3 was phosphopantetheinylated *in vitro* by incubation with sfp from *Bacillus subtilis*,[4] (8 µM), $MgCl_2$ (1 mM), DTT (5 mM) and coenzyme A (1 mM) at 37 °C for 1 h. Thrombin digestion was carried out as follows to simplify mass spectral analysis of the protein, by removing partial glucoronidation post-translational modification.[5] Thrombin (10 µM, from bovine plasma, Sigma) was added to the above mixture and the reaction was incubated for a further 16 h at room temperature. The holo-EryACP3 was then purified using preparative gel filtration using Sephadex S200 and eluting with 100 mM sodium phosphate buffer at pH 7.2, then concentrated using an Amicon Ultra-4 3 kDa cutoff spin concentrator (Millipore). MlsAT9 was expressed and purified by Ni-affinity chromatography following the same protocol as EryACP3. Gel filtration was carried out on Superdex75 (10/300) column (GE Healthcare) in 50 mM sodium phosphate pH 7.5, 150 mM NaCl. The purified protein was concentrated using a 10 kDa cutoff PES Vivaspin spin concentrator (Sartorius) to a concentration of 50 µM. Protein concentration of all EryKS3AT3 variants (Σ 91890 $M^{-1}cm^{-1}$) and MlsAT9 (Σ 31400 $M^{-1}cm^{-1}$) was estimated using absorbance at 280 nm, while that of EryACP3 was estimated using amino acid analysis and Bradford's method.[6] All EryKS3AT3 variants were freshly expressed, purified and stored for a maximum of 3 days at 4 °C prior to use, while EryACP3 was diluted to give a final glycerol concentration of 20%, flash frozen and stored for up to 2 months at -80 °C.

Chemical Synthesis

N-acetylcysteamine substrates **3-19** were prepared using the standard procedure described below, or according to literature procedures, and their identity was confirmed by comparison with literature values, except **7**, **10** and **13**, for which data are given below.[7–14] The appropriate carboxylic acid was dissolved in anhydrous DCM, cooled on ice, and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (1.1 eq) and dimethylaminopyridine (0.2 eq) were

added. After stirring on ice for 15 minutes, cysteamine (1.1 eq) was added, the reaction was allowed to warm to room temperature and stirred for 16 hours. After removal of the solvent *in vacuo*, the residue was dissolved in chloroform and washed with 0.1 M HCl followed by brine, and the organic layer was dried over $MgSO_4$. After removal of the solvent *in vacuo*, the residue was purified by silica chromatography using mixtures of petroleum ether and ethyl acetate.

7 96%, colourless oil, R_f 0.13 (60:40 EtOAc:petroleum ether), δ_H ($CDCl_3$, 400 MHz) 3.41 (2H, q, 6.3 Hz), 3.00 (2H, t, 6.3 Hz), 2.57 (1H, sex, 6.9 Hz), 1.94 (3H, s), 1.71 (1H, m), 1.46 (1H, m), 1.14 (3H, d, 6.9 Hz), 0.89 (3H, t, 7.3 Hz); δ_C ($CDCl_3$, 100 MHz) 204.5, 170.2, 50.1, 39.8, 28.1, 27.1, 23.1, 17.1, 11.5; *m/z* = 204.1051 $[M+H]^+$, 1 ppm (204.1053 calcd for $C_9H_{18}O_2N^{32}S$);

10 73%, yellow oil, R_f 0.13 (60:40 EtOAc:petroleum ether), δ_H ($CDCl_3$, 400 MHz) 6.86 (1H, q, 7.1 Hz), 3.44 (2H, q, 6.0 Hz), 3.06 (2H, t, 6.4 Hz), 1.95 (3H, s), 1.87 (3H, br s), 1.83 (3H, d, 7.1 Hz); δ_C ($CDCl_3$, 100 MHz) 193.8, 170.2, 136.84, 136.78, 39.8, 28.3, 23.2, 14.4, 12.1; *m/z* = 202.0893 $[M+H]^+$, 2 ppm (202.0896 calcd for $C_9H_{16}O_2N^{32}S$);

13 73%, white powder, R_f 0.14 (60:40 EtOAc:petroleum ether), δ_H ($CDCl_3$, 400 MHz) 6.89 (1H, dd, 15.4, 6.6 Hz), 6.07 (1H, dd, 15.4, 1.5 Hz), 3.45 (2H, q, 6.0 Hz), 3.08 (2H, t, 6.4 Hz), 2.46 (1H, m), 1.96 (3H, s), 1.07 (6H, d, 6.8 Hz); δ_C ($CDCl_3$, 100 MHz) 190.7, 170.3, 153.6, 125.7, 39.8, 31.0, 28.3, 23.2, 21.1; *m/z* = 216.1052 $[M+H]^+$, 0.3 ppm (216.1053 calcd for $C_{10}H_{18}O_2N^{32}S$).

N-acylcysteamine substrate stock solutions used: **18**-**20** were dissolved at 60 mM in water, **6** was dissolved at 60 mM in DMSO and all other substrates were dissolved at 300 mM in DMSO.

Condensation assays using methylmalonyl-CoA

Methyl malonyl coenzyme A (1 mM), tris(2-carboxyethyl)phosphine (500 µM), EryACP3 (30 µM) and EryKS3AT3 or its mutants (3 µM) were incubated in sodium phosphate buffer (100 mM, pH 7.2) for 30 mins at room temperature. *N*-acetylcysteamine substrates (6 mM) were then added and the reaction incubated at room temperature for a further hour. These reactions were carried out on a 10 µL scale, and were then quenched with 40 µL 0.1% aqueous TFA and the samples were analyzed by LCMS within 16 hours (stored at 10 °C). The reaction conditions used for **6** were adjusted to minimise precipitation observed on addition of the substrate. Methyl malonyl coenzyme A (0.4 mM), tris(2-carboxyethyl)phosphine (200

μM), EryACP3 (12 μM) and EryKS3AT3 or its mutants (1.2 μM) were incubated in sodium phosphate buffer (100 mM, pH 7.2) for 1 hour at room temperature. **6** (3 mM) was added and the reaction incubated at room temperature for one hour. The reaction was carried out on a 25 μL scale, then quenched with 25 μL 0.1% aqueous TFA and analyzed by LCMS as described in the general analytical procedures.

Malonylation of EryACP3 with MlsAT9

Malonyl coenzyme A (1 mM), tris(2-carboxyethyl)phosphine (5 mM), EryACP3 (30 μM) and MlsAT9 (3μM) were incubated in sodium phosphate buffer (100 mM, pH 7.2) for 1 hour at room temperature, then quenched and analysed as above (Figure S2b). A control reaction lacking the MlsAT9 was carried out (Figure S2a). To confirm that the reaction proceeds in the presence of EryKS3AT3, the above reaction was repeated with the addition of EryKS3AT3 (3 μM), and to confirm EryKS3AT3 cannot use malonyl coenzyme A as a substrate the above reaction was carried out with EryKS3AT3 (3 μM) in place of MlsAT9 (Figures S2c and S2d).

Condensation assays using malonyl-CoA

Malonyl coenzyme A (1 mM), tris(2-carboxyethyl)phosphine (500 μM), EryACP3 (30 μM) and MlsAT9 (3μM) were incubated in sodium phosphate buffer (100 mM, pH 7.2) for 30 mins at room temperature. *N*-acetylcysteamine substrates (6 mM) and EryKS3AT3 or its mutants (3 μM) were then added and the reaction incubated at room temperature for a further hour. These reactions were carried out on a 10 μL scale, and were quenched and analyzed as for the methylmalonyl condensation assay above. The reaction conditions used for **6** were adjusted as above; malonyl coenzyme A (0.4 mM), tris(2-carboxyethyl)phosphine (200 μM), EryACP3 (12 μM) and MlsAT9 (1.2 μM) were incubated in sodium phosphate buffer (100 mM, pH 7.2) for 1 hour at room temperature. **6** (3 mM) and EryKS3AT3 or its mutants (1.2 μM) were added and the reaction incubated at room temperature for one hour. The reaction was carried out on a 25 μL scale, then quenched with 25 μL 0.1% aqueous TFA and analyzed as above.

Condensation assays using alkyl-ACP as an internal standard

EryACP3 was alkylated with iodoacetamide following the protocol described by Jenner *et al.*.[15] Methyl malonyl coenzyme A (1 mM), tris(2-carboxyethyl)phosphine (500 μM),

EryACP3 (30 µM) and wild type or Ala154Trp EryKS3AT3 (3 µM) were incubated in sodium phosphate buffer (100 mM, pH 7.2) for 30 mins at room temperature. *N*-acetylcysteamine substrates (6 mM) were then added and the reaction incubated at room temperature for a further hour (10 µL final volume). The reactions were quenched with 40 µL 0.1% aqueous TFA, alkylated EryACP3 was added to a final concentration of 6 µM and samples were analyzed by LCMS as described in the general analytical procedures.

Time course assays for wild type and Ala154Trp eryKS3AT3

Methyl malonyl coenzyme A (1 mM), tris(2-carboxyethyl)phosphine (500 µM), EryACP3 (30 µM) and wild type EryKS3AT3 or Ala154Trp (3 µM) were incubated in sodium phosphate buffer (100 mM, pH 7.2) at room temperature for 30 mins. Substrate **4** (6 mM) was added and aliquots (10 µL) were then extracted at various time points, quenched with 40 µL 0.1% aqueous TFA and analyzed by LCMS within 16 hours (stored at 10 °C). No significant side reactions were observed on this occasion, and since the condensation product peak was well resolved from the holo/methylmalonyl peak, estimation of peak areas was carried out spectrophotometrically (210 nm) rather than by deconvolution of LCMS data. Data was then fitted to equation 1 using OriginPro 8.[16] Wild type EryKS3AT3 had a v_0 value of 10.6 ± 0.4 mMs^{-1} and an η value of 0.009 ± 0.001 s^{-1}, while Ala154Trp had a v_0 value of 47.2 ± 0.8 mMs^{-1} and an η value of 0.048 ± 0.001 s^{-1}.

$$[P] = \frac{v_0}{\eta}(1 - e^{-\eta t}) \quad (1)$$

Data analysis

The EryKS3AT3 and its mutants had an approximate retention time of 16 mins, all EryACP3 derivatives had retention times between approximately 17 and 23 mins, and MlsAT9 had a retention time of 24 mins. To allow quantification of the various EryACP3 derivatives formed during the reaction, deconvolution was carried out on the mass spectrum generated by summing all data after the EryKS3AT3, or its mutants, had been eluted and before elution of MlsAT9. The proportions of each EryACP3 derivative present were then analyzed (with a possible output range of 9000-13000 Da). In those cases where very low levels of condensation product were obtained, deconvolution of the peak containing the

condensation product was carried out to confirm its presence, and a yield of <1% was quoted. For EryKS3AT3 and all mutants, with the exception of $A_{154}W$, the most abundant ACP derivatives observed were the holo-ACP (M_w 10477) and methylmalonyl-loaded ACP (M_w 10577). Other commonly observed minor derivatives were propionoyl loaded ACP, presumed to be formed by spontaneous decarboxylation of methylmalonyl loaded ACP (M_w 10533), and ethyl loaded ACP formed by decarboxylation of malonyl loaded ACP (M_w 10519), oxidised holo-ACP (M_w 10493) and methylmalonyl-loaded ACP (M_w 10593), presumed to be oxidized on one of the two methionine side chain sulfurs, and holo-ACP that has formed a disulfide bridge with an additional molecule of coenzyme A (M_w 11240). Acylation of the unloaded holo-EryACP3 via a (non-enzyme catalyzed) trans-thioesterification reaction with the *N*-acetylcysteamine substrates was commonly observed in small quantities, while for α,β-unsaturated *N*-acetylcysteamine substrates, ACP derivatives derived from conjugate addition of the phosphopantetheine sulfur to the substrate was often a major product of the reaction (up to 77%). Neither of these reactions was observed for the apo-EryACP3. For one batch of EryACP3, a holo-ACP + 32 Da derivative was observed (M_w 10507), presumed to represent a doubly oxidised phosphopantetheine sulphur (sulfinic acid), due to the absence of a methylmalonyl loaded equivalent. While the reducing agent TCEP was included to reduce the formation of holo-ACP dimers, its concentration was minimized to avoid its conjugate addition onto Claisen condensation products when α,β-unsaturated SNAc substrates were employed.

Table S2. ESI-MS analysis of proteins

Primers for cloning EryKS3AT3 and EryACP3

EryKS3AT3 Fwd ATATATAT**CATATG**ACTGACAGCGAGAAGGTGG

EryKS3AT3 Rev TATATATA**GAATTC**GCCCTTACGCGGTAGGCCAGCTCGTC

EryACP3 Fwd ATATATAT**CATATG**CGGCTCGCGGGGCTTTCC

EryACP3 Rev TATATATA**GAATTC**TTAGGCGTCACCGACGAGCCGGGC

Primers used for cloning of MlsAT9

MlsAT9 Fwd

AGCAGCGGCGAAAACCTGTATTTTCAGGGCCCCGACACCACACAAACC

MlsAT9 Reverse GCCAACTCAGCTTCCTTTCGGGCTTTGTTAGAAAGCGTAGGTGGGCAGTG

pET28a(TEV) Fwd TAACAAAGCCCGAAAGGAAGC

pET28a(TEV) Rev GCCCTGAAAATACAGGTTTTCG

Primers used for quikchange mutagenesis of EryKS3AT3

A_{154}G sense	CTTCCTCGGAGTGG**G**GAAGTTCGGCTACG
A_{154}G antisense	CGTAGCCGAACTTC**C**CCACTCCGAGGAAG
A_{154}W sense	GTCTTCCTCGGAGTG**TG**GAAGTTCGGCTACGG
A_{154}W antisense	CCGTAGCCGAACTTC**CA**CACTCCGAGGAAGAC
K_{155}A sense	CCTCGGAGTGGCG**GC**GTTCGGCTACGGC
K_{155}A antisense	GCCGTAGCCGAAC**GC**CGCCACTCCGAGG
F_{156}Q sense	CTCGGAGTGGCGAAG**CAA**GGCTACGGCGAGGAC
F_{156}Q antisense	GTCCTCGCCGTAGCC**TTG**CTTCGCCACTCCGAG
V_{173}M sense	CGAGGGCTACTCG**A**T**G**ACCGGTGTGGCGC
V_{173}M antisense	GCGCCACACCGGT**C**A**T**CGAGTAGCCCTCG
A_{230}T sense	TCGGCGGTGCC**A**CGGTGATGGCG
A_{230}T antisense	CGCCATCACCG**T**GGCACCGCCGA
F_{263}T sense	CGGCGCCGACGGG**AC**CGGCTTCTCCGAA
F_{263}T antisense	TTCGGAGAAGCCG**GT**CCCGTCGGCGCCG
F_{265}W sense	GCCGACGGGTTCGGCT**GG**TCCGAAGGCGT
F_{265}W antisense	ACGCCTTCGGA**CC**AGCCGAACCCGTCGGC

Table S2. ESI-MS analysis of proteins

	Calculated		Observed	
	+ Met	- Met	+ Met	- Met
EryKS3AT3	99989	99857	99992	99870
holo-EryACP3	10477	-	10477	-
$A_{154}G$	99975	99843	99996	99878
$A_{154}W$	100104	99972	100094	99973
$K_{155}A$	99932	99800	99940	99802
$F_{156}Q$	99970	99838	99957	99837
$A_{230}T$	100019	99887	100011	99890
$F_{263}T$	99943	99811	99928	99803
$F_{265}W$	100028	99896	100050	99909
MlsAT9	-	52288	-	52279

Values for proteins with (+ Met) and without (-Met) an N-terminal methionine are given

Figure S1. KS domain sequence alignment Sequence alignment of KS domains of known structure with a representative mycolactone KS domain. The positions of variable surface loops identified in X-ray crystal structures[17–21] are highlighted with the 'outer clasping loop' in yellow, the 'dimer interface loop' in blue and the 'active site cap' in red. Regions previously identified as involved in ACP docking events[22] are outlined in yellow for the "upstream" ACP and green for the intramodular ACP. Residues identified as possible specificity determinants[17] are highlighted with a closed circle. Residues that when mutated obstruct the intramodular ACP docking region[22] are indicated with open triangles. The seven EryKS3 residues mutated in this work are indicated with asterisks.

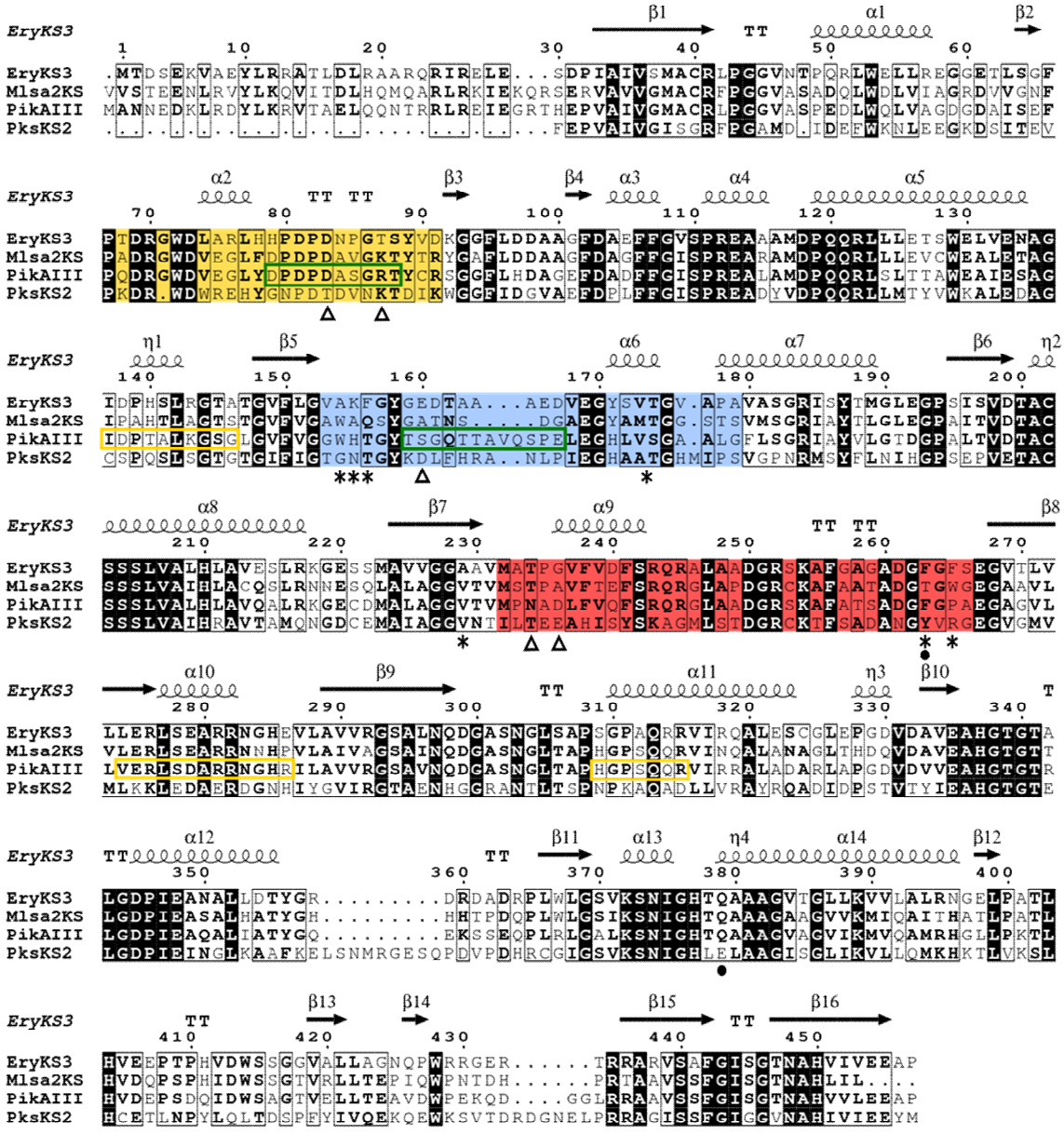

Figure S2. EryKS3AT3 crystal structure and Phyre model comparison Active site residues chosen for mutagenesis in the ketosynthase (KS) domain of Ery PKS (DEBS) module 3 (PDB accession number 2QO3) based on comparison with a Phyre2[2] model of a representative mycolactone PKS KS (MlsA2). Conserved active site residues are shown in yellow (EryKS) or orange (MlsA2) while those that differ are shown in blue (EryKS) or green (MlsA2). Residue Val173' lies at the KS-AT dimer interface and is contributed by the other subunit. The catalytic triad is shown in red and electron density observed for the ketosynthase inhibitor cerulenin bound to the catalytic Cys is shown in purple.

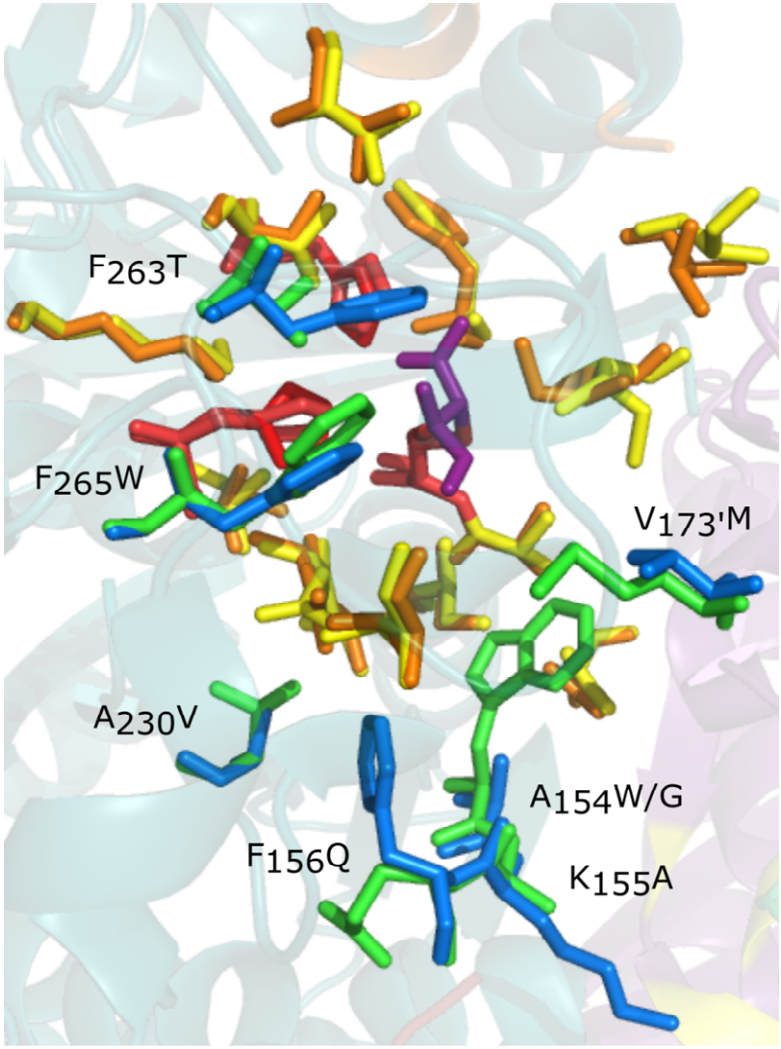

Figure S3. Sequence variability in *cis*-AT PKS domains. The percentage of sequences from 199 aligned PKS domains that have the consensus residue at each position is plotted against residue number. The regions shown in colour correspond to "variable surface loops" identified in X-ray crystal structures of both *trans*-AT and *cis*-AT KS domains.[25,28-31] Yellow = outer clasping loop; blue = dimer interface loop; and red = active site cap.

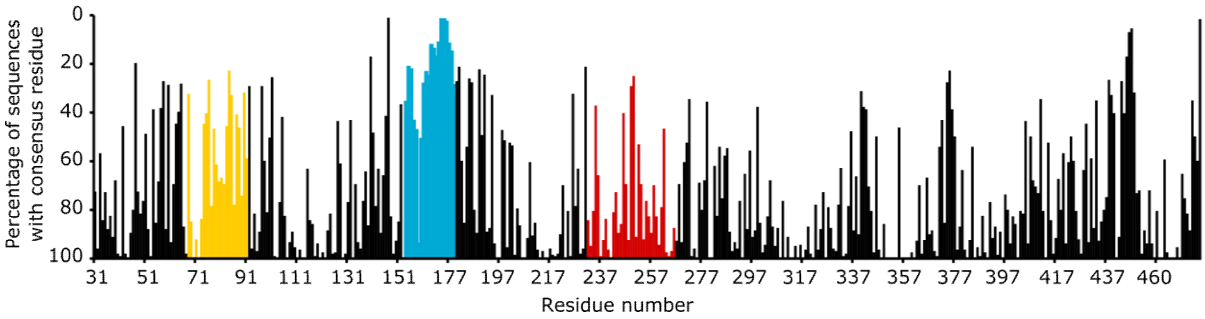

Figure S4. SDS-PAGE analysis of proteins (4-12% Bis-Tris)

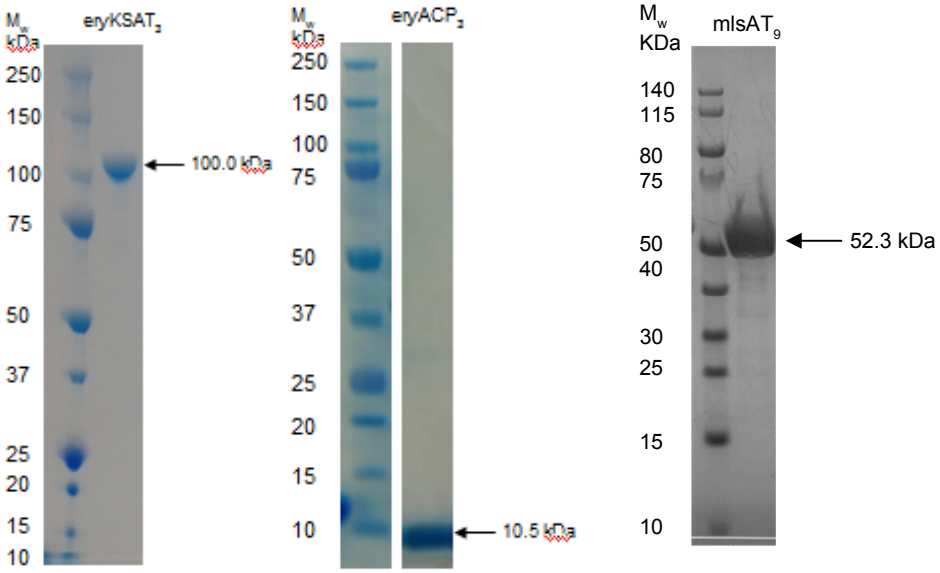

Table S3 Summary of EryACP3 condensation products

Substrate	R_t (mins)	Methyl malonyl (Da)	Malonyl (Da)
3	18.5	10604	10590
4	19.0	10618	10604
5	21.0	10660	10646
6	22.8	10688	10674
7	18.8	10618	10604
8	19.4	10644	10630
9	18.0	10602	10588
10	18.4	10618	10604
11	18.6	10616	10602
12	19.1	10630	10616
13	19.1	10630	10616
14	18.7	10628	10614
18-**20**	17.8	10646	10632

Figure S5. Selected MS analyses of KS-catalysed production of β-ketoacyl-EryACP3. EryACP3 after incubation with methylmalonyl-CoA (MeMal-CoA), substrate **6** and (A) wild type EryKS3AT3 or (B) Ala154Trp; ACP3 after incubation with MeMal-CoA, substrate **11** and (C) wild type EryKS3AT3 or (D) Ala154Trp; ACP3 after incubation with MeMal-CoA, substrate **13** and (E) wild type EryKS3AT3 or (F) Ala154Trp; ACP3 after incubation with MeMal-CoA, substrate **14** and (G) EryKS3AT3 or (H) Ala154Trp. Claisen condensation products for the respective substrates are shown in green, while side products (non-enzyme catalysed) arising from thioester exchange of ACP with SNAc substrate and conjugate addition of ACP onto α,β-unsaturated thioester substrates are shown in purple and orange respectively. An alkylated ACP used as an internal standard is shown in yellow.

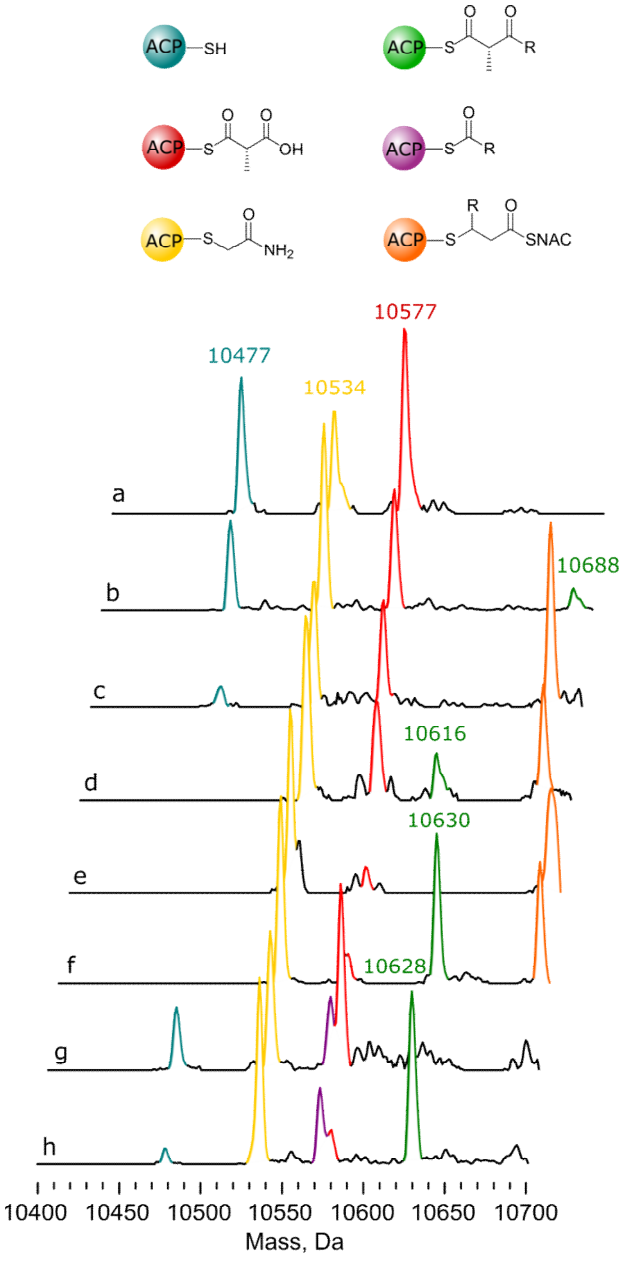

Figure S6. Time course of condensation of **4** with methylmalonyl-ACP, catalysed by wild type EryKS3AT3 and mutant Ala154Trp.

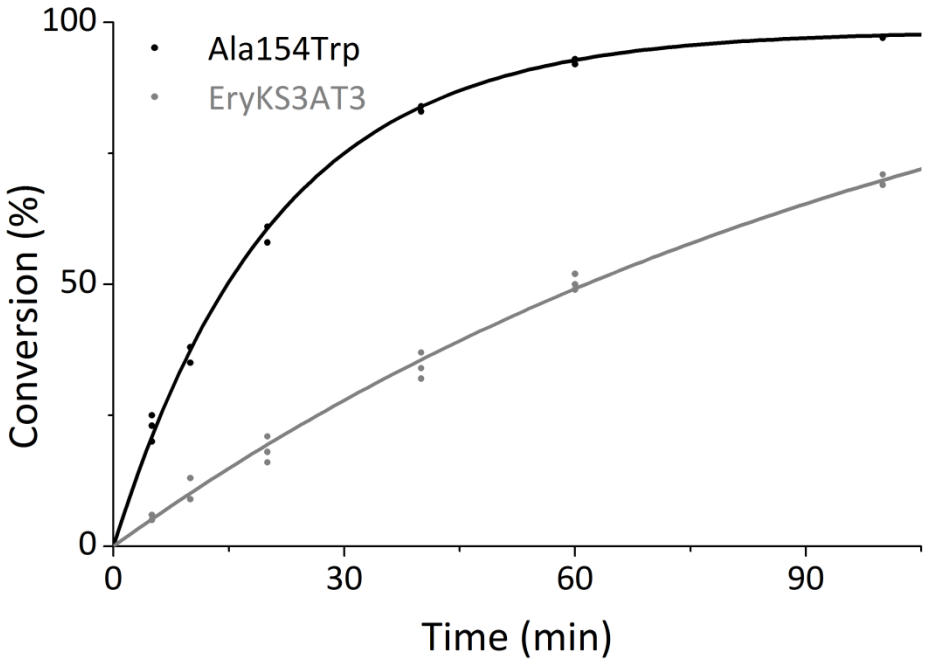

Figure S7. Malonylation of EryACP3 with MlsAT9.

Deconvoluted traces of mass spectrometric analysis of the incubation of malonyl coenzyme with EryACP3 in (a) the absence of an acyltransferase; in the presence of (b) MlsAT9; (c) MlsAT9 and EryKS3AT3; and (d) EryKS3AT3.

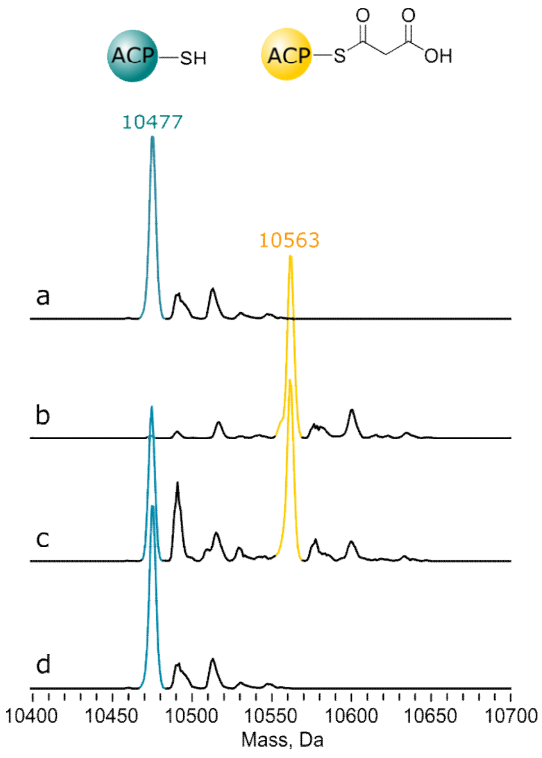

Table S4. Condensation assays using malonyl-CoA

Percentage of EryACP3 bearing Claisen condensation product after incubation with EryKS3AT3 or its mutants, malonyl CoA and synthetic SNAc substrate.

Substrate	EryKS3AT3	$A_{154}G$	$A_{154}W$	$K_{155}A$	$F_{156}Q$	$V_{173}M$	$A_{230}T$	$F_{263}T$	$F_{265}W$
3	3	T	6	5	3	3	2	T	
4	5	2	11	8	5		2	T	
5	T		6	2	2	T	T		
6			T						
7			4	2	T	T	T		
8	3		9	3	T	5	T		
9			T						
10	11	2	16	9	3	6	4		
11			3						
12			T						
13			5						
14			5	T	3				
15	T		T	4	T	T	T		
16			T						
17			T	T	T				
18									
19									

Figure S8. Dimer interface loop location

The location of the EryKS3AT3 dimer interface loop (blue sticks) including the A154 residue mutated during this study (orange) for one KS subunit is shown alongside the catalytic triad (red sticks). The surface of this subunit is shown in blue for the KS and red for the AT and post-AT linker, while the surface of the second subunit is shown in yellow.

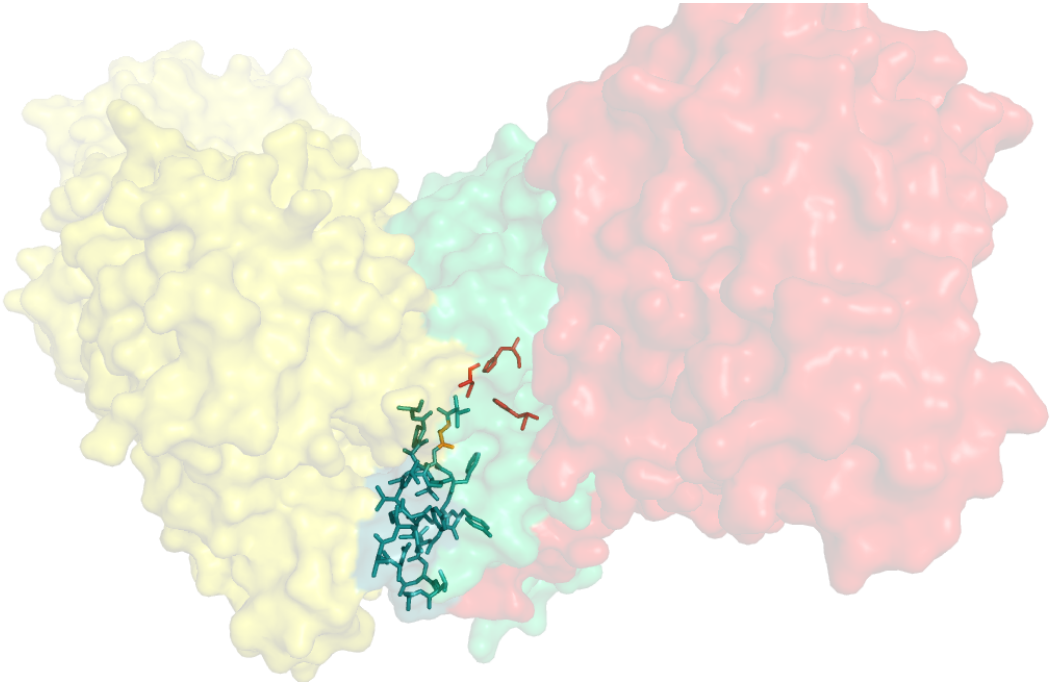

Figure S9. Sequence alignment of PKS type I *cis*-AT ketosynthases

199 sequences were chosen to represent the various type I KS clades identified by Jenke-Kodama et al., and were aligned using Clustal Omega (EMBL-EBI).[23] All non-loading domain ketosynthases from the following clusters were used: mycolactone, erythromycin, amphotericin, avermectin, borrelidin, epothilone, lasalocid, rapamycin, nyastatin, spinosyn, tylactone, niddamycin, monensin, oligomycin, rifamycin, pikromycin and stambomycin.

```
mycol_001_KS_001.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_001_KS_002.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_001_KS_003.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_001_KS_004.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_001_KS_005.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_001_KS_006.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_001_KS_007.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVAGLFDPD-PDAVGKTY
mycol_001_KS_008.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_002_KS_001.seq      ERVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_003_KS_001.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_003_KS_002.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_003_KS_003.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_003_KS_004.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_003_KS_005.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_003_KS_006.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
mycol_003_KS_007.seq      EPVAVVGMACRFPGGVASADQLWDLVIAGRDVVGNFPADRGWDVEGLFDPD-PDAVGKTY
eryth_001_KS_001.seq      EPVAVVAMACRLPGGVSTPEEFWELLSEGRDAVAGLPTDRGWDLDSLFHPD-PTRSGTAH
eryth_001_KS_002.seq      EPIAIVGMACRLPGEVDSPERLWELITSGRDSAAEVPDDRGWVPDELMASD-AA----GT
eryth_002_KS_001.seq      DPIAIVSMACRLPGGVNTPQRLWELLREGGETLSGFPTDRGWDLARLHHPD-PDNPGTSY
eryth_002_KS_002.seq      EPIAIVGIGCRFPGGIGSPEQLWRVLAEGANLTTGFPADRGWDIGRLYHPD-PDNPGTSY
eryth_003_KS_001.seq      EPIAIVGMACRFPGDVDSPESFWEFVSGGGDAIAEAPADRGWEPDP-------------D
eryth_003_KS_002.seq      DPIAIVGMACRFPGGVHNPGELWEFIVGRGDAVTEMPTDRGWDLDALFDPD-PQRHGTSY
ampho_002_KS_001.seq      EPVAIVGMSCRLPGGVSSPESLWELVESGGDAISGFPADRGWDIEALTGKG-EG--SSST
ampho_002_KS_002.seq      --IVIVGMSCRFPGGANTPEELLRLALDGADVISEFPADRGWDAHGLYDPD-PDRQGRTY
ampho_003_KS_001.seq      EPIAVVGIGCRFPGGVVSPEDLWELVAEGRDVIGPFPTDRGWDLDRLAGGG-E---GSSL
ampho_003_KS_002.seq      DPIVIVGMACRYPGGVTSPEDLWRLVSEGADATGPFPVNRGWDVDNLYDPD-PDRPGHTY
ampho_003_KS_003.seq      DPIVIVGMACRFPGDVTSPEDLWRLVSEGTDAISDFPANRGWDVDSLYHPD-PEHSGTSY
ampho_003_KS_004.seq      DPIVIVGMACRFPGGITSPEDLWRLLGEGSDVISDFPVNRGWDVDSLYDPD-PAHTGTSY
ampho_003_KS_005.seq      DPIVIVGMACRYPGGVASPEDLWRLVTEGTDAVSAFPVNRGWDVENLYHPD-PDNIGTSY
ampho_003_KS_006.seq      DPIVIVGMACRYPGGVASPEDLWRLLLEGTDAITEFPVNRGWDIESVYHPD-PEHLGTSY
ampho_004_KS_001.seq      EPIAIVGMTCRYPGGVSSPEDLWRMVEAGENGVTPFPTDRGWDLEALASS---------P
ampho_004_KS_002.seq      DPIVIVGMACRYPGGVTSPEDLWRLVTEGTDAVSGFPVNRGWDVENLYHPD-PDHPGTAY
ampho_004_KS_003.seq      EPIAIIGMSCRYPGGVSSPEQLWDLVLSGTDAITDFPVNRGWNTAGLYDPD-PDHPGTTY
ampho_004_KS_004.seq      DPIAIVGMSCRYPGGITSAEQLWRVSLEEVDAVSVFPADRGWDAEALYDPD-PDASGRTY
ampho_004_KS_005.seq      DPIAVVGLACRFPGGVSTPEELWQLLAEGRDGITGFPDDRNWDIAALGAGA-------SD
ampho_004_KS_006.seq      DPIVIVGMACRFPGGVNSPEDLWQLVLDEVDAVGDFPADRGWDLDALAGDG-PG--RSAT
ampho_005_KS_001.seq      EPIAVVGMGCRFPGGVDSPQALWEMVAGGTDVISEFPDDRGWDLEALRTSG-IDDRDTSV
ampho_005_KS_002.seq      DPVVIVGMGCRFPGGVGSPEDLWDLLASGSDAITGFPADREWESSRLVAGE--A--GGVS
ampho_005_KS_003.seq      DPVVIVGMGCRFPGGVGSPEELWDLVASGTDAITGFPADREWESST-IGGE--P--GDLS
ampho_006_KS_001.seq      EPVAIVGMACRFPGGVRSPEDLWEMLADGRDAISGFPADRGWDLETLAGDG-AG--GSST
averm_001_KS_001.seq      DPIAIIGMACRFPGGVRSPKDLWELAASGGDAIGPFPTDRGWPTEQRHAQD-PTQPGTFY
averm_001_KS_002.seq      EPIAIVGMACRFPGGVTSADDFWDLISSEQDAIGGFPTDRGWDLDTLYDPD-PDHPGTCY
averm_002_KS_001.seq      EPIAIVGMACRYPGGATSPTRLWHLVKSQTDAIGEFPTDRGWNLEQLYDPD-PDRSGTSY
averm_002_KS_002.seq      EPIAIVGMACRYPGDVRTVDDLWQVVSGGHDAIGGFPTNRGWDLDTLYNPD-PDHHGTSY
averm_002_KS_003.seq      EPIAIVGMACRFPGGVTSADDFWDLISSEQDAIGGFPTDRGWDLDTLYDPD-PDHPGTCY
averm_002_KS_004.seq      EPIAIIGMACRFPGGVRSADDLWELLASGKDAIGVFPTDRGWDLDTLYDPD-PDHPGTCY
averm_003_KS_001.seq      EPIAIVGMACRFPGGVCSPEELWELVASGGDAIGEFPAGRGWDLEGLFDSD-PDRSGTSY
averm_003_KS_002.seq      EPVAIIGMACRYPGGVTSAEELWELLASGRDTVGEFPTDRGWDLEALFDPE-PGRPGTSY
averm_003_KS_003.seq      EPIAIVGMACRFPGGVCSPEELWELVASGGDAIGEFPAGRGWDLEGLFDSD-PDRSGTSY
averm_004_KS_001.seq      EPIAIVGMACRFPGGVESAEDFWELIASGRDAVGEFPVDRGWDVEAFYDPE-PGRAGSSY
averm_004_KS_002.seq      EPIAIVGMACRFPGGVESAEDLWELVASGRDAVGEFPVDRGWDVEAFYDPE-PGRAGSSY
averm_004_KS_003.seq      EPIAIVGMACRFPGGVESAEDFWELIASGRDAVGEFPVDRGWDVEAFYDPE-PGRAGTSY
borre_002_KS_001.seq      EPIAIVGMSCRLPGGVNSPGEFWSLLEAGTDAVSEFPRDRGWDVENLYDPD-PDAPGRSY
borre_003_KS_001.seq      EPIAIVGMSCRLPGGVNSPEEFWSLLEAGTDAVSEFPRDRGWDVERLYDPD-PDAPGKSY
borre_003_KS_002.seq      DAIVIVGMAGRFPGGVRTPEGLWDLVHSGTDAISEWPTDRGWDVENLYDPD-PDAVGKSY
borre_004_KS_001.seq      DPIAIIGMGCRLPGGVRTPEDLWDLVEKKHDAIGPFPADRGWDLENLYDPD-PDAPGKAY
borre_005_KS_001.seq      EPIAIVGMGCRYPGGVRSPEEFWNLLDEGVDAVAGFPEDRGWDLENLYDPD-PDEPGKCY
borre_006_KS_001.seq      EPVAIVGMSCRLPGDVRTPERFWELLDTGTDALTPLPTDRGWNLDTAFDDE------RPY
epoth_002_KS_001.seq      --IAVIGMSGRFPGA-RDLDEFWRNLRDGTEAVQRFSEQE-LAA-SGVDPA--LVLDPNY
epoth_003_KS_001.seq      EPIAIVGIGCRFPGGADAPEAFWELLDAERDAVQP-LDRR-WALV-GVAPV-EAVPH---
```

```
epoth_003_KS_002.seq      DPIAIVGAACRFPGGVEDLESYWQLLTEGVVVSTEVPADR-WNGADGRGPGSGEAPRQTY
epoth_003_KS_003.seq      EPIAIVGAACRFPGGVEDLESYWQLLAEGVVVSAEVPADR-WDAADWYDPD-PEIPGRTY
epoth_003_KS_004.seq      EPIAVIGIGCRFPGGVTDPESFWRLLEEGSDAVVEVPHER-WDIDAFYDPD-PDVRGKMT
epoth_004_KS_001.seq      EPIAIVGIGCRFPGGAGTPEAFWELLDDGRDAIRP-LEER-WALV-GVDPG-DDVPR---
epoth_004_KS_002.seq      DDIAIVGAACRFPGGDEGLETYWRHLAEGMVVSTEVPADR-WRAADWYDPD-PEVPGRTY
epoth_005_KS_001.seq      EPIAIIGIGCRFPGGADTPEAFWELLDSGRDAVQP-LDRR-WALV-GVHPS-EEVPR---
lasal_001_KS_002.seq      DPVAIVAMACRFPGGVASPDDLWRLVAEGGDAISGFPTDRGWDMDALYDPD-PTRTGKNY
lasal_001_KS_003.seq      EPIAIVGMACRYPGDVRSPEDLWQMVAEGRDVIAGFPADRGWDLEGLYDPQ-SARPGTSY
lasal_002_KS_001.seq      EPIAIVGMACRFPGGADTPERLWELVAEGRDAITGFPTDRGWDLDALFDPD-PDRSGTSY
lasal_002_KS_002.seq      EPVAIVGMACRFPGGVRSPEDLWRLVVDGGDAVSDFPADRGWDLDALFDPD-PDHPGTSY
lasal_002_KS_003.seq      EPLAIVSMACRFPGGIESPDDLWQYVVQGRDAVSDFPVDRGWDLERLYDPD-PERSGTSY
lasal_003_KS_001.seq      EPIAIVGMGCRYPGDVRSADDLWGLLAEGGDAISKFPDDRGWNTEELYDPD-PDRPGRNY
lasal_004_KS_001.seq      EPIAIIGMACRLPGGVDSPEGLWELVDSGTDAIAGFPLDRGWDVEGMYDPD-AEAPGKTY
lasal_005_KS_001.seq      EPVAIVGMGCRFPGDVSSPEELWDLVAEGRDAVSSFPADRGWDLDALRDAD-PDRPGTSY
lasal_005_KS_002.seq      EPLAIVSMACRFPGGVRSPEDLWRVVADGLDVVSGFPDDRGWDLSRLYDPD-PGQAGTSY
lasal_006_KS_001.seq      EPIAIVGMACRFPGGADSPDTFWELIAEGRDAISGMPTDRGWDLDALYDDD-PSHAGTSY
lasal_007_KS_001.seq      EPVAIVGMACRFPGGVESPEQLWRLLTEGRDAVADFPSDRGWDLDSLYDPD-PDHPGTSY
rapam_001_KS_001.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAISGFPTDRGWDVDGLFDPD-PDASGKSY
rapam_001_KS_002.seq      EPLAIVGMACRLPGGVSSPEDLWRLLESGTDAVSGFPTDRGWDVENLFGPA----VGNSY
rapam_001_KS_003.seq      EPLAIVGMACRLPGGVSSPEDLWRLLESGTDAVSGFPTDRGWDVENLYDM-----AGKSH
rapam_001_KS_004.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAVSGFPTDRGWDVENLYDSD-PEAAGKSY
rapam_002_KS_001.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAISGFPADRGWDVENLYDPD-PEAAGKSY
rapam_002_KS_002.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAISGFPADRGWDAESLFDPD-PDAVGKSY
rapam_002_KS_003.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAVSGFPTDRGWDVEGLFDPD-PDAAGKSY
rapam_002_KS_004.seq      EPLAIVGMACRLPGGVSSPEDLWRLLESGTDAVSGFPTDRGWDVENLFGPA----AGDSY
rapam_002_KS_005.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAISGFPTDRGWDVENLFDSD-PDAAGKSY
rapam_002_KS_006.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAISGFPTDRGWDVENLYDPD-PDAPGKSY
rapam_003_KS_001.seq      EPLAIVGMACRLPGGVSSPEDLWRLVESGTDAISDFPADRGWDVENLYDPD-PDASGKSY
rapam_003_KS_002.seq      EPLAIVGMACRLPGGVSSPEDLWRLLESGTDAVSGFPTDRGWDVENLYDM-----AGKSH
rapam_003_KS_003.seq      EPLAVVGMACRLPGGVSSPEDLWRLVESGTDAISGFPADRGWDAESLFDPD-PDASGKSY
rapam_003_KS_004.seq      EPIAIVAMACRLPGGVSSPEGLWHLVESGTDAISGFPTDRGWDVEGLFDPD-PDAAGKSY
nysta_002_KS_001.seq      EPIAIVGMGCRLPGGVNSPESLWDLVRSGGDAISGFPVDRGWDLETLTGNG-DG--SSAT
nysta_002_KS_002.seq      DPIAVIGMSCRYPGGAATPEELLRLALDGADVISEFPADRGWDARGLYDPD-PDRPGHTY
nysta_003_KS_001.seq      EPVAVVGIGCRFPGGVTSPEDLWELVAEGRDVIGPFPQDRGWDLEKLAGGG-E---GGSL
nysta_003_KS_002.seq      DPIVIVGMACRYPGGIASPEDLWRLVSQGADATGPFPTNRGWDLDNLYDPD-PDRPGRTH
nysta_003_KS_003.seq      DPIVVVGMACRFPGGVTSPEDLWRLVDDGTDAITTFPTNRGWDLDNLYDPD-PEHFGTSY
nysta_003_KS_004.seq      DPIVIIGMSCRFPGDIGSPEDLWRLLGDGADVVTDFPTNRGWDLDNLYDPD-PAHAGTSY
nysta_003_KS_005.seq      DPIVIVGMACRYPGGVASPDDLWRLVTDGVDAVSPFPTNRGWDVESLYHPD-PDHLGTSY
nysta_003_KS_006.seq      DPIVVVGMACRFPGGVTSPEDLWRLVDAGTDAITTFPTNRGWDLESLYDPD-PAHLGTSY
nysta_004_KS_001.seq      EPVAIIGMTCRYPGGVRSPEDLWRMVEAGEHGVTPFPTDRGWDLEALAAA---------P
nysta_004_KS_002.seq      DPIVIVGMACRYPGGVSSPEDLWRVLTEGTDAVSGFPTNRGWDVESLYHPD-PDHPGTSY
nysta_004_KS_003.seq      EPIAIIGMACRFPGGVSSPEQLWDLVASGTDAISEFPVNRGWQTGHLFDPD-PDRPGTTY
nysta_004_KS_004.seq      DPIAIVGMSCRYPGGVGSAEDLWRIALDEVDAISGFPADRGWDAEGLYDPD-PDRPGRTY
nysta_004_KS_005.seq      DPIAVIGIGCRFPGGVTTPEELWQLLDEGRDGISRFPDDRGWDLAALGAGA-------SD
nysta_004_KS_006.seq      DAIVIVGMACRYPGGVTSPEDLWQLVGDEVDAVGDFPTDRGWDLAALAGDG-PG--RSAT
nysta_005_KS_001.seq      EPIAVVGMACRFPGAVDSPEALWEMVATGTDVISGFPDDRGWDLEALRNSG-TDARDTDV
nysta_005_KS_002.seq      DPVVIVGMSCRFPGGVSSPEELWDLVASGTDAITGFPADRAWDRHPQLAGA--P--GART
nysta_005_KS_003.seq      DPIVVVGMACRFPGGVTTPEEFWQLLAEGRDGIDAFPTDRGWDLDVLGRRR-PG--PQRP
nysta_006_KS_001.seq      EPVAIVGMACRFPGGVRSPEDLWELLSAGRDGIGPFPADRGWDLAALAGDG-PG--RSAT
spino_002_KS_001.seq      DPIAIVAMSCRYPGQVSSPEDLWQLAAGGVDAISEVPGDRGWDLDGVFVPD-SDRPGTSY
spino_003_KS_001.seq      EPIAIVAMGCRYPGGVQDPEGLWKLVASGGDAIGEFPADRGWHLDELYDPD-PDQPGTCY
spino_003_KS_002.seq      EPIAIVAMSCRFPGGAHSPEDLWRLVAAGTEVIGEFPSDRGWDAEGLYDPD-ASRPGTTY
spino_004_KS_001.seq      EPIAIVSASCRLPGGVDSPEALWQLVRTGTDAISEFPADRGWDLGRLYDPD-PNHQGTSY
spino_004_KS_002.seq      --VAIVGMGCRFPGGVSSPEELWRLVAGGVDAVAGFPDDRGWDLAALYDPD-PDRLGTSY
spino_004_KS_003.seq      DPVVVVAMGCRYPGDVCSPEELWQLVSAGRDAVSTFPVDRGWDCNTLFDPD-PDRAGSTY
spino_005_KS_001.seq      DPIAIVSMGCRYPGGVSTPEELWRLVVDGGDAIANFPEDRGWNLDELFDPD-PGRAGTSY
spino_005_KS_002.seq      DPVVIVGMGCRYPGGVSSPEELWRLVAGGLDAVAEFPDDRGWDQAGLFDPD-PDRLGTSY
spino_005_KS_003.seq      DLVAIVGMGCRFPGQVSSPEELWRLVAGGVDAVADFPADRGWDLAGLFDPD-PERAGKTY
tylac_001_KS_002.seq      EPIAVVGMGCRFPGGVGSPEALWRLVVEGVDAVSPFPGDRGWDVEGLYDPE-PGVAGKSY
tylac_001_KS_003.seq      DPVVIVGMACRFPGGAGTPEALWKLVTEERDVIGAAPTDRGWDLDSVYDPE-PGVAGKTY
tylac_002_KS_001.seq      EPIAVVAMSCRYPGGVTTPEDLWRLLADERDAVSGLPRDRGWDLDALYDPD-G-GPGTSY
tylac_003_KS_001.seq      EPIAVVGMGCRFPGGVGSPEALWRLVVEGVDAVSPFPGDRGWDVEGLYDPE-PGVAGKSY
tylac_003_KS_002.seq      EPIAIVSMACRMPGGVSTPEELWRLLRDGKDAIGPFPANRGWDLENLYDPD-PDADGRTY
tylac_004_KS_001.seq      EPIAVVGMACRYPGGANTPEQFWELLDTGTDAAAPMPSDRGWDTHGLYDPD-PAAAGRTY
tylac_005_KS_001.seq      EPIAVVGMACRYPGGVTGPEELWELVAGGRDAIGPFPVDRGWDVASVYDPD-PESKGTTY
nidda_001_KS_002.seq      DPIAVIGIGCRFPGGVATPDDLWELVASGTDAISTFPTDRGWDLDGLYDPD-PSTPGKSY
nidda_001_KS_003.seq      EPVAVVGMACRYPGGVATPDDLWELVASGTDAISTFPTDRGWDLDGLYDPD-PSTPGKSY
nidda_002_KS_001.seq      EPVAVVGMACRYPGGVATPDDLWELVASGTDAISTFPTDRGWDLDGLYDPD-PSTPGKSY
nidda_003_KS_001.seq      EPVAVVGMACRYPGGVATPDDLWELVASGTDAISTFPTDRGWDLDGLYDPD-PSTPGKSY
nidda_003_KS_002.seq      EPIAIIGMACRLPGGVRSPEDLWELLSDGADAITAFPTDRGWDNDALYDPD-PAASRRTY
nidda_004_KS_001.seq      EPIAIVGMACRLPGGVRGPEDLWELLRSETDAVSPAPSDRGWDVTAMYSPD-PDRPGTTY
nidda_005_KS_001.seq      EPIAIVGMACRYPGGVSDPEELWQLAAEGVDAITRFPTDRGWDEAAVYSPE-H-RPGTTY
monen_001_KS_002.seq      DPVAIVAMSCRYPGGAGTPEDLWRLVADGADAIGDFPTDRGWDLARLFHPD-PDRSGTSC
monen_002_KS_001.seq      EPVALVGMACRFPGGVASPEDLWDLVAAGKDAIEDFPTDRGWDLEALYDPD-PAAYGTSY
monen_003_KS_001.seq      EPVAVVEAACRYPGGIRTPEDLWDLVAAGGNALGAFPDNRGWDLRRLFHPD-PDHPGTTY
monen_003_KS_002.seq      DPIAIVGMACRYPGGVTSPEQLWQLVATGTDAIGPFPEDRGWDTAGLFDPD-PDQVGHSY
monen_004_KS_001.seq      EPIAIVSMACRFPGDTRTPEALWDLVAEGGDAIDDFPTNRGWDLESLYHPD-PDHPGTSY
```

```
monen_004_KS_002.seq      DPVAIVAMTCRFPGGVASPDDLWDLLDARKDAMGAFPTDRGWDLERLFHPD-PDHPGTSY
monen_005_KS_001.seq      EPVAVVGMACRFPGGIETPEGLWELVAAGDDAIEPFPTDRGWDLEGIYHPD-PDHPGTCY
monen_005_KS_002.seq      EPVAIVAMACKYPGGATSPEGLWDLVAEGVDAVGAFPTGRGWDLERLFHPD-PDHPGTSY
monen_006_KS_001.seq      EPIAIVAMACRYPGGVASPEQLWDLVASRGDAIEEFPADRGWDVAGLYHPD-PDHPGTTY
monen_007_KS_001.seq      EPVAVIGMACRLPGGVRSPEEFWELVSTGGDAVAPLPGNRNWDLDSLYDPD-PESTGTSY
monen_008_KS_001.seq      EPVAIVGMACRYPGGARSAEDLWELVRDGGDAVAGFPDDRGWDLESLYHPD-PEHPATSY
monen_008_KS_002.seq      EPIAIVGMACRYPGDVRSAEELWDLVAAGKDAMGAFPDDRGWDLETLYDPD-PESRGTSY
oligo_001_KS_002.seq      DPIVIVGMGCRLPGGVRSPEDLWRFVAAGGDAISALPVDRGWEFAA------------DF
oligo_001_KS_003.seq      DPVVIVSAACRFPGGVTTPEELWQLVLDGGDAIGPFPADRGWDLDRLYDPD-PGNPGTSY
oligo_001_KS_004.seq      DPVVIVAMACRFPGEATSPEKLWDLIAAGKDGIGDFPADRGWSVAADAAFS---------
oligo_002_KS_001.seq      EPIAVVSMGCRFPGGIASPEDLWECVRTGTDTISGFPTDRGWRTGHLRG---------DF
oligo_002_KS_002.seq      DPIAIVGMACRYPGGVRGPEDLWDLVHCGRDGVGDFPTDRGWDLPALRRLA-PD-----L
oligo_003_KS_001.seq      EPIAIVGMACRYPGGVSSPEDLWRLVTDETDAIGPFPTDRGWDLDGLHHPD-PDHPGTCY
oligo_003_KS_002.seq      EPLVIVGMACRYPGGITSPEQLWDLVIGERDAVSGFPEDRGWSTGAPADG--------SV
oligo_004_KS_001.seq      EPIAIVGMACRLPGGVGSPDQLWDLVRRGEDAVTGMPTDRGWELGELYDPD-PDHPGTSY
oligo_004_KS_002.seq      EPIAIVGMACRYPGGVASPEELWELVAAGRDAIGDMPEDRGWNVEELYDPE-LARPGTSY
oligo_005_KS_001.seq      EPIAIVGMACRFPGGAQTPEDLWHMVSEGVDAISGFPDDRGWDLDRLYRTD-AEESGAST
oligo_006_KS_001.seq      EPLAIVGIGCRYPGGVRDTEDLWRLITEGRDAMSGFPTDRGWDRWDIPT-----------
oligo_006_KS_002.seq      EPLAIVGMACRYPGGVRDADDLWRLIVEGRDELSDFPTDRGWDRWGAPA-----------
oligo_006_KS_003.seq      EPLAIVGMACRYSGGVSSPEDLWQVVVENKDGMSDFPSYRGWDRWNLAS-----------
oligo_007_KS_001.seq      EPLAVVGMACRLPGGVDSPEQLWQMVMDGADGVSGFPDDRGWDLSSLFDSD-PDRPGTTY
oligo_007_KS_002.seq      DPIVIVGMACRFPGGVRGPEDLWRLLVDSRDEMTDFPADRGWHGLAMNAFI-EE---SGG
rifam_001_KS_001.seq      EPIAIVGMACRLPGGVASPEDLWRLVAERVDAVSEFPGDRGWDLDSLIDPD-RERAGTSY
rifam_001_KS_002.seq      EPIAIVAMACRFPGGVHSPEDLWRLVADGADAVTEFPADRGWDTDRLYHED-PDHEGTTY
rifam_001_KS_003.seq      EPIAIVGMSTRLPGGADSPEELWKLVAEGRDAVSGFPVDRGWDLDGLYHPD-PAHAGTSY
rifam_002_KS_001.seq      EPVAIVSMACRYAGGIRGPEDFWRVVSEGADVYTGFPEDRGWDVEGLYHPD-PDNPGTTY
rifam_002_KS_002.seq      EPIAIVGMACRLPGGVMSPDDLWRMVAEGRDGMSPFPGDRGWDLDGLFDSD-PERPGTAY
rifam_002_KS_003.seq      EPIAIVGMACRLPGGVAGPEDLWRLVAEGRDAVSSFPTDRGWDTDSLYDPD-PARPGKTY
rifam_003_KS_001.seq      EPIAIVGMACRLPGGVTGPGDLWRLVAEGGDAVSGFPTDRCWDLDTLFDPD-PDHAGTSY
rifam_004_KS_001.seq      EPIAIVAMACRYPGGVSSPEDLWRLVAEGTDAVSAFPGDRGWDVDGLVDPD-PDRPGTTY
rifam_005_KS_001.seq      EPLAVVGMACRFPGGVSSPEDLWQLVAGGVDALSDFPDDRGWELDGLFDPD-PDHPGTSY
rifam_005_KS_002.seq      EPIAIVGMACRLPGGVTDPEGLWRLVRDGLEGLSPFPEDRGWDLENLFDDD-PDRSGTTY
pikro_001_KS_002.seq      EPVAIVGMACRLPGGVASPEDLWRLVAGGGDAISEFPQDRGWDVEGLYHPD-PEHPGTSY
pikro_001_KS_003.seq      DPIAIVAMSCRYPGDIRSPEDLWRMLSEGGEGITPFPTDRGWDLDGLYDAD-PDALGRAY
pikro_002_KS_001.seq      EPVAIVGMACRLPGGVASPEDLWRLVAGGEDAISEFPQDRGWDVEGLYDPN-PEATGKSY
pikro_002_KS_002.seq      EPVAIVGMACRLPGGVASPEDLWRLVAGGEDAISGFPQDRGWDVEGLYDPD-PDASGRTY
pikro_003_KS_001.seq      EPVAIVGMACRLPGGVASPEDLWQLVAGDGDAISEFPQDRGWDVEGLYDPD-PDASGRTY
pikro_004_KS_001.seq      EPMAIVGMSCRFAGGIRSPEDLWDAVAAGKDLVSEVPEERGWDIDSLYDPV-PGRKGTTY
Stamb_001_KS_002.seq      DPVVIVSMACRFPGGIDSPAALWEALSEGREVLTGLPEDRGWQPALAALE-------TPL
Stamb_001_KS_003.seq      EPIAVVAMSCRYPGGVDGPEELWRMVREGRDGITGFPTDRGWDLARLHDPD-PSRPGTTY
Stamb_001_KS_004.seq      DPVAVVAMGCRFPGA-ATPEEFWDLLSRGLDLVGDFPADRGWDLRDAPD----------F
Stamb_002_KS_001.seq      EPIAIVSMACRFPGGVRTPEQLWHLVTEGRDAVGPFPADRGWDLDALLDPD-PERPGGSY
Stamb_002_KS_002.seq      EPIAIIGMSCRLPGGVRSPADLWRLVHAAEDAITAFPTDRGWDIDTTDL---------PY
Stamb_003_KS_001.seq      EPIAIVAMSCRLPGGVGAPEDLWRLLADGTDAIVDFPDDRGWTQTVAET--------GGY
Stamb_003_KS_002.seq      EPIAIVGMSCRYPGGANSAEQLWDLVASGRDVVTDFPTDRGWPEG-GDG---------DY
Stamb_003_KS_003.seq      EPIAIVGMACRLPGGVASPADLWPLVSEGMDATSGFPADRGWPLDIVG----------AE
Stamb_003_KS_004.seq      EPIAIVGMSARYPGGVRNAEQLWELVAHGADGTAEFPADREWPSDVVAG----------A
Stamb_003_KS_005.seq      EPLALVGMACRLPGGVGSPDELWDLLIGGVDAMTTFPTDRGWDLTALPDSP---------
Stamb_004_KS_001.seq      EPIAIIGMSCRYPGDVHTPEDLWRLVAGEQDAISGFPEDRGWDLARLFDGD-PDGEGHSS
Stamb_004_KS_002.seq      DPIAIVGMACRFPGGADTPEALWRLVAEETDAVGDFPTDRAWDLAHIYGES-EKADD-RR
Stamb_004_KS_003.seq      EPLAVIGMSCRYAGGVDSPEDLWRLVLDGTDAMGPFPSDRGWDVENLYHPD-PDHPGTSY
Stamb_005_KS_001.seq      EPIAVIGMACRYPGGVRSPEDLWKLVESGTDALGPLPTDRGWDPARTSA-----------
Stamb_006_KS_001.seq      EPIAIVGMACRLPGGVDSPASLWDLVSAGGDGISEFPADRGWDLGSLYDPH-SRGEGTSY
Stamb_006_KS_002.seq      EPVAIIGMACRYPGGVTSPDELWQLVAEGRDGVTDFPEDRGWDTARLYEPD-G-TPGTSY
Stamb_007_KS_001.seq      EPVAVVGMGCRYPGGVHSPEDLWHLVSEGRDGISGFPLDRGWDT-AVYDDD-PDAAGTTY
Stamb_007_KS_002.seq      EPLAIVGMACRYPGGASSPEELWRLVTEGIDAVGGFPSDRGWPTDDDTG------PGTSY
Stamb_008_KS_001.seq      EPMAVVSVACTLPGGVRSPEDLWRLLASGTDAVTPFPGDRGWDDGPLVGAE-GV--AAPE
Stamb_008_KS_002.seq      EPIAIVAMSCRYPGEVNSPEDLWRLVATGSDAIGPFPDNRGWDTTALYDPD-PDAAGKSY
Stamb_009_KS_001.seq      EPIAIIGMSCRFPGGADSPEALWRLVRDGIDAMGPFPADRGWDMTHFETGV-A--ADVEF
Stamb_009_KS_002.seq      DPIVIVGMACRFPGGVRSPEDLWKLVESGTDAMGPFPADRGWDLDAL---------GGDY
Stamb_009_KS_003.seq      EPVAIIGMSCRFPGGVDSPEALWRLVRDGVDAMGDFPADRGWDLDTLLGHD-----GPSH
Stamb_009_KS_004.seq      DPVVIVSMSCRYPGGADSPENLWEVVAAGTDAMSPFPVDRGWDLPDA----------GDF
                            :.::       *                        .

mycol_001_KS_001.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_001_KS_002.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_001_KS_003.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_001_KS_004.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_001_KS_005.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_001_KS_006.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_001_KS_007.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_001_KS_008.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_002_KS_001.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_003_KS_001.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_003_KS_002.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_003_KS_003.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_003_KS_004.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
```

```
mycol_003_KS_005.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_003_KS_006.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
mycol_003_KS_007.seq      TRYGAFLDDAAGFDAGFFGISPREARAMDPQQRLLLEVCWEALETAGIPAHTLAGTS-TG
eryth_001_KS_001.seq      QRGGGFLTEATAFDPAFFGMSPREALAVDPQQRLMLELSWEVLERAGIPPTSLQASP-TG
eryth_001_KS_002.seq      RAHGNFMAGAGDFDAAFFGISPREALAMDPQQRQALETTWEALESAGIPPETLRGSD-TG
eryth_002_KS_001.seq      VDKGGFLDDAAGFDAEFFGVSPREAAAMDPQQRLLLETSWELVENAGIDPHSLRGTA-TG
eryth_002_KS_002.seq      VDKGGFLTDAADFDPGFFGITPREALAMDPQQRLMLETAWEAVERAGIDPDALRGTD-TG
eryth_003_KS_001.seq      ARLGGMLAAAGDFDAGFFGISPREALAMDPQQRIMLEISWEALERAGHDPVSLRGSA-TG
eryth_003_KS_002.seq      SRHGAFLDGAADFDAAFFGISPREALAMDPQQRQVLETTWELFENAGIDPHSLRGSD-TG
ampho_002_KS_001.seq      H-EGGFLYEATDFDAGFFGISPREATAMDPQQRLLLEVSWEALERAGIAPTSLRSSA-TG
ampho_002_KS_002.seq      SVHGGFLHEAAGFDAGFFGISPREALAMDPQQRLLLETSWEVLERAGIDPESLHSTA-TG
ampho_003_KS_001.seq      AQVGGFLDDAAGFDPGFFGISPREAVAMDPQQRILLEIAWEALERAGIDPSALRGSR-TG
ampho_003_KS_002.seq      VRAGGFLHDAGSFDADFFGMSPREALSTDSQQRLLLELSWEAVERAGIDPHSLRGSR-TG
ampho_003_KS_003.seq      TRSGGFLHDAGDFDPDFFGMSPREALATDSQQRLLLETSWEAIERAGIDPLSLRDSR-TG
ampho_003_KS_004.seq      TRSGGFLHNAADFDADFFGMSPREAMATDSQQRLLLESTWEAIERAGIDPVTLRGSR-TG
ampho_003_KS_005.seq      TRSGGFLHEAGEFDPGFFGMSPREALSTDSQQRLLLETSWEAIERAGIDPVGLRGSA-TG
ampho_003_KS_006.seq      TRSGGFLHEAGEFDPGFFGMSPREALATDSQQRLLLESSWEAIERAGIDPVGLRGSA-TG
ampho_004_KS_001.seq      TSRGGFLHDAPEFDADFFGISPREAVAMDPQQRVVLESAWEAFERAGINPTSVKGSR-TG
ampho_004_KS_002.seq      TRSGGFLHEAGEFDPGFFGMSPREALATDSQQRLLLEASWEAIERAGIDPVGLRGSA-TG
ampho_004_KS_003.seq      STQGGFLHEADEFDPMFFGISPREALVMDPQQRLLLETTWEAFERAGLTPDTLRGSL-TG
ampho_004_KS_004.seq      SVQGGFLRDVADFDPGFFGISPREALSMDPQQRLLLETAWEVFENAGLDPVAQRGSR-TG
ampho_004_KS_005.seq      TLHGGFLAQVADFDARFFGISPREALAMDPQQRLLLETTWEALERAGIDPAALRGST-TG
ampho_004_KS_006.seq      D-QGGFLYDATDFDPGLFGVSPREAMVMDPQQRILLEASWEALERAGIDPGTLRGSGTTG
ampho_005_KS_001.seq      SQRGGFLDSIADFDPGFFGISPREAVTMDPQQRLLLETAWEAIERARIDATRLRGTR-TG
ampho_005_KS_002.seq      A-QGGFLSDIAGFDADFFGISPREALAMDPQQRILLEVTWEAIERAGVDPTALRGSR-TG
ampho_005_KS_003.seq      G-QGGFLSDIADFDADFFGIAPREALAMDPQQRILLEVTWEAVERAGLDPTALRGSR-TG
ampho_006_KS_001.seq      Q-EGGFLHDVADFDPGFFDISPREALAMDPQQRLLLETAWEAVERAGIAPGSLRGSR-TG
averm_001_KS_001.seq      PQGGGFLHDAAHFDAGFFGISPREALAMDPQQRLLLETSWEAFERAGIDPLSVRGSR-TG
averm_001_KS_002.seq      TRNGGFLYDAGHFDAEFFGISPREALAMDPQQRLLLETAWETIEHAGINPHTLHGTP-TG
averm_002_KS_001.seq      TRSGGFLYDAGDFDAAFFELSPREALAMDPQQRLLLETTWETFEQGGIDPRSMRGSR-TG
averm_002_KS_002.seq      TRSGGFLYDAGNFDPDFFGISPREALAMDPQQRLLLETAWESIEHACINPDSLRGTP-TG
averm_002_KS_003.seq      TRNGGFLYDAGHFDAEFFGISPREALAMDPQQRLLLETAWETIEHAGINPHTLHGTP-TG
averm_002_KS_004.seq      TRNGGFLYGAGHFDAEFFGISPREALAMDPQQRLLLETAWETIEHAGINPHTLHGTP-TG
averm_003_KS_001.seq      ARYGGFLYEAGEFDADFFGISPREALAMDPQQRLLLETSWEAFERAGIDPLSMRGSR-TG
averm_003_KS_002.seq      TRCGSFLYDAGEFDAGFFGISPREALAMDPQQRLLLEASWEAMEQAGIDPTTVRGSQ-TG
averm_003_KS_003.seq      ARYGGFLYEAGEFDADFFGISPREALAMDPQQRLLLETSWEAFERAGIDPLSMRGSR-TG
averm_004_KS_001.seq      TRRGGFLEGAAEFDAGFFGISPREALAMDPQQRLMLEVSWEALERAGIDPATLRGSR-TG
averm_004_KS_002.seq      TRRGGFLEGAAEFDAGFFGISPREALAMDPQQRLMLEVSWEALERAGIDPATLRGST-TG
averm_004_KS_003.seq      TRCGGFLQGAAEFDAGFFGISPREALAMDPQQRLMLEVSWEALERAGIDPATLHGST-TG
borre_002_KS_001.seq      VREGGFLDGAGQFDAAFFGISPREALAMDPQQRLLLECSWEAIERSRIDPKTLHGSR-TG
borre_003_KS_001.seq      VREGGFLDGAGRFDPAFFGISPREAVVMDPQQRLLLECSWEAIERSRIDPKTLHGSR-AG
borre_003_KS_002.seq      VRHGGFLHDVAGFDAGFFGISPREALAMDPQQRLLLECSYEALERAGIDPATLRGSR-SG
borre_004_KS_001.seq      VREGGFVHDVAGFDAGFFGISPREALAMDPQHRLLLECSWEALERAGIDPSSLEGTR-TG
borre_005_KS_001.seq      AREGGFLYDAGEFDAAFFGISPREALSMDPQQRLLLECSWSALERAGIDPGSLRGKD-VG
borre_006_KS_001.seq      RREGGFLYDAGRFDAEFFGISPREALAMDPQQRLLLESSWEAIEHARIDPRSLHGSR-TG
epoth_002_KS_001.seq      VRAGSVLEDVDRFDAAFFGISPREAELMDPQHRIFMECAWEALENAGYDPTAYEGS--IG
epoth_003_KS_001.seq      -WAGLLTEPIDCFDAAFFGISPREARSLDPQHRLLLEVAWEGLEDAGIPPRSIDGSR-TG
epoth_003_KS_002.seq      VPRGGFLREVETFDAAFFHISPREAMSLDPQQRLLLEVSWEAIERAGQDPSALRESP-TG
epoth_003_KS_003.seq      VTKGAFLRDLQRLDATFFRISPREAMSLDPQQRLLLEVSWEALESAGIAPDTLRDSP-TG
epoth_003_KS_004.seq      TRFGGFLSDIDRFEPAFFGISPREATTMDPQQRLLLETSWEAFERAGILPERLMGSD-TG
epoth_004_KS_001.seq      -WAGLLTEAIDGFDAAFFGIAPREARSLDPQHRLLLEVAWEGFEDAGIPPRSLVGSR-TG
epoth_004_KS_002.seq      VAKGAFLRDVRSLDAAFFSISPREAMSLDPQQRLLLEVSWEAIERAGQDPMALRESA-TG
epoth_005_KS_001.seq      -WAGLLTEAVDGFDAAFFGTSPREARSLDPQQRLLLEVTWEGLEDAGIAPQSLDGSR-TG
lasal_001_KS_002.seq      VREGGFLDSAADFDAAFFGISPREALAMDPQQRLLLETSWEALERAGVAPGSLRGTR-TG
lasal_001_KS_003.seq      AREGGFLHDAGHFDPAFFGISPREALAMDPQQRLLLETSWEVFERAGLDPTALKGSR-TG
lasal_002_KS_001.seq      TRSGGFLLDAAEFDPAPFGISPREALAMDPQQRLMLEVSWEVLERAGIAPDSLKGSA-TG
lasal_002_KS_002.seq      ARQGGFLLDAADFDAGFFGISPREALSMDPQQRLLLETSWEVFERAGIDPATLKGSR-TG
lasal_002_KS_003.seq      TRSGSFLYDAADFDAQLFGISPREALAMDPQQRLLLETSWEVLERAGINPRSLRGSQ-TG
lasal_003_KS_001.seq      VREGGFLHDATEFDADFFGISPREALAMDPQQRIMLEVAWETFEHAGLDPQSFGGSD-TG
lasal_004_KS_001.seq      VKEAGFLYDAGEFDAGFFGISPREAVSMDPQQRLMLEASWEAFERAGLDPARQRGTA-TG
lasal_005_KS_001.seq      TGSGGFMDVAGDFDAALFGISPREALAMDPQQRLLLETSWEVLERSGLDPTSLAGTR-TG
lasal_005_KS_002.seq      TSQGGFLYDAGQFDPLLFGISPREALAMDPQQRLLLETSWEVLERAGIDPAALRGTD-TG
lasal_006_KS_001.seq      AREGGFLPDAGHFDAALFGISPREALAMDPQQRLVLEASWEALERAGINPQSLRGSR-TG
lasal_007_KS_001.seq      TSQGGFLYDAGQFDPLLFGISPREALAMDPQQRLLLETSWEVFERAGIDPSSLGGTR-TG
rapam_001_KS_001.seq      CVQGGFLDTAAGFDASFFGISPREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGSD-TG
rapam_001_KS_002.seq      RLQGGFLDAAAGFDASFFGISPREALAMDPQQRLVLEVSWEAFERAGIKPGSVRGTD-TG
rapam_001_KS_003.seq      RAEGGFLDAAAGFDAGFFGISPREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGSD-TG
rapam_001_KS_004.seq      CVQGGFLDTAAGFDAGFFGISPREALAMDPQQRLLLEVSWEAFERAGIEPGSVRGSD-TG
rapam_002_KS_001.seq      CVQGGFLDSAGGFDASFFGISPREALAMDPQQRLVLEASWEAFERAGIEPGSLRGSD-TG
rapam_002_KS_002.seq      CVEGGFLDSAASFDAGFFGISPREALAMDPQQRLIMEVSWEAFERAGIEPGSVRGSD-TG
rapam_002_KS_003.seq      RAEGGFLDTAAGFDAGFFGISPREALAMDPQQRLLLEVSWEAFERAGIEPGSVRGSD-TG
rapam_002_KS_004.seq      RLQGGFLDAAAGFDASFFGISPREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGTD-TG
rapam_002_KS_005.seq      CVEGGFLATAANFDASFFGISPREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGSD-TG
rapam_002_KS_006.seq      SVQGGFLDAAAGFDASFFGISPREALAMDPQQRLMLEVSWEAFERAGIEPGSVRGSD-TG
rapam_003_KS_001.seq      CVQGGFLDSAGGFDASFFGISPREALAMDPQQRLVLEVSWEAFERAGIEPGSLRGSD-TG
rapam_003_KS_002.seq      RAEGGFLDAAAGFDAGFFGISPREALAMDPQQRLVLEVSWEAFERAGIEPGSVRGSD-TG
rapam_003_KS_003.seq      CVEGGFLDSAGSFDAGFFGISPREALAMDPQQRLIMEVSWEAFERAGIEPGSVRGTH-RR
```

```
rapam_003_KS_004.seq      CVQGGFLDTAADFDAPFFGISPREALGMDPQQRLLLETTWEAIERAQIDPKSLRGRD-VG
nysta_002_KS_001.seq      H-EGGFLYDAAEFDAAFFGISPREATAMDPQQRLLLEVAWEALERAGIAPTALRGSR-SG
nysta_002_KS_002.seq      SVQGGFLHEAAGFDPGFFGISPREAVAMDPQQRLLLETSWEAFERAGIDPASLRGSA-AG
nysta_003_KS_001.seq      AQVGGFVEDAAGFDPGFFGISPREAVAMDPQQRILLEITWEALERAGIDPSTLRGTP-TG
nysta_003_KS_002.seq      VRAGGFLHDAGSFDADFFGMSPREAMATDSQQRLLLELSWEAVERAGIDPASLRDSG-TG
nysta_003_KS_003.seq      TRSGGFLHEAGEFDPAFFGMSPREALATDSQQRLLLESSWEAIERAGIDPLTLRGSA-TG
nysta_003_KS_004.seq      ARTGGFLHDAADFDADFFGMSPREAMATDSQQRLLLESSWEAIERAGIDPLTLRDSR-TG
nysta_003_KS_005.seq      TRSGGFLHEAGEFDPGFFGMSPREALATDSQQRLLLESSWEAIERAGIDPVSLRGSR-TG
nysta_003_KS_006.seq      TRSGGFLHEAGEFDPAFFGMSPREALATDSQQRLLLESSWEAIERAGIDPLTLRGSA-TG
nysta_004_KS_001.seq      TASGGFLHDAPDFDADFFGISPREAVAMDPQQRVVLESAWEAFERAGIDPTSVKGSR-TG
nysta_004_KS_002.seq      TRSGGFLHEAGEFDPGFFGMSPREALATDSQQRLLLESSWEAIERAGIDPVSLRGSR-TG
nysta_004_KS_003.seq      STQGGFLHEADEFDPTFFGISPREALVMDPQQRLLLETTWESFERAGIRPETLRSTL-TG
nysta_004_KS_004.seq      SVQGGFLRDVAEFDPGFFGISPREALSMDPQQRLLLETAWEAFEHAGIDPVGQRGSR-TG
nysta_004_KS_005.seq      TLEGGFLTGVADFDARFFGISPREALAMDPQQRLLLETTWEALERAGIDPTTLRGST-TG
nysta_004_KS_006.seq      A-QGGFLYDATDFDPGLFGISPREALVMDPQQRILLETSWEALERAGIDPATLRGSGTTG
nysta_005_KS_001.seq      SQRGGFLDCIADFDPGFFGISPREAVTMDPQQRLLLTTAWEAVERAGIDATTLRATR-TG
nysta_005_KS_002.seq      G-QGGFLRDIADFDAAFFGISPREALAMDPQQRILLEVAWEAAERAGIDPQTLRGSD-TG
nysta_005_KS_003.seq      PRSAASSYDAAAFDPGFFDISPREALAMDPQQRLLLETAWEAVERTGTDPTRLRGSR-TG
nysta_006_KS_001.seq      Q-EGGFLPDAAAFDPGFFDISPREALAMDPQQRLLLETAWEAVERSGIDPAGLRGSR-TG
spino_002_KS_001.seq      ACAGGFLQGVSEFDAGFFGISPREALAMDPQQRLLLEVAWEVFERAGLEQRSTRGSR-VG
spino_003_KS_001.seq      TRHGGFLHDAGEFDAGFFDISPREALAMDPQQRLLLEISWETVESAGMDPRSLRGSR-TG
spino_003_KS_002.seq      ARMAGFLYDAGEFDADLFGISPREALAMDPQQRLVLEIAWEALERAGIDPLSLKGSG-VG
spino_004_KS_001.seq      TRAGGFLAGAGDFDPAMFGISPREALAMDPQQRLLLELSWEALERAGIDPTSLRGSK-TG
spino_004_KS_002.seq      VCEGGFLRDAAEFDADMFGISPREALAMDPQQRLLLEVAWETLERAGIDPFSLHGSR-TG
spino_004_KS_003.seq      VREGAFLTGADRFDAGFFGISPREARAMDPQQRLLLEVAWEVFERAGIAPLSLRGSR-TG
spino_005_KS_001.seq      VREGGFLRGVADFDAGLFGISPREAQAMDPQQRLLLEISWEVFERAGIDPFSLRGTK-TG
spino_005_KS_002.seq      VCEGGFLRDAAEFDAGFFGISPREALAMDPQQRLLLEVAWETVERAGIDPLSLRGSR-TG
spino_005_KS_003.seq      VREGAFLTDADRFDAGFFGISPREALAMDPQQRLLLELSWEAIERAGIDPGSLRGSR-TG
tylac_001_KS_002.seq      VREGGFLHDAAEFDAEFFGISPREAVAMDPQQRLLLETSWEAIERAGIDPHSLHGSR-TG
tylac_001_KS_003.seq      VREGGFLHDAAEFDAEFFGISPREAVAMDPQQRLLLETSWEAIERAGIDPHSLHGSR-TG
tylac_002_KS_001.seq      AREGGFLSHCAGFDAEFFGISPREALAMDPQQRLLLETSWEALERAGVTADRARGSR-TG
tylac_003_KS_001.seq      VREGGFLHDAAEFDAEFFGISPREAVAMDPQQRLLLETSWEAIERAGIDPHSLHGSR-TG
tylac_003_KS_002.seq      VREGGFLHEAPDFDPSFFGISPREALAMDPQQRLLLETSWEALERAGIDPARLRGSR-TG
tylac_004_KS_001.seq      CREGGFLHDAGDFDADFFGISPREAVAMDPQQRLLLETSWEAIEAAGIDPRGLRGSR-TG
tylac_005_KS_001.seq      CREGGFLEGAGDFDAAFFGISPREALVMDPQQRLLLEVSWEALERAGIDPSSLRGSR-GG
nidda_001_KS_002.seq      VRHGGFLHDAAQFDAEFFGISPREATAMDPQQRLLLETSWEALERAGVVPESLRGGR-TG
nidda_001_KS_003.seq      VRHGGFLHDAAQFDAEFFGISPREATAMDPQQRLLLETSWEALEHAGIAPHTLRTTR-TG
nidda_002_KS_001.seq      VRHGGFLHDAAQFDAEFFGISPREATAMDPQQRLLLETSWEALEHAGIAPNTLRTTR-TG
nidda_003_KS_001.seq      VRHGGFLHDAAQFDAEFFGISPREATAMDPQQRLLLETSWEALEHAGIAPHTLRTTR-TG
nidda_003_KS_002.seq      ARAGGFLHDAAEFDPGFFGISPREALAMDPQQRLSLETAWESFERAGLDPTRLRGSR-TG
nidda_004_KS_001.seq      CREGGFVRGVDRFDPGLFGISPNEALAMDPQQRLLLETVWESLERAGIDPKSLRGSR-TG
nidda_005_KS_001.seq      CREGGFLAGAGDFDAAFFGISPNEALVMDPQQRLLLETSWEALERAGVVPESLRGGR-TG
monen_001_KS_002.seq      TRQGGFLYDAADFDAEFFDISPREALAVDPQQRLLLECAWEAFERAGLDPRALKGSP-TG
monen_002_KS_001.seq      VRHGGFVDDAGSFDADFFGISPREALAMDPQQRLMLETSWELFERAGIEPVSLKGSR-TG
monen_003_KS_001.seq      AREGGFLHDADLFDPEFFGISPREAAVLDPQQRLLLECAWEALERAGIDPRSLQGSR-TG
monen_003_KS_002.seq      TREGGFLYDAARFDAGFFGISPREAAATDPQQRLLLETAWQAFEHAGIDPAALRGTP-CG
monen_004_KS_001.seq      VRRGGFLYDAPAFDASFFGISPREALAMDPQQRVLMETAWQLLERAGIDPASLKLSA-TG
monen_004_KS_002.seq      TDQGGFLPDAGDFDAAFFGINPREALAMDPQQRLLLEASWEVLERAGIDPTTLKGTP-TG
monen_005_KS_001.seq      VREGGFLAAPDRFDSDFFGFSPREALASSPQLRLLLETSWEALERAGINPASLKGSP-TG
monen_005_KS_002.seq      ADEGAFLPDAGDFDAAFFGINPREALAMDPQQRLLLEASWEVLERAGIDPTTLKGTP-TG
monen_006_KS_001.seq      VREAGFLRDAARFDADFFGINPREALAADPQQRVLLEVSWELFERAGIDPATLKDTL-TG
monen_007_KS_001.seq      VREGGFVYDAGDFDPTFFGIGPTEAAAMAPQQRLALETAWEAIERAGIDPLSLRSSD-TS
monen_008_KS_001.seq      VRDGAFLYDAGHFDAEFFGISPREATAMDPQQRLLLETAWEAIEHAGMNPHALKGSD-TG
monen_008_KS_002.seq      VREGGFLYDAGDFDAGFFGISPREAVAMDPQQRLLLETAWEAIERAGLDRETLKGSD-AG
oligo_001_KS_002.seq      PGEGGFLADVAEFDAGFFGISPREALAMDPQQRLLLETSWEALERAGVDALKLRGSR-TG
oligo_001_KS_003.seq      VREGGFLTDVADFDAAFFGISPREALAMDPQQRLLLETSWEALERAGIVPASLAGSR-TG
oligo_001_KS_004.seq      -QTGGFLPDVAEFDAGFFGISPREALAMDPQQRLLLETSWEALERAGVDALKLRGSR-TG
oligo_002_KS_001.seq      RLAGGFLADAAGFDAGLFGISPREALAMDPQQRLLLETSWEVLERAGLDPTTVRGAD-GG
oligo_002_KS_002.seq      AMKGGFLPDAADFDAAFFGISPREATAMDPQQRLLLEVSWQALEGAGLDPAALRGSR-TG
oligo_003_KS_001.seq      AEAGGFVEDIASFDAAFFGISPREAQSMDPQQRLLLETSWEALEHAGLDIHALRGSR-TG
oligo_003_KS_002.seq      PQQGGFLDRVAEFDAAFFGISPREALTMDPQQRLLLETSWEALERAGIAPGTLRGSR-TG
oligo_004_KS_001.seq      ATEGGFLRGATEFDAEFFGVSPREALAMDPQQRLLLETAWEALENTGLDPRSLAGSR-TG
oligo_004_KS_002.seq      VREGGFLYDAAEFDPEFFGISPREALAMDPQQRLLLETSWEALERAGIDPRAVRGSR-TG
oligo_005_KS_001.seq      TLEGGFLHGAGEFDADFFGISPREAIATDPQQRLLLETSWEALERAGVDPRSLRGSV-TG
oligo_006_KS_001.seq      ARTGAFLHDAGDFDPAFFRISPREAMAMDPQQRLLLETSWEALERAGIDPATLKGSR-TG
oligo_006_KS_002.seq      VGQAGFLHEAGEFDAAFFGISPREAASMDPQQRLLLETSWEAVEQAGIDPWSLRGSS-TG
oligo_006_KS_003.seq      LRRAGFLHEAGDFDAGFFGISPREAMAMDPQQRLLLEASWEAVERTGIDPHSLRGSD-TG
oligo_007_KS_001.seq      TREGAFLRGAGDFDAGFFGISPREALTMDPQQRLLLETSWEALERAGIDPHTLRGSR-TG
oligo_007_KS_002.seq      ARQGAFLAEAGDFDAAFFGISPREALAMDPQQRLLLETSWEALERAGFDPVGLRGSR-TG
rifam_001_KS_001.seq      VGQGGFLHDAGEFDAGFFGISPREAVAMDPQQRLLLETSWEALENAGVDPIALKGTD-TG
rifam_001_KS_002.seq      VRHGAFLDDAAGFDAAFFGISPNEALAMDPQQRLLLETSWELFERAAIDPTTLAGQD-IG
rifam_001_KS_003.seq      TRSGGFLHDAAQFDAGLFGISPREALAMDPQQRLLLETSWEALERAGVDPLSARGSD-VG
rifam_002_KS_001.seq      VREGAFLQDAAQFDAGFFGISPREALAMDPQQRQLLEVSWETLERAGIDPHSVRGSD-IG
rifam_002_KS_002.seq      IRQGGFLHEAALFDPGFFGISPREALAMDPQQRLLLEASWEALERAGIDPTKARGDA-VG
rifam_002_KS_003.seq      TRHGGFLHEAGLFDAGFFGISPREAVAMDPQQRLLLEASWEAMEDAGVDPLSLKGND-VG
rifam_003_KS_001.seq      TDQGGFLHDAALFDPGFFGISPREALAMDPQQRLLLEASWEALEGVGLDPASLQGTD-VG
rifam_004_KS_001.seq      TDQGGFLHEAGLFDAGFFGISPREAVAMDPQQRLLLETSWEAIERTGTDPLSLKGSD-IG
```

```
rifam_005_KS_001.seq      TSQGGFLRGAGLFDAGLFGISPREALVMDPQQRVLLETSWEALEDAGVDPLSLKGSD-VG
rifam_005_KS_002.seq      TSRGGFLDGAGLFDAGFFGISPREALAMDPQQRLLLEAAWEALEGTGVDPGSLKGAD-VG
pikro_001_KS_002.seq      VRQGGFIENVAGFDAAFFGISPREALAMDPQQRLLLETSWEAVEDAGIDPTSLRGRQ-VG
pikro_001_KS_003.seq      VREGGFLHDAAEFDAEFFGVSPREALAMDPQQRMLLTTSWEAFERAGIEPASLRGSS-TG
pikro_002_KS_001.seq      AREAGFLYEAGEFDADFFGISPREALAMDPQQRLLLEASWEAFEHAGIPAATARGTS-VG
pikro_002_KS_002.seq      CRAGGFLDEAGEFDADFFGISPREALAMDPQQRLLLETSWEAVEDAGIDPTSLQGQQ-VG
pikro_003_KS_001.seq      CRSGGFLHDAGEFDADFFGISPREALAMDPQQRLSLTTAWEAIESAGIDPTALKGSG-LG
pikro_004_KS_001.seq      VRNAAFLDDAAGFDAAFFGISPREALAMDPQQRQLLEASWEVFERAGIDPASVRGTD-VG
Stamb_001_KS_002.seq      PTAGGFLDDIAGFDAALFGISPREAVAMDPQQRLLLEVTWEALERLGTDPTDLRGSR-TG
Stamb_001_KS_003.seq      ARHGGFLHDAADFDAELFGISPREALAMDPQQRLLLELSWEAFERAGIDPTGLRGSD-TG
Stamb_001_KS_004.seq      ARRGGFLPDAAGFDADLFGVSPREAQAMDPQQRLMLEISWETFERAGVDPASLQGSR-VG
Stamb_002_KS_001.seq      GHEGGFLHEAAEFDAGLFEISPREAVAMDPQQRLLLETSWEVFERAGIDPKSVRGAR-VG
Stamb_002_KS_002.seq      APLGGFLDDVAGFDAGLFGISPREALAMDPQQRLLLETAWEAFEAAGLDPLAVAGEP-IG
Stamb_003_KS_001.seq      TRRGGFVPDAAEFDAALFGINPREALAMDPQQRMLLEAAWEVAERAGIDPHSLAGSD-TG
Stamb_003_KS_002.seq      ARRGGFVDGATEFDAALFGISPREALAMDPQQRLLLEASWEAFEHAGIDPLSQNHTR-TG
Stamb_003_KS_003.seq      AARGGFLDEATDFDAALFGISPREALAMDPQQRLLLETAWEAFESAGIATSDAHGSS-TG
Stamb_003_KS_004.seq      AARGGFVDGATEFDAGLFGISPREALAMDPQQRLLLEASWEAFESAGLNPRSLRGRP-VG
Stamb_003_KS_005.seq      TLRGGFVDGVAEFDAGLFGISPREALAMDPQQRLLLETVWETFESAGVDPRAVRGRS-VG
Stamb_004_KS_001.seq      ARQGGFVHDATAFDADFFGVSPREALAMDPQQRLLLRAAWEVFEHAGIDPQSVRGTR-TG
Stamb_004_KS_002.seq      TFEGGFLQDAGGFDPAFFGISPREALAMDPQQRLLLETSWEAFERAGVDPAELRASS-TG
Stamb_004_KS_003.seq      VKEGGFLADASHFDASLFGISPREAVATDPQQRLLLEATWEAMERATIDPSSLRGSP-TG
Stamb_005_KS_001.seq      APRGGFVHDVGDFDAGLFGISPREAAATDPQQRQLLEASWEAVERAGIDMRTLRGGR-TG
Stamb_006_KS_001.seq      SRQGGFVHDAGRFDADFFSISPREAAAMDPQQRLLLEASWEAFERAGIDPKEVKGSP-VG
Stamb_006_KS_002.seq      VREGGFLGDVAEFDAPFFGISPREAVATDPQQRLMLQVAWEALEHAGVDPTALRGSD-TA
Stamb_007_KS_001.seq      VREGGFLHDAGEFDAGLFGISPHEALAMDPQQRLMLETSWEALERAGIDPLSLRGSR-TG
Stamb_007_KS_002.seq      TRVGGFVHDATSFDADLFGISPREAMAMDPQQRLLLETSWELFERAGLSPHAVRSTP-TG
Stamb_008_KS_001.seq      ASQGGFLHDAGMFDPAFFGINPREALAMDPQQRLVLESVWEAMERAGIHPDTLRGSR-TG
Stamb_008_KS_002.seq      VLQGGFVHDADRFDPVLFGISPREALAMDPQQRLMLEASWEAFERAGLDPTSLRGSS-TG
Stamb_009_KS_001.seq      SREGGFLDGVADFDPSLFNISPHEALAMDPQQRLMLETSWEAVERAGIDPLSLRGSR-TG
Stamb_009_KS_002.seq      PRTGGFLPDATDFDAELFGISPREALAMDPQQRLLLEASWEALERAGIHPRSLAGSR-TA
Stamb_009_KS_003.seq      TRSGGFLDDIAAFDAGLFGISPHEALAMDPQQRMLLESAWEAFERAGIDPTSVRRSD-TG
Stamb_009_KS_004.seq      TTEGGFLHDATEFDADLFGISPREAVAMDPHQRLLLESAWELLERGGIAPTAVRGTR-TG
                             .        ::   *   * **     : *  :   :.  *

mycol_001_KS_001.seq      VFVGAWAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_001_KS_002.seq      VFVGAWAQSYGATNS--------DDAEGYAMTGGAISVMSGRIAYTLGLEGPAITVDTAC
mycol_001_KS_003.seq      VFVGAGAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_001_KS_004.seq      VFAGAWAQSYGATNS--------DDAEGYAMTGGSTSVMSGRIAYTLGLEGPAITVDTAC
mycol_001_KS_005.seq      VFVGAGAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_001_KS_006.seq      VFAGAWAQSYGATNS--------DDAEGYAMTGGSTSVMSGRIAYTLGLEGPAITVDTAC
mycol_001_KS_007.seq      VFVGAGAQSYGATNS--------DDAEGYAMTGGAISVMSGRIAYTLGLEGPAITVDTAC
mycol_001_KS_008.seq      VFVGAGAQSYGATNS--------DGAEGYAMTGGAISVMSGRIAYTLGLEGPAITVDTAC
mycol_002_KS_001.seq      VFVGAWAQSYGATNS--------DGAEGYAMTGGSTSVMSGRIAYTLGLEGPAITVDTAC
mycol_003_KS_001.seq      VFVGAWAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_003_KS_002.seq      VFVGAWAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_003_KS_003.seq      VFVGAWAQSYGATNS--------DDAEGYAMTGGAISVMSGRIAYTLGLEGPAITVDTAC
mycol_003_KS_004.seq      VFAGAWAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_003_KS_005.seq      VFAGAWAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_003_KS_006.seq      VFAGAWAQSYGATNS--------DDAEGYAMTGGATSVMSGRIAYTLGLEGPAITVDTAC
mycol_003_KS_007.seq      VFAGAWAQSYGATNS--------DDAEGYAMTGGSTSVMSGRIAYTLGLEGPAITVDTAC
eryth_001_KS_001.seq      VFVGLIPQEYGPRLAEGG-----EGVEGYLMTGTTTSVASGRIAYTLGLEGPAISVDTAC
eryth_001_KS_002.seq      VFVGMSHQGYATGRPRPE-----DGVDGYLLTGNTASVASGRIAYVLGLEGPALTVDTAC
eryth_002_KS_001.seq      VFLGVAKFGYGED-TAAA-----EDVEGYSVTGVAPAVASGRISYTMGLEGPSISVDTAC
eryth_002_KS_002.seq      VFVGMNGQSYMQLLAGEA-----ERVDGYQGLGNSASVLSGRIAYTFGWEGPALTVDTAC
eryth_003_KS_001.seq      VFTGVGTVDYGPRPDEAP-----DEVLGYVGTGTASSVASGRVAYCLGLEGPAMTVDTAC
eryth_003_KS_002.seq      VFLGAAYQGYGQDAVVPE-----D-SEGYLLTGNSSAVVSGRVAYVLGLEGPAVTVDTAC
ampho_002_KS_001.seq      VFVGSYH--WGAPSADTA-----TELHGHALTGTAASVLSGRLSYVYGLEGPAVTVDTAC
ampho_002_KS_002.seq      AFFGASYQDYSSSVQNGT-----EGAEVHMMTGTAASVLSGRISYLLGLEGPAVTVDTAC
ampho_003_KS_001.seq      VFMGTTGQDYGEVIKASD-----EDAEVYSTTGHAASVISGRLSYFLGAEGPAVTVDTGC
ampho_003_KS_002.seq      VFAGVMYNDYGTIL-------TGEEYEAFRGNGSAPSVASGRVSYTLGLEGPAVTVDTAC
ampho_003_KS_003.seq      VFAGVMYHDYGSLL-------GGKEFEGFQGQGSAGSVASGRVSYTFGFEGPAVTVDTAC
ampho_003_KS_004.seq      VFAGVMYSGYGTAL-------DGDEFEGFQGQGSALSVASGRVSYTFGFEGPAMTIDTAC
ampho_003_KS_005.seq      VFAGVMYSDYSAML-------ASPEFEGFQGSGSSPSVASGRVSYTFGFEGPAVTVDTAC
ampho_003_KS_006.seq      VFAGVMYSDYGSIL-------GGKEFEGLQGQGSAGSVASGRVSYTFGFEGPAVTVDTAC
ampho_004_KS_001.seq      VFIGAMAQDYRVGP--------ADGAEGFQLTGNTGSVLSGRISYTFGTVGPAVTVDTAC
ampho_004_KS_002.seq      VFAGVMYSDYSAML-------GSPEFEGFQGSGSSPSLASGRVSYTLGLEGPAVTVDTAC
ampho_004_KS_003.seq      TFIGSSYQEYGMGAG--------DGAEGHLVTGTSPSVLSGRLAYVFGLEGPAVTVDTAC
ampho_004_KS_004.seq      TFIGASYQDYGAAVP-GS-----EGSEGHMITGSLPSVLSGRVSYLFGLEGPAVTLDTAC
ampho_004_KS_005.seq      VFVGTNGQDYPNLLRRST-----SDVSGYVATGNTASVMSGRLSYALGLEGPAVTIDTAC
ampho_004_KS_006.seq      VFVGGGSGDYRPPAE-----------YGQWQTAQSASLLSGRLAYTFGIQGPTVSVDTAC
ampho_005_KS_001.seq      TFIGTNGQDYAYLLVRSL-----DDATGDVGTGIAASAVSGRLSYTFGLEGPAITVDTAC
ampho_005_KS_002.seq      VFMGVNGQDYSSLVMGSR-----DDVAGHATAGLAVSVVSGRLSYALGLEGPALSVDTAC
ampho_005_KS_003.seq      VFMGVSGQDYSGLVMRSR-----DDIASHATTGLAVSVVSGRLSYTLGLEGPALSVDTAC
ampho_006_KS_001.seq      VFVGTNSQDYAHLVLASD-----DDMGGYAGNGLAASVMSGRLSFALGFEGPAVTLDTAC
averm_001_KS_001.seq      VFAGALSFDYGPRMDTASS-EGAADVEGHILTGTTGSVLSGRIAYSFGLEGPAITVDTGC
averm_001_KS_002.seq      VFTGTNGQDYALRVHN-----AGQSTDGFALTGTAGSVISGRISYTFGFEGPAVSVDTAC
averm_002_KS_001.seq      VFVGINPEDYTTGYTHQ----PSNAVEGYLLTGSAASIASGRISYNFGLEGPAITIDTAC
```

```
averm_002_KS_002.seq      VFAGLTYHDYAARFPT-----APAGFEGYLGHGSAGSIASGRVAYALGLEGPALTVDTAC
averm_002_KS_003.seq      VFTGTNGQDHAAHIRQ-----APSGTEGFVLTGAATSIASGRISYILGLEGPAVTLDTAC
averm_002_KS_004.seq      VFAGINAQDHAAHIRQSR---DVETIEGYALTGSSGSVASGRVAYTLGLEGPAVSVDTAC
averm_003_KS_001.seq      VFAGVMYHDYGSRLGT-----IPEGFEGYIGNGSGGAVASGRVAYTLGLEGPAVSVDTAC
averm_003_KS_002.seq      VFAGLIPQAYGPRLHEN----AAADTEGYVLTGTSGSVASGRISYTFGFEGPAVSVDTAC
averm_003_KS_003.seq      VFAGVMYHDYAARLHH-----VPEGFEGLIANGSAGSVATGRVAYSFGLEGPAVTVDTAC
averm_004_KS_001.seq      VFAGLMSQDYATRLLS-----VPDDLAGYLGNGNAGSILSGRVAYTFGFEGPAVTVDTAC
averm_004_KS_002.seq      VFAGMCSQDYADLVRR-----ATEDLEGYAMTGLSSSVTSGRVAYTLGLEGPAVTVDTAC
averm_004_KS_003.seq      VFAGVSQQDYAELLRR-----GTQDHEGYALTGVSNSVVSGRLSYTFGFEGPAVTVDTAC
borre_002_KS_001.seq      VFAGSNWQDYNTLLLN-----AEERSQSYLATGASGSVLSGRVSYTLGMEGPAITVNTAC
borre_003_KS_001.seq      VFVGSNGQDYGTLLLR-----ADDRSHAYLATGASASVLSGRISYTLGLEGPAVTISTAC
borre_003_KS_002.seq      VYAGVMYHEYASRLGA-----TPAGFEGTLGTGSSGSIASGRISYTFDLTGPAVTVDTAC
borre_004_KS_001.seq      VYTGLMTHEYATRLPS-----IDEELEGVIGIGNAGSVASGRVSYTLGLNGPAVTVDTAC
borre_005_KS_001.seq      VYVGAWNSNYGRGG-------GAESSEGHLLTGNASSVVSGRVAYVLGLEGPAVTIDTAC
borre_006_KS_001.seq      VWFGTIGQDYFSLFAASG-----GEHANYLATACSASVMSGRVSYVLGLEGPAVTVDTAC
epoth_002_KS_001.seq      VYAGANMSSYLTSNLHEHPAMMRWPGWFQTLIGNDKDYLATHVSYRLNLRGPSISVQTAC
epoth_003_KS_001.seq      VFVGAFTADYAR-TVARLPREE---RDAYSATGNMLSIAAGRLSYTLGLQGPCLTVDTAC
epoth_003_KS_002.seq      VFVGAGPNEY-AERVQDLADEA---AGLYSGTGNMLSVAAGRLSFFLGLHGPTLAVDTAC
epoth_003_KS_003.seq      VFVGAGPNEYYTQRLRGFTDGA---AGLYGGTGNMLSVAAGRLSFFLGLHGPTLAMDTAC
epoth_003_KS_004.seq      VFVGLFYQEYAALAGGI------EAFDGYLGTGTTASVASGRISYVLGLKGPSLTVDTAC
epoth_004_KS_001.seq      VFVGVCATEYLHAAVAHQPREE---RDAYSTTGNMLSIAAGRLSYTLGLQGPCLTVDTAC
epoth_004_KS_002.seq      VFVGMIGSEH-AERVQGLDDDA---ALLYGTTGNLLSVAAGRLSFFLGLHGPTMTVDTAC
epoth_005_KS_001.seq      VFLGACSSDYSH-TVAQQRREE---QDAYDITGNTLSVAAGRLSYTLGLQGPCLTVDTAC
lasal_001_KS_002.seq      VFVGTTFQDYGPR-LHQ----GTETTEGYLLTGGTPSVASGRIAYTLGLEGPAVTVDTAC
lasal_001_KS_003.seq      VFAGLMHHDYAGP-GDR----LPEGVEGYALTGTQGSVVSGRIAYVYGLEGPAVTVDTAC
lasal_002_KS_001.seq      VYVGAGNLGYLVG-MQQ----ARPGVEGHSLTGNIGSIISGRISYTFGLEGPAVTVDTAC
lasal_002_KS_002.seq      VFAGAMYQEYAAN---V----APETVEGHFLSGTSSSVVSGRVSYTFGLEGPAVTIDTAC
lasal_002_KS_003.seq      VFVGAVNTGYGTG-LPH----TPEDIEGYLGTGTSSSVASGRVAYTFGLEGPAVTVDTAC
lasal_003_KS_001.seq      VFVGISNVDYTWG-HAR----VPEAVEGYFGTGNFASVLSGRLAYTFGLEGPALTVDTAC
lasal_004_KS_001.seq      VFVGATATGYVSPAAEVP-----EGAEGFAITGNMTAVTSGRISYTLGLQGPAVTIDTAC
lasal_005_KS_001.seq      VFVGVSPSGYGSG-AQD----VPEGAEGYTMTGVAPAVASGRISYAFGLEGPAVTVDTAC
lasal_005_KS_002.seq      VFAGASASGYGSG-AQE----ATEGAEGYAMTGAATSVISGRVAYTFGLEGPAVTLDTAC
lasal_006_KS_001.seq      VFVGHVASGYGAHVTDAP-----DGVEGYLGTGTSASVASGRVAYTLGLEGPAVTVDTAC
lasal_007_KS_001.seq      VFAGAIASGYAIG-LPE----FPEGVQGYVGTGTSASVLSGRISYAFGLEGPAVTVDTAC
rapam_001_KS_001.seq      VFMGGFPGGYGAG---A-------DLEGFGATAGAASVLSGRVSYFFGLEGPAITVDTAC
rapam_001_KS_002.seq      VFMGAYPGGYGIG---A-------DLGGFGTTAGAVSVLSGRVSYFFGFEGPAFTVDTAC
rapam_001_KS_003.seq      VFMGAYPGGYGAG---A-------DLGGFAATASATSVLSGRVSYFFGLEGPAFTVDTAC
rapam_001_KS_004.seq      VFIGAFPVGYGAG---F-------DREGYGATSG-PSVLSGRVSYVFGLEGPAITMDTAC
rapam_002_KS_001.seq      VFMGAYPGGYGVG---A-------DLGGFGATAGAVSVLSGRVSYFFGLEGPAVTVDTAC
rapam_002_KS_002.seq      VFMGAYAGGYGAG---A-------DLGGFAATASATSVLSGRVSYFFGLEGPAITVDTAC
rapam_002_KS_003.seq      VFIGAFPVGYGAG---A-------AREGYGATAA-PNVLSGRLSYFFGLEGPAITMDTAC
rapam_002_KS_004.seq      VFMGAYPGGYGIG---A-------DLGGFGATASAVSVLSGRVSYFFGLEGPAITVDTAC
rapam_002_KS_005.seq      VFMGAFPGGYGIG---A-------DLEGYGATAG-LNVLSGRLSYFFGLEGPAVTVDTAC
rapam_002_KS_006.seq      VFIGAYPGGYGIG---A-------DLGGFGTTAGAASVLSGRVSYFFGLEGPAFTVDTAC
rapam_003_KS_001.seq      VFIGAYPGGYGAG-AGA-------DLEGYGTTSG-PSVLSGRVSYFFGLEGPAITVDTAC
rapam_003_KS_002.seq      VFMGAYPGGYGIG---A-------DLGGFGATASSVSVLSGRVSYFFGLEGPAFTVDTAC
rapam_003_KS_003.seq      -LHGRVRGGYGAG---A-------DLGGFAATASATSVLSGRVSYFFGLEGPAITVDTAC
rapam_003_KS_004.seq      VYVGGAAQGYGVGVDQQ---------HDNGITGSSVSLLSGRVSYALGLEGPGVTVDTAC
nysta_002_KS_001.seq      VFVGSYH--WGAPSADAA-----TELHGHALTGTAASVLSGRLAYTLGLEGPAVTVDTAC
nysta_002_KS_002.seq      TFFGASYQDYSSTVQN-----GTGESEAHMVTGTAASVLSGRVSYLLGLEGPAVTVDTAC
nysta_003_KS_001.seq      VFVGTTGQDYGEVIKASA-----EDVEVYSTTGHAASVISGRLSYTLGAEGPAVTVDTGC
nysta_003_KS_002.seq      VFAGVMYNDYGTTL-------TGDEYEAFRGNGSAPSVASGRVSYTLGLEGPAVTVDTAC
nysta_003_KS_003.seq      VFAGVMYSDYGSIL-------GGKEFEGFQGQGSAGSVASGRVSYALGFEGPAVTVDTAC
nysta_003_KS_004.seq      VFAGVMYSGYGTRL-------DGAEFEGFQGQGSALSVASGRVSYTFGFEGPAMTVDTAC
nysta_003_KS_005.seq      VFAGVMYSDYSAML-------ASPEFEGFQGSGSSPSLASGRVAYTLGLEGPAVTVDTAC
nysta_003_KS_006.seq      VFAGVMYSDYGSIL-------GGKEFEGFQGQGSAGSVASGRVSYTLGFEGPAVTVDTAC
nysta_004_KS_001.seq      VFIGAMAQDYRVGP--------ADGAEGFQLTGNTGSVLSGRISYTFGTVGPAVTVDTAC
nysta_004_KS_002.seq      VFAGVMYSDYSAML-------ASPEFEGFQGSGSSPSLASGRVAYTLGLEGPAVTVDTAC
nysta_004_KS_003.seq      TFVGSSYQEYGLGAG--------DGTEGHMVTGSSPSVLSGRLSYVFGLEGPAVTVDTAC
nysta_004_KS_004.seq      TFVGASYQDYASGVP-NS-----EGSEGHMITGTLSSVLSGRVSYLFGFEGPAVTLDTAC
nysta_004_KS_005.seq      VFVGTNGQDYPTLLRRSA-----SDVAGYVATGNTASVMSGRLSYALGLEGPAVTIDTAC
nysta_004_KS_006.seq      VFVGGGSGDYRPPEE-----------AGQWQTAQSASLLSGRLAYTFGIQGPTVSVDTAC
nysta_005_KS_001.seq      AFIGTNGQDYAYLLVRSL-----DDATGDVGTGIAASAASGRLSYTLGLEGPALTVDTAC
nysta_005_KS_002.seq      VFMGVSGQDYAGLVMRSR-----DDIAGHATTGLAVSVVSGRLAYALGLEGPALSVDTAC
nysta_005_KS_003.seq      VFVGTNGQDYAGLVLRAQ-----EDVEGHAGTGLAASVISGRLAYAFGFEGPAVTVDTAC
nysta_006_KS_001.seq      VFVGTNGQDYAHLVLAAQ-----DDMGGYAGNGLAASVLSGRLAFALGLEGPAVTLDTAC
spino_002_KS_001.seq      VFVGTNGQDYASWL-RT----PPPAVAGHVLTGGAAAVLSGRVAYSFGFEGPAVTVDTAC
spino_003_KS_001.seq      VFAGLMYEGYDTGA-HR----AGEGVEGYLGTGNAGSVASGRVAYAFGFEGPAVTVDTAC
spino_003_KS_002.seq      TYIGAGSRGYATDV-RQ----FPEEAEGYLLTGTSASVLSGRVAYSFGFEGPAVTVDTAC
spino_004_KS_001.seq      VFGGVTPQEYGPSLQE-----MSRNAGGFGLTGRMVSVASGRVAYSFGFEGPAVTVDTAC
spino_004_KS_002.seq      VFAGLMYHDYGARFITR----APEGFEGHLGTGNAGSVLSGRVAYSFGFEGPAVTVDTAC
spino_004_KS_003.seq      VFAGTNGQDHGAKVAA------APEAAGHLLTGNAASVLAGRLSYTFGLEGPAVAVDTAC
spino_005_KS_001.seq      VFAGLIYHDYASRF-RK----TPAEFEGYFATGNAGSVASGRVAYTFGLEGPAVTVDTAC
spino_005_KS_002.seq      VFAGLMHHDYGARFITR----APEGFEGYLGNGSAGGVFSGRVAYSFGFEGPAVTVDTAC
spino_005_KS_003.seq      VFAGLMYHDYGARFASR----APEGFEGYLGNGSAGSVASGRIAYSFGFEGPAVTVDTAC
tylac_001_KS_002.seq      VYAGVMPQEYGPR-LAE----GAEGSDGYLLTGTSGSVVSGRVAYTLGLEGPAVTVDTAC
tylac_001_KS_003.seq      VYVGLTHQEYASR-LHE----APEEYEGYLLTGKSASVVSGRISYTLGLEGPSLSIDTAC
```

```
tylac_002_KS_001.seq      VYAGVMYDDYGARVLYGAGAGPPEDLEGYLVNGSAGSIASGRVSYTFGLRGPAVTVNTAC
tylac_003_KS_001.seq      VYAGVMYHDYGTG-QTS-----ATDTSGYSGTGTSGSVVSGRVAYTLGLEGPAVTVDTAC
tylac_003_KS_002.seq      VFVGTNGQHYMPLLQNGG-----DSFDGYLGTGNSASVMSGRLSYVFGLEGPAVTVDTAC
tylac_004_KS_001.seq      VYVGAWDSGYTGQ-AHA----PSAELEADLLTGGVVSFTSGRIAYTLGLEGPALTVDTAC
tylac_005_KS_001.seq      VYVGAAHGSYASDPRLVP-----EGSEGYLLTGSADAVMSGRISYALGLEGPSMTVETAC
nidda_001_KS_002.seq      VFVGTTAPEYGPRLHEGT-----DGYEGFLLTGTTASVASGRIAYALGTRGPALTVDTAC
nidda_001_KS_003.seq      VFTGASQQEYGTQSREAA-----DKYGGHLLTGTLASVMSGRVAYTLGLQGPALTVDTAC
nidda_002_KS_001.seq      VFTGVMYDDYGSRCFPSL-----PEVEGYVVNGSAGSVASGRVAYTLGLQGPALTVDTAC
nidda_003_KS_001.seq      VFTGVMHHDYGSHQVGSA-----ADASGQLGLGTAGSVASGRVAYTLGLQGPALTVDTAC
nidda_003_KS_002.seq      VFVGTNGQHYVPLLQSGG--ADGESFDGYISTGNSASVMSGRLSYVFGLEGPAVTVDTAC
nidda_004_KS_001.seq      VFAGAWESGYQKGLDAAD-----AGLEAQLLAGI-VSFTAGRVSYTLGLEGPAMTVDTAC
nidda_005_KS_001.seq      VFVGAAHTGYASDPARAP-----EGTEGYLLTGDADAVLSGRIAYVLGLEGPAITVETAC
monen_001_KS_002.seq      VFVGMTGQDYGPR-LHE----PSQATDGYLLTGSTPSVASGRLSFSFGLEGPALTVDTAC
monen_002_KS_001.seq      VYAGVSSEDYMSQLPRIP-----EGFEGHATTGSLTSVISGRVAYNYGLEGPAVTVDTAC
monen_003_KS_001.seq      VYAGAALPGFGT--PHID-----PAAEGHLVTGSAPSVLSGRLAYTFGLEGPAVTIDTAC
monen_003_KS_002.seq      VITGIMYDDYGSRFLAR----KPDGFEGRIMTGSTPSVASGRVAYTFGLEGPAITVDTAC
monen_004_KS_001.seq      VYIGAGVLGFGG--AQPD-----KTVEGHLLTGSALSVLSGRISFTLGLEGPSVSVDTAC
monen_004_KS_002.seq      TYVGLMYHDYAKSFPTAD-----AQLEGYSYLASTGSMVSGRVAYTLGLEGPAVTVDTAC
monen_005_KS_001.seq      VYVGAATTGNQT--QGDP-----GGKATEGYAGTAPSVLSGRLSFTLGLEGPAVTVETAC
monen_005_KS_002.seq      TYVGVMYHDYAAG-LAQ----DA-QLEGYSMLAGSGSVVSGRVAYTLGLEGPAVTVDTAC
monen_006_KS_001.seq      VYAGVSSQDHMSG-SRVP-----PEVEGYATTGTLSSVISGRIAYTFGLEGPAVTLDTAC
monen_007_KS_001.seq      TFIGCDGLDYALGASEVP-----EGTAGYFTIGNSGSVTSGRVAYTLGLEGPAVTVDTAC
monen_008_KS_001.seq      VFTGVSAHDYLTLISQTA-----SDVEGYIGTGNLGSVVSGRISYTVGLEGPAVTVDTAC
monen_008_KS_002.seq      VFTGLTIFDYLALVGEQP-----TEVEGYIGTGNLGCVASGRVSYVLGLEGPAMTIDTGC
oligo_001_KS_002.seq      VFVGASPSEYGPR-LHE----TS-DADGHVLTGTAPSVLSGRIAYVLGLEGPALTVDTAC
oligo_001_KS_003.seq      VFVGSNGQDYATL-LGE------SEVEGHVLTGTASSVLSGRIAYTLGLEGPALTVDTAC
oligo_001_KS_004.seq      VFVGAGSHDYGTLL-TSL----EG-GQDYALTGAVGSVLSGRIAYVLGLEGPALTVDTAC
oligo_002_KS_001.seq      VFVGMADQKYGPRDGELL-----DQVKGLVLTGTTSSVASGRIAYSLGLQGPAITIDTAC
oligo_002_KS_002.seq      VFAGAAGSDYGVVLEGAP----EA--DGYLTTGTAASVISGRIAYTLGLHGPAVTVDTAC
oligo_003_KS_001.seq      VFAGLNQQDYGTLLAAAP-----DGLDSYGSTGTSNSVLSGRISYVLGLEGPAVTVDTAC
oligo_003_KS_002.seq      VFVGAATSGYTSLFPRGS----QA-LAGYGVTGASTSVVSGRVSYVLGLEGPAVTVDTAC
oligo_004_KS_001.seq      VFTGLMYHDYASG-PGA----LPDEVEGYLSTGMAASVASGRISYFLGLEGPAVTLDTAC
oligo_004_KS_002.seq      VFAGVMYHDYGAD-QAV----LPEEVEGFIGTGSAGSVASGRIAFTLGLEGPAVTLDTAC
oligo_005_KS_001.seq      VYFGGTGGDFAGLLAASP----QA-LDGYMLTGTSSSVLSGRVSYVLGLEGPAVTVDTAC
oligo_006_KS_001.seq      VFIGGAPQEYGALVMNSA----EG-AGGYALTGAPGSVLSGRISYVLGLEGPAVTVDTAC
oligo_006_KS_002.seq      VFVGGGPQDYPAVLMGSA----EA-GGGYGMTGALGSVMSGRVSYVLGLEGPAVTVDTAC
oligo_006_KS_003.seq      VFVGGTAVEYGALLMNSP-----A-SQGYAVTGSSGSVMSGRISYTLGLEGPAVTVDTAC
oligo_007_KS_001.seq      VFVGGTAIEHTVKLMNSP-----T-DQGYAITGGSASVMSGRVSYVLGLEGPALTVDTAC
oligo_007_KS_002.seq      VFVGGTPQEYTTVLMNSS----EA-GGGYALTGASGSVMSGRVAYALGLEGPAVTVDTAC
rifam_001_KS_001.seq      VFSGLMGQGYG---SGA----VAPELEGFVTTGVASSVASGRVSYVLGLEGPAVTVDTAC
rifam_001_KS_002.seq      VFAGVNSHDYSMR-MHR----A-AGVEGFRLTGGSASVLSGRVAYHFGVEGPAVTVDTAC
rifam_001_KS_003.seq      VFTGIVHHDYVTR-LRE----VPEDVQGYTMTGTASSVASGRVAYVFGFEGPAVTVDTAC
rifam_002_KS_001.seq      VYAGVVHQDYAPDLSGFE-----GFMSLERALGTAGGVASGRVAYTLGLEGPAVTVDTMC
rifam_002_KS_002.seq      VFSGVSIHDYLESLSNMP-----AELEGFVTTATAGSVASGRVSYTFGFEGPAVTVDTAC
rifam_002_KS_003.seq      VFTGMFGQGYVAPGDSV----VTPELEGFAGTGGSSSVASGRVSYVFGFEGPAVTIDSAC
rifam_003_KS_001.seq      VFTGAGGSGYGGG--LTG-----PEMQSFAGTGLASSVASGRVSYVFGFEGPAVTIDTAC
rifam_004_KS_001.seq      VFTGVASMGYGAG-GGV----VAPELEGFVGTGAAPCIASGRVSYVLGFEGPAVTVDTGC
rifam_005_KS_001.seq      VFSGVFTQGYG---AGA----ITPDLEAFAGIGAASSVASGRVSYVFGLEGPAVTIDTAC
rifam_005_KS_002.seq      VFAGVSNQGYGMG--ADP-----AELAGYASTAGASSVVSGRVSYVFGFEGPAVTIDTAC
pikro_001_KS_002.seq      VFTGAMTHEYGPSLRD-----GGEGLDGYLLTGNTASVMSGRVSYTLGLEGPALTVDTAC
pikro_001_KS_003.seq      VFIGLSYQDYAARVPNAP-----RGVEGYLLTGSTPSVASGRIAYTFGLEGPATTVDTAC
pikro_002_KS_001.seq      VFTGVMYHDYATRLTD-----VPEGIEGYLGTGNSGSVASGRVAYTLGLEGPAVTVDTAC
pikro_002_KS_002.seq      VFAGTNGPHYEPLLRN-----TAEDLEGYVGTGNAASIMSGRVSYTLGLEGPAVTVDTAC
pikro_003_KS_001.seq      VFVGGWHTGYTSGQTTAV---QSPELEGHLVSGAALGFLSGRIAYVLGTDGPALTVDTAC
pikro_004_KS_001.seq      VYVGCGYQDYAPDIRVAP-----EGTGGYVVTGNSSAVASGRIAYSLGLEGPAVTVDTAC
Stamb_001_KS_002.seq      VFVGATAQDYGPRMHEPA-----ESSEGYLLTGTTASVASGRIAYAFALDGPAVTVDTAC
Stamb_001_KS_003.seq      VFAGVMYHDYAHA-AEA----LPE-TEGYRATGGAGSVVSGRIAYTYGLEGPAVTVDTAC
Stamb_001_KS_004.seq      VFVGTNGQDYAGLFPATG-----SGLEGHVATGSAASVLSGRISYTYGLEGPAVTVDTAC
Stamb_002_KS_001.seq      VFAATNGQDYATA-LAA----DPEAADGYLATGAVASVLSGRISYVFGLEGPALTVDTAC
Stamb_002_KS_002.seq      VFAGAASSYYGLDKALP------EDVAGYQLTGGATSVLSGRVAYSFGLEGPAVTVDTAC
Stamb_003_KS_001.seq      VFVGASPSGYDTVGRLP------ESTVGYQLTGSAGSVLSGRLSYVLGLEGPAVTVDTAC
Stamb_003_KS_002.seq      VLIGAASSGYGLGTDLF------TTAEGHVLAGGSNSVISGRVAYSFGLEGPAVTVDTAC
Stamb_003_KS_003.seq      VFIGAAHAQYGYGMRLP------DNALGHLMTGTTTSVASGRLAYTFGLEGPAVTVDTAC
Stamb_003_KS_004.seq      VFAGASSSGYGGAGDDL------EGAGGYLLAGTANSVISGRVAYTFGLEGPAVTVDTAC
Stamb_003_KS_005.seq      MFVGTNGQDYPVVLAGSA----DEGLDAHAATGNAAAVLSGRVSYAFGLEGPAVTVDTAC
Stamb_004_KS_001.seq      VFAGGNDQGYLRLLANE------PGSVGHQLTGGATAVISGRVAYTLGLEGPAVTLDTAC
Stamb_004_KS_002.seq      VYVGTATSGYGLGRFAVP-----DGSRPHVLTGTATSVVSGRLAYTFGLEGPAVTVDTAC
Stamb_004_KS_003.seq      VFVGTSFVGYGIGAQQPG-----NEAEGFFLAGTGTAAASGRISYTFGLEGPAATVDTAC
Stamb_005_KS_001.seq      VFAGVMYHDYPSVVDPEA-------LDGYLGTANAGSVLSGRIAYTFGLQGPAVTVDTAC
Stamb_006_KS_001.seq      VYVGAGTSGYGIG-VPV----A-EEAAGYALTGTATSVLCGRVAYSFGFEGPAVTVDTAC
Stamb_006_KS_002.seq      VFAGTNGQDYPIL-LAG----DPDISEGHQGAGNAAAVLSGRIAYTFGFEGPTLTVDTAC
Stamb_007_KS_001.seq      VFIGSSSSAYGTGLRTLP-----QGVEGHLLTGSAPSVVSGRVAYALGLEGLAVTVDTAC
Stamb_007_KS_002.seq      VFIGASPSGYGVG-VAV------PGSEGHSLTGMSGSVLSGRVAYVFGLEGPTLTVDTAC
Stamb_008_KS_001.seq      VFIGTTGQGYGINMQAGM-----AGTEGYQLTGSATAVLSGRIAYTLGLRGPAVSVDTAC
Stamb_008_KS_002.seq      VFAGLMYHDYGAT-VTV----LPEGVEGYLGTGTAGSVLAGRVSYTFGLEGPAMTVDTAC
Stamb_009_KS_001.seq      VFTGVMYNEYASRIMSVP----DE-VAGHLGTGNSGSVVSGRVAYTLGLEGPALTVDTAC
Stamb_009_KS_002.seq      VYAGTGGQDYLAV-LAG----DPVAGEGYLVTGGSPSVLSGRVAYAFGLEGPAVTVDTAC
```

```
Stamb_009_KS_003.seq      VFAGTNGQDYLAL-LLG----SAAATEGHLGTGNSASVLSGRISYTFGLRGPALSVDTAC
Stamb_009_KS_004.seq      VFVGASAGGYAMAGVLP------EGSESHAMTGTSNSVLSGRVSYVFGLEGPAVTVDTAC
                             .                            .        ::::     *   ::.: *

mycol_001_KS_001.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTDFSRQRGLAPDGRCKAFAATADG
mycol_001_KS_002.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTDFSRQRGLAPDGRCKAFAATADG
mycol_001_KS_003.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAIFTEFSRQRGLAPDGRCKAFAATADG
mycol_001_KS_004.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_001_KS_005.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_001_KS_006.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAIFTEFSRQRGLAPDGRCKAFAATADG
mycol_001_KS_007.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTDFSRQRGLAPDGRCKAFAATADG
mycol_001_KS_008.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAIFTEFSRQRGLAPDGRCKAFAATADG
mycol_002_KS_001.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_003_KS_001.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_003_KS_002.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_003_KS_003.seq      SSSLVAIHLACQSLRNNESQLALTGGVTVMSTPAIFTEFSRQRGLAPDGRCKAFAATADG
mycol_003_KS_004.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_003_KS_005.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_003_KS_006.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
mycol_003_KS_007.seq      SSSLVAIHLACQSLRNNESQLALAGGVTVMSTPAVFTEFSRQRGLAPDGRCKAFAATADG
eryth_001_KS_001.seq      SSSLVAVHLACQSLRRGESSLAMAGGVTVMPTPGMLVDFSRMNSLAPDGRCKAFSAGANG
eryth_001_KS_002.seq      SSSLVALHTACGSLRDGDCGLAVAGGVSVMAGPEVFTEFSRQGALSPDGRCKPFSDEADG
eryth_002_KS_001.seq      SSSLVALHLAVESLRKGESSMAVVGGAAVMATPGVFVDFSRQRALAADGRSKAFGAGADG
eryth_002_KS_002.seq      SSSLVGIHLAMQALRRGECSLALAGGVTVMSDPYTFVDFSTQRGLASDGRCKAFSARADG
eryth_003_KS_001.seq      SSGLTALHLAMESLRRDECGLALAGGVTVMSSPGAFTEFRSQGGLAADGRCKPFSKAADG
eryth_003_KS_002.seq      SSSLVALHSACGSLRDGDCGLAVAGGVSVMAGPEVFTEFSRQGGLAVDGRCKAFSAEADG
ampho_002_KS_001.seq      SSSLVALHMAAQSLRNGESSLALVGGVTILAEPSVFVEFSAQGGLAQDGRCKAFSDSADG
ampho_002_KS_002.seq      SSSLVAMHLACQSLRSGESSLALAGGAAVMATPHAFIGFSRQRALAENGRCKPFAEGADG
ampho_003_KS_001.seq      SSSLVSLHLAAQALRGGECSLALAGGASVMATPHPFIAFTSQSGLAADGRCKPFSDDADG
ampho_003_KS_002.seq      SSSLVGMHLAAQALRTGECTLALAGGVTVMSTPSTFVDFSRQRGLAADGRSKAFAEAADG
ampho_003_KS_003.seq      SSSLVALHWAAQALRSGECSLALAGGVTVMSTPSTFIEFSRQQGLSADGRSKAFAESADG
ampho_003_KS_004.seq      SSALVAMHLAAQALRSGECSLALAGGVTVMSIPDTFVEFSRQRGLSPDGRCKPFSESADG
ampho_003_KS_005.seq      SSSLVAMHWAMQALRSGEISLALAGGVTVMSTPGVFVDFARQRGLSPDGRCKAFSDSADG
ampho_003_KS_006.seq      SSSLVAMHLAAQALRTGECSLALAGGVTVISTPGTFVEFSRQRGLSPDGRCKAFSDSADG
ampho_004_KS_001.seq      SSSLVAVHLATQALRAGECTLALAGGVTVMSGPGTFIEMGRQGGLSVDGRCRSFGDTADG
ampho_004_KS_002.seq      SSSLVAMHWAMQALRSGEISLALAGGVTVMSTPAVFVDFARQRGLSPDGRCKAFSDSADG
ampho_004_KS_003.seq      SSSLVALHLACQALRNGESNLAVAGGATVMTTPNAFVAFSRQRALAQDGRCKAFSESADG
ampho_004_KS_004.seq      SSSLVAIHLACQSLRNGESTLALAGGASIMSTPMSFIGFSRQRALAEDGRCKAYAEGADG
ampho_004_KS_005.seq      SSSLVALHWAGHALSAGECDLVVAGGVSVMSTPDSFVEFSTQGGLAPDGHCKPFSDDADG
ampho_004_KS_006.seq      SSSLVALHLAAQALRSGECSIALAGGVTVMATPVGFVEFSAQGALSPDGRCRAFSEDANG
ampho_005_KS_001.seq      SSSLVALHLAVQSLRNGECTLALAGGVNVMSTPGSLVEFSRQGGLAGDGRCKAFSDSADG
ampho_005_KS_002.seq      SSSLVSLHLAAQALRSGECSMALVGGVTVMTTPANFAGFSRMGGLAQDGRCKAFSDSADG
ampho_005_KS_003.seq      SSSLVSLHLAAQALRSGECTMALAGGVTVMTTPANFTGFSKMGGLAHDGRCKAFSDSADG
ampho_006_KS_001.seq      SSALVALHLAAQSVRSGEADLALAGGVTVMTTSSSFVGFSLQGGLATDGRCKAFADSADG
averm_001_KS_001.seq      SASLVTLHLACQSLRSGECTLALAGGVSVMSTLGMFIEFSRQRGLSVDGRCKAYSAAADG
averm_001_KS_002.seq      SSSLVALHLACQALRAGECSMALAGGVTVMSSPGAFVEFSRQRGLAADGHCKAFSAAADG
averm_002_KS_001.seq      SSSLVALHLACQALRSGECTMALAGGASVMATPFVFTEFSRQRGLAADGRCKAFSAAADG
averm_002_KS_002.seq      SSSLVALHLACQALRSGECSMALAGGVTVMSTPAGFVEFSRQRGLAVDGRCKAFSAAADG
averm_002_KS_003.seq      SSSLVALHLACQSLRSGECTMALAGGATVMTTPITFTEFARQRGLAPDGRCKAFSAAADG
averm_002_KS_004.seq      SSSLVALHWAAQALRAGECSMALAGGVTVMSSPGTFVEFSRQRGLAADGRCKAYSAAADG
averm_003_KS_001.seq      SSSLVALHLACQSLRSGECTLALAGGVTVMSTPHLFVEFSRQRGLSVDGRCKSFAGGADG
averm_003_KS_002.seq      SSSLVALHLACQALRAGECSMALAGGVTVMSSPGAFVEFSRQRGLAADGHCKAFSAAADG
averm_003_KS_003.seq      SSSLVALHWAAQALRAGECSMALAGGVTVMSSPGTFVEFSRQRGLAADGRCKAYSAAADG
averm_004_KS_001.seq      SSSLVALHLACQSLRTGESSFALAGGVTVMSTPGMFVEFSRQRGLSPDGRCKAYSAAADG
averm_004_KS_002.seq      SSSLVALHLACQALRSGECSLALAGGVTVMSTPGAFVEFSRQRGLSPDGRCKAYGSGADG
averm_004_KS_003.seq      SSSLVALHLACQALRSGECSLALAGGVTVMSTPGAFVEFSRQRGLSPDGRCKAYGSGADG
borre_002_KS_001.seq      SSSLVAVHLAARSLRAGECDLALAGAVTVMSTPQLPVAFSRQRGLAPDGRSKAFAVSADG
borre_003_KS_001.seq      SSSLVALHLAARALRAGECELALAGGVTVMPTTRLFEVFSRQRGLAGDGRCKAFAAGADG
borre_003_KS_002.seq      STSLVGLHLAVQALRAGECELALAGGVTVMHTPRPFVEFSRQRGLAADGRSKAFAASADG
borre_004_KS_001.seq      SSSLVALHLAAQALRQGQCTLALAGGASVIAAPTVFATFSRQRGLAPDGRCKAFSSTTDG
borre_005_KS_001.seq      SSSLVGLHLAAQALRSGECGLALAGGVTVMSTPLSLVSFSRQRGLAQDGRSKAFSADADG
borre_006_KS_001.seq      SSSLVALHSAVQALRSGECELALAGGATVMATPTVFTAFSHQRGLAGDGRCKAFAAGADG
epoth_002_KS_001.seq      STSLVAVHLACMSLLDRECDMALAGGITVRIPHRAGYVYAEGGIFSPDGHCRAFDAKANG
epoth_003_KS_001.seq      SSSLVAIHLACRSLRAGESDLALAGGVSTLLSPDMMEAAARTQALSPDGRCRTFDASANG
epoth_003_KS_002.seq      SSSLVALHLGCQSLRRGECDQALVGGVNMLLSPKTFALLSRMHALSPGGRCKTFSADADG
epoth_003_KS_003.seq      SSSLVALHLACQSLRLGECDQALVGGVNVLLAPETFVLLSRMRALSPDGRCKTFSADADG
epoth_003_KS_004.seq      SSSLVAVHLACQALRRGECSVALAGGVALMLTPATFVEFSRLRGLAPDGRCKSFSAAADG
epoth_004_KS_001.seq      SSSLVAIHLACRSLRARESDLALAGGVNMLLSPDTMRALARTQALSPNGRCQTFDASANG
epoth_004_KS_002.seq      SSSLVALHLACQSLRLGECDQALAGGSSVLLSPRSFVAASRMRLLSPDGRCKTFSAAADG
epoth_005_KS_001.seq      SSSLVAIHLACRSLRARESDLALAGGVNMLLSSKTMIMLGRIQALSPDGHCRTFDASANG
lasal_001_KS_002.seq      SSSLVALHMACRALRAGECTMALAGGVTVMATPGMFVELSRQRALSPDGRCKSFSSAADG
lasal_001_KS_003.seq      SSSLVALHLACQSLRNGECELALAGGVAVMATPGVFVEFSRQRGLAADGRCKSFAAAADG
lasal_002_KS_001.seq      SSSVVALHSAVQALRNGECSLALAGGVTIMPNPVEFVEFSRQRVLSPDGRCKAFGATADG
lasal_002_KS_002.seq      SSSLVALHLAAQALRNGECELALAGGVAVMTTPVLFTEFSRQRGLAPDSRCKPFAAAADG
lasal_002_KS_003.seq      SSSLVALHLAAQALRNGECTMALASGVAVMATPLEFVEFSRQRVLSADGRCKAFAAAADG
lasal_003_KS_001.seq      SSSLVALHLATQALRRGECGAALVGGVTVMSNPSGFVEFSRQRGLAPDGRCKAFAAAADG
lasal_004_KS_001.seq      SSSLVALHLACQSLRQGECTTALAGGVTVMPTPTAFTEFSRQRGLAPDGRCKSFAAAADG
```

```
lasal_005_KS_001.seq      SSSLVALHLACRALRQEECSLALAGGVAVLALPTLFQEFSRQRGLAADGRCKSFAAAADG
lasal_005_KS_002.seq      SSSLVALHLACQALRNGDCSLALACGVTVMATPQVFTEFSRQRGLAADGRCKAFAAAADG
lasal_006_KS_001.seq      SSSLVALHWAEQALRARECSLAIVAGVTVLPSPLAFIEFSRQRGLAADGRCKPFSAAADG
lasal_007_KS_001.seq      SSSLVALHLACQALRNGECSLALAGGVTVMAVPGLFTEFSRQRGLAPDGRCKAFAAAADG
rapam_001_KS_001.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMPTPQSFVEFSRQRGLSADGRCKAYADAADG
rapam_001_KS_002.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFARQGGLAGDGRSKAFADSADG
rapam_001_KS_003.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMATPELFTEFSRQRGLASDGRCKAFADSADG
rapam_001_KS_004.seq      SSSLVALHLAAQALRNGECSMALAGGVTVMATPEVFTEFARQRGLASDGRCKAFADSADG
rapam_002_KS_001.seq      SSSLVALHQAAYALRQGECSLALVGGVTVMATPQSFVEFSRQRGLASDGRCKAFADSADG
rapam_002_KS_002.seq      SSSLVALYQAGYALRQGECSLALVGGVTVMATPQSFVEFSRKSGLDSDGRCKAFADSADG
rapam_002_KS_003.seq      SSSLVALHLAAQALRNGECSMALAGGVTVMATPEVFTEFARQRGLASDGRCKAFADSADG
rapam_002_KS_004.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFARQGGLAGDGRSKAFADSADG
rapam_002_KS_005.seq      SSSLVALHQAGYALRQGECSLALIGGVTVMATPHTFVEFSRQRGLASDGRCKAFADSADG
rapam_002_KS_006.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMPTPQTFVEFSRQRGLSADGRCKAFADAADG
rapam_003_KS_001.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMATPDVFTEFARQRGLATDGRSKAFADSADG
rapam_003_KS_002.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMATPQTFVEFSRQGGLASDGRCKAFADAADG
rapam_003_KS_003.seq      SSSLVALHQAGYALRQGECSLALVGGVTVMATPQSFVEFSRQRGLASDGRCKAFADSADG
rapam_003_KS_004.seq      SSSLVALHLASQALRQRECSLALVSGVSVMSSPAMFVEFSRQRGLSSDGRCKSFAASADG
nysta_002_KS_001.seq      SSSLVALHLAAQSLRVGESSLAVIGGVTILTEPSVFVEFSAQGGLAPDGRCKAFSDAADG
nysta_002_KS_002.seq      SSSLVALHLACQSLRDGESSLALAGGAAVMATPHAFVGFSRQRALAKDGRCKPFSDTADG
nysta_003_KS_001.seq      SSSLVALHWAVQALRGGECSMALAGGASIMATPGPFVAFTAQSGLAADGRCKPFSDRADG
nysta_003_KS_002.seq      SSSLVALHWAAQALRAGECSLALAGGVTVMSTPSTFVEFSRQRGLAPDGRSKAFAEAADG
nysta_003_KS_003.seq      SSSLVALHWAAQALRAGECSLALAGGVTVMSTPSTFVEFSRQRGLAPDGRSKAFAEAADG
nysta_003_KS_004.seq      SSSLVALHLAAQALRGGECTLALAGGVTVMSIPDTFIEFSRQRGLAPDGRSKPFSESADG
nysta_003_KS_005.seq      SSSLVAMHWAMQALRSGECGLALAGGVTVMSTPAVFVDFARQRGLSPDGRCKAFADAADG
nysta_003_KS_006.seq      SSSLVALHLAAQALRAGECTLALAGGVTVMSTPGTFVEFSRQRGLAPDGRSKAFAEAADG
nysta_004_KS_001.seq      SSSLVAVHLATQALRAGECTLALAGGVTIMSGPGTFIEMGRQGGLSADGRCRSFGDTADG
nysta_004_KS_002.seq      SSSLVAMHWAMQALRSGECGLALAGGVTVMSTPAVFVDFARQRGLSPDGRCKAFADAADG
nysta_004_KS_003.seq      SSSLVALHLACQSLRNGESNLAVAGGATIMTTPNPFIAFSRQRALAKDGRCKAFSDDADG
nysta_004_KS_004.seq      SSSLVAMHLACQSLRNGESSLALAGGVSIMSTPMSFVGFSRQRALAEDGRCKAYADGADG
nysta_004_KS_005.seq      SSSLVALHWAGRALRAGECDLVVAGGVSVMASPDSFVEFSTQGGLAPDGRCKAFSDAADG
nysta_004_KS_006.seq      SSSLVALHLAAQALRAGECSIALAGGVTVMATPVGFVEFSAQGALSPDGRCRAFSDDANG
nysta_005_KS_001.seq      SSSLVALHLAVQALRNGECGMALAGGVNVMATPGSLVEFSRQGGLARDGRCKAFADAADG
nysta_005_KS_002.seq      SSSLVSLHLAAQALRAGECTMALAGGVTVMTTAANFTGFSRMGGLAQDGRCKAFSDSADG
nysta_005_KS_003.seq      SSSLVALHWAVQALRAGECSLALAGGVTVMTTSTSFAGFTRQGGLAPDGHCKAFSDSADG
nysta_006_KS_001.seq      SSSLVTLHLAAQAVRAGECGLALAGGVTVMTTSSSFAGFSLQGGLAPDGRCKAFAEAADG
spino_002_KS_001.seq      SSSLVALHLAGQALRAGECDLALAGGVTVMSTPKVFLEFSRQRGLAPDGRCKSFAAGADG
spino_003_KS_001.seq      SSSLVALHLACQSLRQGECDLALAGGVTVMSTPERFVEFSRQRGLAPDGRCKSFAAAADG
spino_003_KS_002.seq      SSSLVALHLACQSLRSGECDLALAGGVTVMSTPEMFVEFSRQRGLAPDGRCKSFAESADG
spino_004_KS_001.seq      SSSLVALHLACQSLRSGECDLALAGGVTVMATPATFVEFSRQRGLAPDGRCKSFAAAADG
spino_004_KS_002.seq      SSSLVALHLAGQALRAGECEFALAGGVTVMSTPTTFVEFSRQRGLAPDGRCKSFAAAADG
spino_004_KS_003.seq      SSSLVALHLACQSLRSGECDMALAGGVTVMSTPLAFLEFSRQRGLAPDGRCKSFAAAADG
spino_005_KS_001.seq      SSSLVALHLACQSLRLGECDLALAGGISVMATPGAFVEFSRQRALASDGRCKPFADAADG
spino_005_KS_002.seq      SSSLVALHLAGQALRSGECDLALAGGVTVMATPGMFVEFSRQRGLAADGRCKSFAAAADG
spino_005_KS_003.seq      SSSLVALHLAGQSLRSGECDLALAGGVTVMSTPGTFVEFSRQRGLAPDGRCKSFAESADG
tylac_001_KS_002.seq      SSSLVALHLAVQALRGGECDMALAGGVTVMAGPGMFVEFSRQRGLAADGRCKAFADGADG
tylac_001_KS_003.seq      SSSLVALHNAAQALRGGECDMALAGGVTVMAAPGLFVEFSRQRGLAADGRCKAFADGADG
tylac_002_KS_001.seq      SSSLVSLHLAVRALRNGECDMALAGGATVLSTPTVLVDFSRQRGLAPDGRCKAFADSADG
tylac_003_KS_001.seq      SSSLVALHLAVQALRGGECDMALAGGVTVMAGPGMFVEFSRQRGLAADGRCKAFADGADG
tylac_003_KS_002.seq      SASLVALHLAVQAMRRGECDMALVGGATVMSTPEMLVEFSRQRVISANGRSRAFAAGADG
tylac_004_KS_001.seq      SSSLVALHNAAQALRRGECDLALAGGVTVMATPAVFVQFARQRGLAPDGRCKAFADAADG
tylac_005_KS_001.seq      SSSLVALHLAVRALRHGECGLALAGGVAVMADPAAFVEFSRQKGLAADGRCKAFSAAADG
nidda_001_KS_002.seq      SSSLVALHLAVQSLRRGECDLALAGGTTVMSGPGMFVEFSRQRGLAPDGRCKAFSADADG
nidda_001_KS_003.seq      SSSLVALHLAVQSLRRGECDLALAGGSTVMSTPTVFVEFSRQRGLAPDGRCKPFADTADG
nidda_002_KS_001.seq      SSSLVALHLAVQSLRRGECDLALAGGSTVMSTPTVFVEFSRQRGLAPDGRCKSFADTADG
nidda_003_KS_001.seq      SSSLVALHLAVQSLRRGECDLALAGGTTVMATPTVFVEFSRQRGLAPDGRCKPFADTADG
nidda_003_KS_002.seq      SASLTALHLSVQSLRRGECDLALVSGATVMATPEMLVDFARQRAISPDGRCKAFAEAADG
nidda_004_KS_001.seq      SSSLVALHLAVQSLRRRECDLALAGGATVIADPALLVQFSRQRALAPDGRCKAFAEAADG
nidda_005_KS_001.seq      SSSLVALHLAVQSLRRGECDLALAGGVAVMPDPTVFVEFSRQRGLAPDGRCKAFADSADG
monen_001_KS_002.seq      SSSLVTLHLAAQALRRGECDLALAGGATVLATPGMFTEFSRQRGLAPDGRCKPFAAGADG
monen_002_KS_001.seq      SASLVAIHLASQALRQRECDLALAGGVLVLSSPLMFTEFCRQRGLAPDGRCKPFAAAADG
monen_003_KS_001.seq      SSSLVAVHLAAHALRQRECDLALAGGVTVMTTPYVFTEFSRQRGLAADGRCKPFAAAADG
monen_003_KS_002.seq      SSSLVAMHLAAQALRQGECELALAGGVTVMATPNTFVEFSRQRGLAPDGRCKPFAAAADG
monen_004_KS_001.seq      SSSLVSMHLAAQALRQGECDLALAGGVTVMSTPGAFTEFSRQGALSPDGRSKAFAASADG
monen_004_KS_002.seq      SSSLVSIHLATQALRHGECDLALAGGVTVMADPDMFAGFSRQRGLSPDGRCKAYAAAADG
monen_005_KS_001.seq      SSSLVAMHLAANALRQGECDLALAGGVTVMSTPEVFTGFSRQRGLAPDGRCKPFAAAADG
monen_005_KS_002.seq      SSSLVSIHLAAQALRQGECTLALAGGVTVMATPEVFTGFSRQRGLAPDGRCKPFAAAADG
monen_006_KS_001.seq      SASLVAIHLACQALRQGDCGLAVAGGVTVLSTPTAFVEFSRQRGLAPDGRCKPFAEAADG
monen_007_KS_001.seq      SSSLVSLHLATQALRTQECSLALAGGTYVMSSPAPLIGFSELRGLAPDGRCKPFSASSDG
monen_008_KS_001.seq      SSSLVAIHLASQALRQGECSLALAGGSTVMATPGSFTEFSRQRGLAPDGRCKPFAAAADG
monen_008_KS_002.seq      SSSLVAIHQAAHALRQGECSLALAGGATVMATPGSFVEFSLQRGLAKDGRCKPFAAAADG
oligo_001_KS_002.seq      SSSLVALHLAAQALRGGECDLALAGGVTVMATSGMFAEFARQGGLAPDGRCKAFADGADG
oligo_001_KS_003.seq      SSSLVALHLAVQALNAGECDLALAAGVTVMSSPEIFVEFSRQRGLSADGRCKAFGPDADG
oligo_001_KS_004.seq      SSSLVALHLAAQALRGGECDVALAGGVAVMATPDAFDAFARQGGLAPDGRCKAFADGADG
oligo_002_KS_001.seq      SSSLVALHLAVRALRSGECSFALVGGAAVMAEPQLFAEMAEQGGMAGDGRCKAFAAAADG
oligo_002_KS_002.seq      SSSLVALHMAVGALHQGECDLALATGVSVISTPGAFVDFARQSGLAADGRCKAFAASADG
oligo_003_KS_001.seq      SSSLVTLHLAAQALRSGECDLALAGGATTLSTSAVHVALSGQRALAPDGRSKAFSAAADG
```

```
oligo_003_KS_002.seq      SSSLVALHTAAQALRSGDCDLALAGGVAVMTSPFLFDDFAKQGGLSPDGRCKAFAAGADG
oligo_004_KS_001.seq      SSSLVALHLAVQALRDEECDLALAGGATIMSTPATFVENSRQRGLARDGRCKSFAAAADG
oligo_004_KS_002.seq      SSSLVALHLAAQALRTGECDLALAGGVTVMSTPGVFVELSRQGGLSPDGRCRSFAESADG
oligo_005_KS_001.seq      SSSLVSMHLAVQALRKDEISLALAGGVSVLATPGAFPEFSRQGGLASDGRCKAFSSDADG
oligo_006_KS_001.seq      SSSLVALHLAVKSLRTGECDLALAGGVLVLVTPSIFAEFSATGGSAGDGRCKAFAASADG
oligo_006_KS_002.seq      SSSLVALHLAAQSLNNGECELALAGGVTIMATPGAFAEFDTLGGMAGDGRCKAFAASADG
oligo_006_KS_003.seq      SSSLVALHLAAQALRGGECSLALTGGVGLMATPGAFAEFDTLGGLSVDGRTKAFAAAADG
oligo_007_KS_001.seq      SSSLVALHSAAQSLRQGDCSLALAGGVAVMATSSAFVTFARQRGLAADGRCKAFSDDADG
oligo_007_KS_002.seq      SSSLVSLHLAGQALRGGECDLALVGGVTVMATPGAFVEFSRQGGLAGDGRCKAFAAGADG
rifam_001_KS_001.seq      SSSLVAMHLAAQALRQGECSMALAGGVTVMATPGSFVEFSRQRALAPDGRCKAFAAAADG
rifam_001_KS_002.seq      SSSLVALHMAVQALQRGECSMALAGGVMVMGTVETFVEFSRQRGLAPDGRCKAFADGADG
rifam_001_KS_003.seq      SSSLVAMHLAAQALRQGECSMALAGGATVMASPDAFLEFSRQRGLSADGRCKAYAEGADG
rifam_002_KS_001.seq      SSSLVAIHLAAQALRRGECSMALAGGSTVMATPGGFVGFARQRALAFDGRCKSYAAAADG
rifam_002_KS_002.seq      SSSLVAIHLAAQALRQGECTMALAGGVAVMGSPIGVIGMSRQRGMAEDGRVKAFADGADG
rifam_002_KS_003.seq      SSSLVAMHLAAQSLRQGECSMALAGGATVMANPGAFVEFSRQRGLAVDGRCKAFAAAADG
rifam_003_KS_001.seq      SSSLVAMHLAAQALRQGDCSMALAGGAMVMSGPDSFVVFSRQRGLATDGRCKAFASGADG
rifam_004_KS_001.seq      SSSLVAMHLAAQALRRGECSMALAGGAMVMAQPGSFVSFSRQRGLALDGRCKAFSDSADG
rifam_005_KS_001.seq      SSSLVAIHLAAQALRAGECSMALAGGATVMPTPGTFVAFSRQRVLAADGRSKAFSSTADG
rifam_005_KS_002.seq      SSSLVAMHLAGQALRQGECSMALAGGVTVMGTPGTFVEFAKQRGLAGDGRCKAYAEGADG
pikro_001_KS_002.seq      SSSLVALHLAVQALRKGEVDMALAGGVAVMPTPGMFVEFSRQRGLAGDGRSKAFAASADG
pikro_001_KS_003.seq      SSSLTALHLAVRALRSGECTMALAGGVAMMATPHMFVEFSRQRALAPDGRSKAFSADADG
pikro_002_KS_001.seq      SSSLVALHLAVQALRKGEVDMALAGGVTVMSTPSTFVEFSRQRGLAPDGRSKSFSSTADG
pikro_002_KS_002.seq      SSSLVALHLAVQALRKGECGLALAGGVTVMSTPTTFVEFSRQRGLAEDGRSKAFAASADG
pikro_003_KS_001.seq      SSSLVALHLAVQALRKGECDMALAGGVTVMPNADLFVQFSRQRGLAADGRSKAFATSADG
pikro_004_KS_001.seq      SSSLVALHLALKGLRNGDCSTALVGGVAVLATPGAFIEFSSQQAMAADGRTKGFASAADG
Stamb_001_KS_002.seq      SSSLVALHLAAQSLRAGECSTALAGGVTLMATPGVLVEFARQDGLSPDGRCKAFGARADG
Stamb_001_KS_003.seq      SSSLVALHLAARSLRAGECSMAVVGGVTVMSAPGVFVDFSKQRGLAPDGRCKSFSADADG
Stamb_001_KS_004.seq      SSSLVALHLAAQSLRSGECSMALAGGVTVMAAPTAFLEFAHQRGLAADGRCKAFSADADG
Stamb_002_KS_001.seq      SASLVALHLAVQSLRRGECTMALAGGVTVMATPAAFVEFSRQRGLAADGRCKAFGDGADG
Stamb_002_KS_002.seq      SSSLVAVHLAVQALHNGECDMALAGGVTVMAGFGVFSEFSRLDGLAGDARCKAFAEAADG
Stamb_003_KS_001.seq      SSSLVALHLAVQALRQGECGRALVGGVAVITTPAAFTEFGKQGGMASDGRCKSFSAEADG
Stamb_003_KS_002.seq      SSSLVAVHLAVQALRNGECDRALAGGVTVMAAPSIFAEFNRQGGLAGDGRCKAFAGAADG
Stamb_003_KS_003.seq      SSSLVALHLAVQALRNGECAMALAGGVTVLATPGVITEFDRQRGLASDGRCKAFSAAADG
Stamb_003_KS_004.seq      SSSLVALHWAAQALRNGECELALAGGVTVMVSPAAFAEFDRQDGLASDGRCKSFAGGADG
Stamb_003_KS_005.seq      SSSLVALHLAAQALRRGECDLALAGGVSVMSTEAAFTEFARQGGLADDGRCKAFSADADG
Stamb_004_KS_001.seq      SSSLVALHLACRSLRSGESTLALAGGVTVMATPGVFTEFTRQRGLAADGRCKSFSVDADG
Stamb_004_KS_002.seq      SSSLVALHLAIGALRRGECAMALAGGATVMAVPGIFTDAAQGGALAPGGRCRAFSSDAEG
Stamb_004_KS_003.seq      SSSAVAIHLACQALRNNECDLAVAGGAAVLATPASFTEFSRQRGLAADGRCKPFAAAADG
Stamb_005_KS_001.seq      SSSLVALHLAAQALRAGECELALVGGVTVMSGPMMFAGFGLEDGSAADGRCKAFAAAADG
Stamb_006_KS_001.seq      SSSLVALHLAVRALQAGECTMALAGGVTVMATPAAFVEFSRQRGLARDGRCKAFAAAADG
Stamb_006_KS_002.seq      SSSLVALHLAVRALRAGECSMALAGGVTVMSTPTAFVEFSRQLGLAADGRCKAFAAAADG
Stamb_007_KS_001.seq      SSSLVATHLAVQALRNGECDTALAGGVTVMTSPGIFTEFSKQRGLAADGRCKPFSADADG
Stamb_007_KS_002.seq      SSSLVALHLAARSLRDGECTMAVVGGVAVMTSPGVFSEFARQDGLASDGRCRAFSADADG
Stamb_008_KS_001.seq      SSSLTALHLAAQSIRAGDCSMAFAGGVTVMPHPIAFHEFNRQGGLASDGRCKAFSADADG
Stamb_008_KS_002.seq      SSSLVTLHLACQALRNGECDMALAGGVTVLSTPGVFVDFSRQRGLASDGRCRSFSADADG
Stamb_009_KS_001.seq      SSSLVALHLAMRSLRSGESSLALAGGVAIMVTPNTFFDFALSGGLASDGRCKAFSADADG
Stamb_009_KS_002.seq      SSSLVALHLACQSLRHQESDLALVGGVNVISLPSVFAEFSKQRGQAADGRCKAFSASADG
Stamb_009_KS_003.seq      SSSLVALHLAARSLRAGECSMALAGGVTAMPLPGTFIEFSTQGALSSDGRCKAFSADADG
Stamb_009_KS_004.seq      SSSLVALHLAAQSLRNGECSMALAGGVTVMHSPVVFAEFGRQGGLASDGRCRSFAASADG
                          *:. .  : .  .:   :   ..  .                    .: : :   ::*

mycol_001_KS_001.seq      TGWGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_001_KS_002.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_001_KS_003.seq      TGWGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_001_KS_004.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_001_KS_005.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_001_KS_006.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_001_KS_007.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_001_KS_008.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_002_KS_001.seq      TGWGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_003_KS_001.seq      TGWGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_003_KS_002.seq      TGWGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_003_KS_003.seq      TGWGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_003_KS_004.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_003_KS_005.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_003_KS_006.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
mycol_003_KS_007.seq      TGFGEGAAVLVLERLSEARRNNHPVLAIVAGSAINQDGA-SNGLTAPHGPSQQRVINQAL
eryth_001_KS_001.seq      FGMAEGAGMLLLERLSDARRNGHPVLAVLRGTAVNSDGA-SNGLSAPNGRAQVRVIQQAL
eryth_001_KS_002.seq      FGLGEGSAFVVLQRLSDARREGRRVLGVVAGSAVNQDGA-SNGLSAPSGVAQQRVIRRAW
eryth_002_KS_001.seq      FGFSEGVTLVLLERLSEARRNGHEVLAVVRGSALNQDGA-SNGLSAPSGPAQRRVIRQAL
eryth_002_KS_002.seq      FALSEGVAALVLEPLSRARANGHQVLAVLRGSAVNQDGA-SNGLAAPNGPSQERVIRQAL
eryth_003_KS_001.seq      FGLAEGAGVLVLQRLSAARREGRPVLAVLRGSAVNQDGA-SNGLTAPSGPAQQRVIRRAL
eryth_003_KS_002.seq      FGFAEGVAVVLLQRLSDARRAGRQVLGVVAGSAINQDGA-SNGLAAPSGVAQQRVIRKAW
ampho_002_KS_001.seq      TGWSEGAAVLVVERLSDARRAGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_002_KS_002.seq      MTLAEGVGVVLLEKLSDARRNGHRVLAVVRGSALNQDGA-SNGLTAPNGLSQQRVIRQAL
ampho_003_KS_001.seq      TGWAEGAGLFILMRLSDAQREGRKVLAVLRGSALNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_003_KS_002.seq      VAWSEGVGMLLLERRSDAIRNGHEILAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
```

```
ampho_003_KS_003.seq      VGWSEGVGMLLLERQSDAIRNGHEILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_003_KS_004.seq      VGWSEGVGMLLLERQSDAIRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_003_KS_005.seq      VGWSEGVGMLVLERQSDAIRNGHEILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_003_KS_006.seq      VGWSEGVGMLVLERQSDAIRNGHEILAVLRGSAINQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_004_KS_001.seq      TGWAEGVGILVLERLSDAIRNGREILAVVRGTAVNQDGA-SNGLTAPNGPSQQAVIEQAL
ampho_004_KS_002.seq      VGWSEGVGMLVLERQSDAIRNGHQILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_004_KS_003.seq      MTLAEGVGIVLVERLSDARRNGHPVLAVIRGSAINQDGA-SNGLSAPNGPSQQRVIRQAL
ampho_004_KS_004.seq      MTLAEGVGLILLERLSDARRNGHEVLAVIRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_004_KS_005.seq      TAWSEGVGILVLERLSDAIRNGHEVRGIIRGTAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_004_KS_006.seq      TGWSEGVGMLVVERLSDARRNGHRVLAVLRGSAINQDGA-SNGLTAPSGPAQQRVIRQAL
ampho_005_KS_001.seq      TGWSEGAAVLALERLSDAQRNGHPVLAVIRGSAVNQDGA-SNGFTAPNGPSQQRVIRQAL
ampho_005_KS_002.seq      TGWSEGAAVLVVERLSDARRAGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_005_KS_003.seq      TGWSEGAAVLVLERLSDARRAGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
ampho_006_KS_001.seq      TGWSEGVGMILVERLSEARRKGHPVLAVLRGSAVNQDGA-SNGLSAPNGPAQQRVIRDAL
averm_001_KS_001.seq      TGWGEGVGMLLVERLSDAVRLGHRVLAVVRGSAVNQDGA-SNGLTAPNGPAQERVIRQAL
averm_001_KS_002.seq      TGWGEGVGMLLVERLSDAHRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
averm_002_KS_001.seq      TGWSEGVGMLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGRSQVKVIRQAL
averm_002_KS_002.seq      TGWGEGVGMLLVERLSDARRLGHRILAVVRGSAVNQDGA-SNGLTAPNGPSQERVIRLAL
averm_002_KS_003.seq      TGWGEGVGMLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
averm_002_KS_004.seq      TGWAEGVGMLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
averm_003_KS_001.seq      TGMGEGVGMLLVERLSDAVRLGHRVLAVLRGSAVNQDGA-SNGLTAPNGPAQERVIRQAL
averm_003_KS_002.seq      TGWGEGVGMLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
averm_003_KS_003.seq      TGWAEGVGMLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
averm_004_KS_001.seq      TGMSEGVGILLLERLSEAERRGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
averm_004_KS_002.seq      VGWAEGVGVLLVERLSEAERRGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
averm_004_KS_003.seq      VGWAEGVGVLLVERLSEAERRGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
borre_002_KS_001.seq      MGFGEGVGVLVLERLSVARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
borre_003_KS_001.seq      TGWGEGVGVLVLERLSVARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
borre_003_KS_002.seq      VAWAEGAGILVLERLSAARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
borre_004_KS_001.seq      TGFGEGVGVLVLERLSDARRNGHEVLAVVRGSAVNQDGA-SSGFTAPNGPSQQDVIREAL
borre_005_KS_001.seq      MGMAEGVGVLVLERLSEARRNGHEVLAVLRSSAVNQDGA-SNGLSAPNGPAQQRVIQSAL
borre_006_KS_001.seq      AGFAEGVGVLVLERLSVARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
epoth_002_KS_001.seq      TIMGNGCGVVLLKPLDRALSDGDPVRAVILGSATNNDGARKIGFTAPSEVGQAQAIMEAL
epoth_003_KS_001.seq      FVRGEGCGLVVLKRLSDAQRDGDRIWALIRGSAINHDGR-STGLTAPNVLAQETVLREAL
epoth_003_KS_002.seq      YARAEGCAVVVLKRLSDAQRDRDPILAVIRGTAINHDGP-SSGLTVPSGPAQEALLRQAL
epoth_003_KS_003.seq      YARGEGCAVVVLKRLRDAQRAGDSILALIRGSAVNHDGP-SSGLTVPNGPAQQALLRQAL
epoth_003_KS_004.seq      VGWSEGCAMLLLKPLRDAQRDGDPILAVIRGTAVNQDGR-SNGLTAPNGSSQQEVIRRAL
epoth_004_KS_001.seq      FVRGEGCGLIVLKRLSDARRDGDRIWALIRGSAINQDGR-STGLTAPNVLAQGALLREAL
epoth_004_KS_002.seq      FARAEGCAVVVLKRLRDAQRDRDPILAVVRSTAINHDGP-SSGLTVPSGPAQQALLGQAL
epoth_005_KS_001.seq      FVRGEGCGMVVLKRLSDAQRHGDRIWALIRGSAMNQDGR-STGLMAPNVLAQEALLRQAL
lasal_001_KS_002.seq      TGWAEGAGVLLVERLSDARRNGHRVLAVVRGSAINQDGA-SNGLTAPNGLSQQRVIRQAL
lasal_001_KS_003.seq      TGWSEGVGLLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIHQAL
lasal_002_KS_001.seq      TGFSEGVGVLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
lasal_002_KS_002.seq      TAWSEGVGVLLVERLSDALRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
lasal_002_KS_003.seq      TGFSEGVGVLLVERLSDARRNGHQVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
lasal_003_KS_001.seq      FGPSEGAGVLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLSAPNGPSQQRVIRQAL
lasal_004_KS_001.seq      TNWAEGVAVLVVERLSDARRNGHRVLAVVRGTAINQDGA-SNGLSAPNDLAQERVIRSAL
lasal_005_KS_001.seq      TGFSEGVGMLLLERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPSGPSQQQVILQAL
lasal_005_KS_002.seq      TGFSEGVGVLLVERLSDARRNGHQVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
lasal_006_KS_001.seq      FGFAEGVGALLVQRLTDARREGREVLAVIRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
lasal_007_KS_001.seq      TGFSEGVGVLLVERLSDARRNGHQVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
rapam_001_KS_001.seq      TEWAEGVGVLLVERLSDAQAKGHQVLAVVRGSAVNQDGA-SNGLSAPNGPSQQRVIRAAL
rapam_001_KS_002.seq      AGFSEGVGVLLVERLSDARRNGHQVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
rapam_001_KS_003.seq      TGWAEGVGVLLVERLSDAQAKGHQVLAVVRSSAVNQDGA-SNGLTAPNGPSQQRVIQAAL
rapam_001_KS_004.seq      AGFSEGAGLLLVERLSDARRNGHQVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
rapam_002_KS_001.seq      TGWAEGAGVLLVERLSDAQRKGHQVLAVVRGSAVNQDGA-SNGLSAPNGPSQQRVIRAAL
rapam_002_KS_002.seq      TGWAEGVGVLLVERLSDAQAKGHQVLAVVRSSAVNQDGA-SNGLSAPNGPSQQGVIRQAL
rapam_002_KS_003.seq      AGFSEGAGLLLVERLSDARRNGHQVLAVVRGSAVNQDGA-SNGFTAPNGPAQQRVIRQAL
rapam_002_KS_004.seq      AGFSEGVGVLLVERLSDAQAKGHQVLAMLRSSAVNQDGA-SNGLTAPNGPSQQRVIQAAL
rapam_002_KS_005.seq      TGWSEGVGVLLVERLSDAQAKGHQVLAVVRSSAVNQDGA-SNGLSAPNGPSQQRVIRQAL
rapam_002_KS_006.seq      TGWAEGVGVLLVERLSDAQANGHQILAVVRSSAVNQDGA-SNGLSAPNGPSQQRVIRAAL
rapam_003_KS_001.seq      AGFSEGIGVLLVERLSDAEAKGHQVLAVVRSSAVNQDGA-SNGLTAPNGPSQQRVIQTAL
rapam_003_KS_002.seq      TGWAEGVGVLLVERLSDARRNGHQVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
rapam_003_KS_003.seq      TGWAEGVGVLLVERLSDAQAKGHQVLAVVRSSAVNQDGA-SNGLTAPNGPSQQRVIQAAL
rapam_003_KS_004.seq      TIWSEGVGVLVVERLSDARRLGHRVLATVRGSAVNSDGA-SNGLTAPNGTSQQRVIRQAL
nysta_002_KS_001.seq      TGWAEGVGVLVAERLSDAQRNGHPVLAVLRGSAVNQDGA-SNGLTAPNGPSQERVIQQAL
nysta_002_KS_002.seq      MTLAEGVGVVLLERLSHARANGHRVLAVIRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_003_KS_001.seq      TGWGEGAGMLVLMRLSDAQREGRPVLAVLRGSAINQDGA-SNGLTAPNGPSQQRVIRAAL
nysta_003_KS_002.seq      VAWSEGVGMLVLERQSDAVRNGHEILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_003_KS_003.seq      VGWSEGVGILVLERQSDAVRNGHEILAVIRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_003_KS_004.seq      VGWSEGVGMLLLERQSDAVRNGHQILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_003_KS_005.seq      VGWSEGVGVLVLERQSDAVRNGHEILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_003_KS_006.seq      VGWSEGVGILVLERQSDAVRNGHEILAVIRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_004_KS_001.seq      TGWAEGVGILVLERLSDAVRNGHEILAVVRGTAVNQDGA-SNGLTAPNGPSQQQVIQQAL
nysta_004_KS_002.seq      VGWSEGVGVLVLERQSDAVRNGHEILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_004_KS_003.seq      MTLAEGVGVVLVERLSDAQRNGHPVLAVLRGSAINQDGA-SNGLTAPNGPSQQRVIRQAL
nysta_004_KS_004.seq      MTLAEGVGLVLLERLSDARANGHQVLAVIRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
```

```
nysta_004_KS_005.seq      TAWSEGVGILVLERLSAARRNGHQVLGLIRGTAVNQDGA-SNGLTAPNGLSQQRVIAQAL
nysta_004_KS_006.seq      TGWSEGVGMLVVERLSDARRNGHRVLAVLRGSAINQDGA-SNGLTAPSGPAQQRVIRQAL
nysta_005_KS_001.seq      TGWSEGAGVLLLERLSDAQRNGHPVLAVVRGSAVNQDGA-SNGFTAPNGPSQQRVIRQAL
nysta_005_KS_002.seq      TGWSEGAAVLVLERLSDARRAGHRVLAVVRGSAVNQDGA-SNGLTAPNGPAQQRVIRQAL
nysta_005_KS_003.seq      TGWSEGVGVLVVERRSDALRNGHEILAVVRGSAVNQDGA-SNGLTAPNGPAQQRVIRQAL
nysta_006_KS_001.seq      TGWSEGIGLLLVERLSDAQRNGHPVLAVLRGSAVNQDGA-SNGLSAPNGPSQQRVIRQAL
spino_002_KS_001.seq      TGWGEGAGLLLLERLSDARRNGHEVLAVVRGSAVNQDGA-SNGLTAPNGSSQQRVITQAL
spino_003_KS_001.seq      TGWGEGAGLVLLERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGLAQERVIQQVL
spino_003_KS_002.seq      TGWGEGAGLLLLERLSDAHRNGHRVLAVVRGSAVNQDGA-SNGLAAPNGPSQQRVINQAL
spino_004_KS_001.seq      TGWGEGAGLVLLERLSDARRNGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVITQAL
spino_004_KS_002.seq      TGWGEGAGLVLLERLSDARRNGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVITQAL
spino_004_KS_003.seq      TGWGEGAGLVLLERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
spino_005_KS_001.seq      TGWGEGAGMLLLERLSDARRNGHPVLAAVVGSAINQDGT-SNGLTAPSGPAQQRVIRQAL
spino_005_KS_002.seq      TGWGEGAGLVLLERLSDARRNGHAVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVITQAL
spino_005_KS_003.seq      TGWGEGAGLVLLERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIQQAL
tylac_001_KS_002.seq      TAWAEGAGVVLVERLSDARRLGHPVLAVVCGSAVNQDGA-SNGLTAPSGPSQERVIRQAL
tylac_001_KS_003.seq      TAWAEGAGVVLVERLSDARRLGHPVLAVVCGSAVNQDGA-SNGLTAPSGPSQERVIRQAL
tylac_002_KS_001.seq      TSWAEGAGMLLLQRLSDARREGRPVLAVIRGSAVNQDGA-SNGLTAPNGRAQRQVIEDAL
tylac_003_KS_001.seq      TAWAEGAGVVLVERLSDARRLGHPVLAVVCGSAVNQDGA-SNGLTAPSGPSQERVIRQAL
tylac_003_KS_002.seq      VALGEGVGVLLVERLSDAERNGHPVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
tylac_004_KS_001.seq      FGPAEGVGMVLVERLSDARRLGHPVLAVVCGSAVNQDGA-SNGLTAPSGPSQERVIRQAL
tylac_005_KS_001.seq      TGWAEGVGVLVLERLSDARRAGHTVLGLVTGTAVNQDGA-SNGLTAPNGPAQQRVIAEAL
nidda_001_KS_002.seq      TAWAEGVGMLLVERLSDAERLGHRVLAVVRGTAVNQDGA-SNGLTAPSGPAQQQVIRDAL
nidda_001_KS_003.seq      TAWAEGVGVLLVERLSDAERLGHRVLAVVRGTAVNQDGA-SNGLTAPSGPAQQQVIRDAL
nidda_002_KS_001.seq      TAWAEGVGVLLVERLSDAERLGHRVLAVVRGTAVNQDGA-SNGLTAPSGPAQQQVIRDAL
nidda_003_KS_001.seq      TAWAEGVGMLLVERLSDAERLGHRVLAVVRGTAVNQDGA-SNGLTAPSGPAQQQVIRDAL
nidda_003_KS_002.seq      VGLAEGAVALLVERLSDAERLGHRVLAVVRGTAVNQDGA-SNGLTAPSGPAQQNVIRDAL
nidda_004_KS_001.seq      FGPAEGAGMLLVERLSDAERLGHRVLAVVRGTAVNQDGA-SNGLTAPSGPAQQKVIRDAL
nidda_005_KS_001.seq      TAWAEGVGMLLVERLSDAERLGHRVLAVVRGTAVNQDGA-SNGLTAPSGPAQQKVIRDAL
monen_001_KS_002.seq      TGWAEGVGLVLLERLSEARRKGHAVLAVIRGSAINQDGA-SNGLTAPNGPSQQRVIRAAL
monen_002_KS_001.seq      TGFSEGIGLLLLERLSDARRNGHKVLAVIRGSAVNQDGA-SNGLTAPNDAAQEQVIRAAL
monen_003_KS_001.seq      TAFSEGAGLLVLERLSDARRAGHRVLAVIRGSAVNQDGA-SNGLTAPNGPAQQRVIRAAL
monen_003_KS_002.seq      TGWGEGAGLVVLERLSDARRKGHRVLALLRGSAVNQDGA-SNGMTAPNGPSQERVIRTAL
monen_004_KS_001.seq      TGFSEGAGLLLLERLSDARRNGHKVLAVIRGSAVNQDGA-SNGLTAPNGPSQERVIRAAL
monen_004_KS_002.seq      VGFSEGVGVLLLERLSDARRHGRRVLGVVRGSAVNQDGA-SNGLTAPNGPSQERVIRQAL
monen_005_KS_001.seq      TGWGEGAGLILLERLSDARRKGHKVLAVIRGSAINQDGA-SNGFTAPNGPSQRRVIRQAL
monen_005_KS_002.seq      TGWGEGVGVLLLERLSDARRHGRRVLGVVRGSAVNQDGA-SNGLTAPNGPSQERVIRQAL
monen_006_KS_001.seq      TGFSEGVGLILLERLSDARRNGHQVLGVVRGSAVNQDGA-SNGLTAPNDVAQERVIRQAL
monen_007_KS_001.seq      MGMAEGTGVVLLERLSDARRKGHKVLAVIRGSAINQDGA-SNGLTAPNGPAQERVIRAAL
monen_008_KS_001.seq      TGWGEGAGVVALELLSEARRRGHKVLAVIRGSATNQDGT-SNGLAAPNGPSQERVIRAAL
monen_008_KS_002.seq      TGWAEGVGLVVLERLSEARRNGHNVLAVIRGSAINQDGT-SNGLTAPNGQAQQRVIRQAL
oligo_001_KS_002.seq      TGWGEGVGMLAVQRLSDAVRDGRPVLAVVRGSAVNSDGA-SNGLTAPNGPSQQRVIRQAL
oligo_001_KS_003.seq      TGWGEGVGTVLLERLSDARRLGHDVLGVVRGTAVNQDGA-SNGLSAPNGPAQQRVIRQAL
oligo_001_KS_004.seq      TGWGEGVGVLVLTRLSEARRRGHDVLAVVRGSAVNSDGA-SNGLTAPNGPSQQRVIRQAL
oligo_002_KS_001.seq      TGWGEGVGVLLLQPLSTAREQGLPVLATVRGTAVNQDGA-SNGLTAPNGPAQRRVIRKAL
oligo_002_KS_002.seq      TNWAEGVGVLVVERLSDARRNGHRVLATVRGSAVNSDGA-SNGLSAPNGGAQQRVIRQAL
oligo_003_KS_001.seq      AGWSEGVGVLAVERLSDARRLGHRVLAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
oligo_003_KS_002.seq      TGWAEGVGVLAVERLSDARRLGHRVLAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
oligo_004_KS_001.seq      VGWGEGSALLVVERLSDARRRGHQVLAIVRGSAVNQDGA-SNGLTSPNGPSQQRVIQQAL
oligo_004_KS_002.seq      TGWGEGAGLLLVERLSDARRNGHPVLAVVRGTAVNQDGA-SNGLTAPNGPAQQRVIEQAL
oligo_005_KS_001.seq      TGWGEGVGVLVLERLSDARRAGHPVLAVVRETGINQDGA-SNGLTAPSGPAQRRLILRTL
oligo_006_KS_001.seq      TGWGEGAGVLAVQRLTDARRDGHPVLAVIRGSAVNQDGA-SNGLSAPNGPSQQRVIRQAL
oligo_006_KS_002.seq      TGWGEGVGMVVVERLSDARRNGHEVLAVIRGSAINQDGA-SNGLSAPNGPSQQRVIRQAL
oligo_006_KS_003.seq      IGWGEGVGMVVVERLSDARRNGHEVLAVIRGSAINQDGA-SNGLSAPNGPSQQRVIRQAV
oligo_007_KS_001.seq      IGWGEGAAVVLLERLADARRNGHPVLAVIRGSAVNQDGA-SNGLSAPNGPSQQRVIRQAV
oligo_007_KS_002.seq      TGWGEGVGMLAVQRLSDAVRDGRPVLAVVRGSAVNSDGA-SNGLTAPNGPSQQRVIRQAL
rifam_001_KS_001.seq      TGWSEGVGVVVLERLSVARERGHRILAVLRGSAVNQDGA-SNGLTAPNGLSQQRVIRRAL
rifam_001_KS_002.seq      TGWSEGVGLLLVERLSEAQRRGHQVLAVVRGSAVNSDGA-SNGLTAPNGPSQQRVIRKAL
rifam_001_KS_003.seq      TGWAEGVGVVVLERLSVARERGHRVLAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRGAL
rifam_002_KS_001.seq      SGWAEGVGVLLLERLSVARERGHQVLAVIRGSAVNQDGA-SNGLTAPNGPAQQRVIRKAL
rifam_002_KS_002.seq      TVLSEGVGIVVLERLSVARERGHRVLAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRSAL
rifam_002_KS_003.seq      TGWAEGVGVVILERLSVARERGHRILAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRRAL
rifam_003_KS_001.seq      MVLAEGISVVVLERLSVARERGHRVLAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
rifam_004_KS_001.seq      MGLAEGVGVIALERLSVARERGHRVLAVLRGIAVNQDGA-SNGLTAPNGPSQQRVIRAAL
rifam_005_KS_001.seq      TGWAEGAGVLVLERLSVAQERGHRILAVLRGSAVNQDGA-SNGLTAPNGPSQQRVIRKAL
rifam_005_KS_002.seq      TGWAEGVGVVVLERLSVARERGHRVLAVLRGSAVNSDGA-SNGLTAPNGPSQQRVIRRAL
pikro_001_KS_002.seq      TSWSEGVGVLLVERLSDARRNGHQVLAVVRGSALNQDGA-SNGLTAPNGPSQQRVIRRAL
pikro_001_KS_003.seq      FGAAEGVGLLLVERLSDARRNGHPVLAVVRGTAVNQDGA-SNGLTAPNGPSQQRVIRQAL
pikro_002_KS_001.seq      TSWSEGVGVLLVERLSDARRKGHRILAVVRGTAVNQDGA-SSGLTAPNGPSQQRVIRRAL
pikro_002_KS_002.seq      FGPAEGVGMLLVERLSDARRNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRRAL
pikro_003_KS_001.seq      FGPAEGAGVLLVERLSDARRNGHRILAVVRGSAVNQDGA-SNGLTAPHGPSQQRVIRRAL
pikro_004_KS_001.seq      LAWGEGVAVLLLERLSDARRKGHRVLAVVRGSAINQDGA-SNGLTAPHGPSQQHLIRQAL
Stamb_001_KS_002.seq      TGWSEGAGMLVLRRLSDALADGQHVLAVLRGSAVNSDGA-SNGLTAPNGLAQRRVIRAAL
Stamb_001_KS_003.seq      TGWSEGAGVLLVERLSDARRHGHDVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRAAL
Stamb_001_KS_004.seq      TGWGEGAGLVLLERLSDARAQGHRVLGLLRGSAVNQDGA-SNGLTAPNGPAQQRVIRAAL
Stamb_002_KS_001.seq      TGWGEGVGVLLVERLSDAVRNGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRDAL
Stamb_002_KS_002.seq      TGWSEGVGLLLVERLSDAVRRGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQERVIRQAL
```

```
Stamb_003_KS_001.seq      TGWSEGVGVLLVERLSDARRNGHDVLAVIRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
Stamb_003_KS_002.seq      TGWAEGVGVLVLERLSDAVRRGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQERVIRQAL
Stamb_003_KS_003.seq      TGMSEGVGLLLVERLSDAERLGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQERVIRQAL
Stamb_003_KS_004.seq      TGWGEGVGVLVLERLSDARRNGHEILAVVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
Stamb_003_KS_005.seq      TGWGEGVGVLLVERLADARRNGHRVLALVRGSAVNQDGA-SNGLTAPNGPSQQRVIRQAL
Stamb_004_KS_001.seq      TGWSEGVGVLLVERLSDAVRNGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQQRVILQAL
Stamb_004_KS_002.seq      TGWGEGVGMLLVERLSDARRNGHPVLAVVRGSAVNQDGA-SNGLTAPHGPAQQRVIRAAL
Stamb_004_KS_003.seq      TGWGEGVGMILVERLSDALRAGHPVLALIRGTAVNQDGA-SNGLTAPSGRAQQKVIRQAL
Stamb_005_KS_001.seq      TGWGEGVGVLLVERLSDARRHGHRVLAVVRGSAVNQDGA-SSGLTAPNGPAQQRVIRQAL
Stamb_006_KS_001.seq      TGWSEGVGVLLVERLSDARRHGHRVLAVVRGSAVNSDGA-SNGLTAPNGPAQRRVILDAL
Stamb_006_KS_002.seq      TGWGEGVGVLLVERLSDARRNGHRVLAVVRGSAINSDGA-SNGLTAPNGPAQRRVIRRAL
Stamb_007_KS_001.seq      TGWSEGAGVLVLRRLSDARRRGDRVLAVIKGSAVNQDGA-SNGLSAPNGPSQERVIRAAL
Stamb_007_KS_002.seq      TGWGEGVGLLLVERLSEARRNGHEVLAVVRGSAVNSDGA-SNGLTAPNGPSQQRVIRAAL
Stamb_008_KS_001.seq      TGWSEGVGVLLVERLSDARRNGHPVLALIRGSAVNQDGA-SNGLTAPSGPAQERVIRAAL
Stamb_008_KS_002.seq      TGWAEGVGVLLVERLSDAVRNGHEVLAVVRGSAVNQDGA-SNGLTAPNGPSQERVIRAAL
Stamb_009_KS_001.seq      TGWAEGVGVLLVERLSDAVRNGHEVLAVVRGSAVNQDGA-SSGLTAPNGPSQERVIRAAL
Stamb_009_KS_002.seq      AGWGEGVGVLLVERLSDARRRGHEVLAVVRGSAVNSDGA-SNGLTAPNGPAQQRVIRAAL
Stamb_009_KS_003.seq      TGWGEGVGLLLLERLSDAERNGHRVLAVVRGSAVNQDGA-SNGLTAPNGPSQERVIRAAL
Stamb_009_KS_004.seq      TGWGEGVGLVLLERLSDARRNGHRVLGVVRGSAINQDGA-SNGLTAPNGPAQQRVIRAAL
                             .:*   .       *     : . :   . * **  . *:  *   .*   :  .

mycol_001_KS_001.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEAGALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_001_KS_002.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEAGALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_001_KS_003.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEAGALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_001_KS_004.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_001_KS_005.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEAGALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_001_KS_006.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_001_KS_007.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEAGALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_001_KS_008.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_002_KS_001.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_003_KS_001.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_003_KS_002.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_003_KS_003.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_003_KS_004.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_003_KS_005.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_003_KS_006.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
mycol_003_KS_007.seq      ANAGLTHDQVDAVEAHGTGTTLGDPIEASALHATYGHHHTPDQPLWLGSIKSNIGHTQA-
eryth_001_KS_001.seq      AESGLGPADIDAVEAHGTGTRLGDPIEARALFEAYGRDR--EQPLHLGSVKSNLGHTQA-
eryth_001_KS_002.seq      ARAGITGADVAVVEAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQA-
eryth_002_KS_001.seq      ESCGLEPGDVDAVEAHGTGTALGDPIEANALLDTYGRDRDADRPLWLGSVKSNIGHTQA-
eryth_002_KS_002.seq      AASGVPAADVDVVEAHGTGTELGDPIEAGALIATYGQDR--DRPLRLGSVKTNIGHTQA-
eryth_003_KS_001.seq      ENAGVRAGDVDYVEAHGTGTRLGDPIEVHALLSTYGAERDPDDPLWIGSVKSNIGHTQA-
eryth_003_KS_002.seq      ARAGITGADVAVVEAHGTGTRLGDPVEASALLATYGKSRGSSGPVLLGSVKSNIGHAQA-
ampho_002_KS_001.seq      ANAGLRPGDVDAVEAHGTGTPLGDPIEAQALIATYGSDRDPQQPLLLGSVKSNIGHTQS-
ampho_002_KS_002.seq      ANAGLGTADIDVVEAHGTGTKLGDPIEAGALLATYGQGRDGGQPLLLGSVKSNIGHTQA-
ampho_003_KS_001.seq      AHAGLQASDVDAVEAHGTGTPLGDPIEAQALLATYGQNR--ERPLLLGSVKSNIGHAQA-
ampho_003_KS_002.seq      ASGGLSPEDVDVVEAHGTGTTLGDPIEAQALIATYGRDRDPERPLWLGSVKSNIGHTQA-
ampho_003_KS_003.seq      ASGGLSTGDVDAVEAHGTGTTLGDPIEAQALIATYGQNRTDEQPLWLGSLKSNIGHTQA-
ampho_003_KS_004.seq      ASGGLSTGDVDAVEAHGTGTTLGDPIEAQALLATYGQNRPEDRPLLLGSVKSNIGHTQA-
ampho_003_KS_005.seq      ASGGLSAQEIDVVEAHGTGTTLGDPIEAQALLATYGRDRDPEQPLWLGSVKSNIGHTQA-
ampho_003_KS_006.seq      ASGGLSTGDVDVVEAHGTGTTLGDPIEAQALLATYGRDRDPEQPLLLGSVKSNIGHTQA-
ampho_004_KS_001.seq      VNARLSAGDIDVVEAHGTGTTLGDPVEAQALLATYGQQRDEDKPLLLGSVKSNISHTQA-
ampho_004_KS_002.seq      ASGGLSAGDVDVVEAHGTGTTLGDPIEAQALLATYGRDRDPEQPLLLGSVKSNIGHTQA-
ampho_004_KS_003.seq      ANARVAPGEIDLLEAHGTGTPLGDPIEAQALFATYGRTRTPETALLLGSVKSNIGHSQS-
ampho_004_KS_004.seq      ANAGVEANDIDVLEGHGTGTALGDPIEAQALFATYGKDRDPERPVLLGSVKSNIGHTQM-
ampho_004_KS_005.seq      ADARLAPSEIDAVEAHGTGTTLGDPIEAQALLNTYGPGRDTDQPLLLGTVKSNIGHTQA-
ampho_004_KS_006.seq      ANARLEAADVDAVEAHGTGTKLGDPIEAQALLATYGQDR--ERPVLVGSLKANIGHTQA-
ampho_005_KS_001.seq      SNAGLNPADVDVVEAHGTGTPLGDPIEAQSILATYGQDR--EQPLLLGSIKSNIGHTQSA
ampho_005_KS_002.seq      ANAGLRPGDVDAVEAHGTGTPPGDPIEAQALLATYGSDRDPQQPLLLGSVKSNIGHTQA-
ampho_005_KS_003.seq      ANAGLRPGDVDAVEAHGTGTPLGDPIEAQALIATYGSDRDPQQPLLLGSVKSNIGHTQS-
ampho_006_KS_001.seq      ASAGLSPADVDAVEAHGTGTTLGDPIEAQALLATYGQDRDASRPLRLGTVKSNIGHTQA-
averm_001_KS_001.seq      ANAGLSVADVDVVEGHGTGTTLGDPIEAQALLATYGQR-AGDRPLWLGSLKSNIGHTMA-
averm_001_KS_002.seq      ANAGLSAGDVDAVEAHGTGTTLGDPIEAQALLATYGQDRAGEGPLWLGSVKSNVGHTQA-
averm_002_KS_001.seq      ANAHLSPADVDAVEAHGTGTTLGDPIEAQALVEAYGQDRPNGRPLWLGTLKSNIGHSMA-
averm_002_KS_002.seq      ANADLTPADVDAVEAHGTGTTLGDPIEAQALLATYGQDRPGNEPLWLGSMKSNIGHAQA-
averm_002_KS_003.seq      ANADLTPADVDAVEAHGTGTTLGDPIEAQAILATYGQDRPGNGPLWLGSVKSNVGHTQA-
averm_002_KS_004.seq      ANAGLTPADVDAVEGHGTGTTLGDPIEAQALLAAYGQHRPHHRPLWLGSLKSNIGHAQA-
averm_003_KS_001.seq      ANAGLSVADVDVVEGHGTGTTLGDPIEAQALLATYGQR-AGNRPLWLGSVKSNIGHAQA-
averm_003_KS_002.seq      ANAGLSAGDVDAVEAHGTGTTLGDPIEAQALLATYGQDRAGEGPLWLGSVKSNVGHTQA-
averm_003_KS_003.seq      ANAGLTPADVDAVEGHGTGTTLGDPIEAQALLAAYGQHRPHHRPLWLGSLKSNIGHAQA-
averm_004_KS_001.seq      ACAGLSVADVDVVEGHGTGTTLGDPIEAQALLATYGQ-RAGDTPVWLGSVKSNIGHAQA-
averm_004_KS_002.seq      ACAGLSVADVDVVEGHGTGTTLGDPIEAQALLATYGQGRSGERPVWLGSVKSNIGHAQA-
averm_004_KS_003.seq      ACAGLSVADVDVVEGHGTGTTLGDPIEAQALLATYGQGRSGERPVWLGSVKSNIGHAQA-
borre_002_KS_001.seq      ASAGLGPADVDVVEAHGTGTRLGDPIEAQALLATYGRGRDAERPLWLGSVKSNIGHAQA-
borre_003_KS_001.seq      ASARLAPEDVDAVEAHGTGTSLGDPIEAQALLATYGRGRDAERPLWLGSVKSNIGHAQA-
borre_003_KS_002.seq      ASAGLGPADVDVVEAHGTGTALGDPIEAQALLATYGRGRDADRPLWLGSVKSNIGHTQA-
borre_004_KS_001.seq      ADGRLTPADVDVVEGHGTGTRLGDPIEAQALLATYGRGRDADRPLWLGSVKSNIGHTQA-
```

```
borre_005_KS_001.seq      TVGRLAPSDIDVVEAHGTGTALGDPIEAQALLATYGRGRDADRPLWLGSVKSNIGHTQA-
borre_006_KS_001.seq      ANARLAPEDVDAVEGHGTGTSLGDPIEAQALLATYGRGRDAERPLWLGSVKSNIGHAQA-
epoth_002_KS_001.seq      ALAGVEARSIQYIETHGTGTLLGDAIETAALRRVFDRDASTRRSCAIGSVKTGIGHLES-
epoth_003_KS_001.seq      RSAHVEAGAVDYVETHGTGTSLGDPIEVEALRATVGPARSDGTRCVLGAVKTNIGHLEA-
epoth_003_KS_002.seq      AHAGVVPADVDFVECHGTGTALGDPIEVRALSDVYGQARPADRPLILGAAKANLGHMEP-
epoth_003_KS_003.seq      SQAGVSPVDVDFVECHGTGTALGDPIEVQALSEVYGPGRSEDRPLVLGAVKANVAHLEA-
epoth_003_KS_004.seq      EQAGLAPADVSYVECHGTGTTLGDPIEVQALGAVLAQGRPSDRPLVIGSVKSNIGHTQA-
epoth_004_KS_001.seq      RNAGVEAEAIGYIETHGAATSLGDPIEIEALRTVVGPARADGARCVLGAVKTNLGHLEG-
epoth_004_KS_002.seq      AQAGVAPAEVDFVECHGTGTALGDPIEVQALGAVYGRGRPAERPLWLGAVKANLGHLEA-
epoth_005_KS_001.seq      QSARVDAGAIDYVETHGTGTSLGDPIEVDALRAVMGPARADGSRCVLGAVKTNLGHLEG-
lasal_001_KS_002.seq      ADARLSGADVDVVEAHGTGTRLGDPIEADALLATYGQDRPADRPLRLGSVKSNIGHTQC-
lasal_001_KS_003.seq      ANARLTPGEVDAVEAHGTGTRLGDPIEAQALLATYGQERAEDRPLWLGSVKSNMGHAQA-
lasal_002_KS_001.seq      ANARLAPADVDAVEAHGTGTALGDPIEAQALLATYGQDRPTEQPLWLGSIKSNIGHTQA-
lasal_002_KS_002.seq      ANARLTPADVDAVEAHGTGTTLGDPIEADALLATYGQDRPEDRPLWLGSVKSNIGHTQA-
lasal_002_KS_003.seq      ANARLTPADVDAVEAHGTGTALGDPIEAQALLATYGQDRPTEQPLWLGSIKSNIGHTQA-
lasal_003_KS_001.seq      RNARLSPEQVDAVEAHGTGTTLGDPIEAQALLATYGQGRPEDRPLWLGSVKSNIGHTSA-
lasal_004_KS_001.seq      DNAGLTASDVDAVEAHGTGTTLGDPIEAQALLAAYGHERPAHRPLRVGSLKSNIGHAGP-
lasal_005_KS_001.seq      ANARLSGADVDAVEAHGTGTTLGDPIEAQALLATYGQERADGRPLWLGSVKSNIGHTQA-
lasal_005_KS_002.seq      ANARLTPADVDAVEAHGTGTTLGDPIEAQALLATYGQNRTEDQPLWLGSVKSNIGHTQC-
lasal_006_KS_001.seq      TNARLTPGDIDMVEAHGTGTALGDPIEAQALLATYGQDRPQDRPLRLGSVKSNIGHTQA-
lasal_007_KS_001.seq      ANARLTPADVDAVEAHGTGTTLGDPIEAQALLATYGQNRTEDQPLWLGSVKSNIGHTQA-
rapam_001_KS_001.seq      SNAGLAPHEVDVVEAHGTGTTLGDPIEAQAVIATYGQDRE--QPVLLGSLKSNIGHAQA-
rapam_001_KS_002.seq      SNAGLSTAEVDVVEAHGTGTTLGDPIEAQALMATYGQDRE--QPLLLGSVKSNLGHTQA-
rapam_001_KS_003.seq      SNAGLAAHEVDVVEAHGTGTTLGDPIEAQAVIATYGQDRE--RPLLLGSLKSNIGHAQA-
rapam_001_KS_004.seq      SNAGLSTADVDVVEAHGTGTTLGDPIEAQALLATYGQDRE--QPLLLGSLKSNIGHTQA-
rapam_002_KS_001.seq      SNAGLSTAEVDVVEAHGTGTTLGDPIEAQALLATYGQDRE--QPLLLGSVKSNLGHTQA-
rapam_002_KS_002.seq      ANAGLTTAEVDVVEAHGTGTTLGDPIEAQAVIATYGQDRE--RPLLLGSLKSNIGHAQA-
rapam_002_KS_003.seq      ANAGLTTAEVDVVEAHGTGTTLGDPIEAQAVIATYGQDRE--QPLLLGTLKSNVGHTQA-
rapam_002_KS_004.seq      SNAGLAAHEVDVVEAHGTGTTLGDPIEAQALLATYGQDRE--QPLLLGSVKSNLGHTQA-
rapam_002_KS_005.seq      ANAGLTTAEVDVVEAHGTGTTLGDPIEAQAVIATYGQDRD--QPVLLGSVKSNVGHTQA-
rapam_002_KS_006.seq      SNAGLAPHEVDVVEAHGTGTTLGDPIEAQAVIATYGQGRG--EPLLLGSLKSNVGHTQA-
rapam_003_KS_001.seq      SNAGLTTAEVDVVEGHGTGTTLGDPIEAQAVIATYGQDRE--QPLLLGSLKSNIGHTQA-
rapam_003_KS_002.seq      SNAGLSTAEVDVVEAHGTGTTLGDPIEAQALIATYGQDRD--QPVLLGSVKSNLGHTQA-
rapam_003_KS_003.seq      SNAGLAAHEVDVVEAHGTGTTLGDPIEAQALIATYGQDRE--RPLLLGSLKSNIGHAQA-
rapam_003_KS_004.seq      ANAGLTASDVDVVEAHGTGTKLGDPIEAEAILATYGQER--SAPAWLGSLKSNIGHAMA-
nysta_002_KS_001.seq      ARTGLTPADIDAVEAHGTGTRLGDPIEAQALLATYGQGHTPDQPLWLGSLKSNIGHTQA-
nysta_002_KS_002.seq      ANAGLTGADVDAVEAHGTGTKLGDPIEAQALLATYGQDRDAERPLLLGSVKSNIGHTQA-
nysta_003_KS_001.seq      DSAHLTAADIDAVEAHGTGTTLGDPIEAQALLATYGQDR--PRPLWLGSVKSNIGHTQA-
nysta_003_KS_002.seq      ASGGLSTADVDAVEAHGTGTTLGDPIEAQALLATYGRDRDPENPLLLGSIKSNIGHTQA-
nysta_003_KS_003.seq      ASGGLSTADVDAVEAHGTGTTLGDPIEAQALLATYGRDRDPENPLWLGSLKSNIGHTQA-
nysta_003_KS_004.seq      ASGGLSTADVDAVEAHGTGTTLGDPIEAQALLATYGRDRDPENPLLLGSIKSNLGHTQA-
nysta_003_KS_005.seq      ASGGLTAGDVDVVEAHGTGTTLGDPIEAQALLATYGRDREPERPLLLGSVKSNLGHTQA-
nysta_003_KS_006.seq      ASGGLSTADVDAVEAHGTGTTLGDPIEAQALLATYGRDRDPENPLLLGSIKSNLGHTQA-
nysta_004_KS_001.seq      VNARLAAGDIDVVEAHGTGTTLGDPVEAQALLATYGQNRPADRPLLLGSVKSNLSHTQA-
nysta_004_KS_002.seq      ASGGLTAGDVDVVEAHGTGTTLGDPIEAQALLATYGRDREPERPLLLGSVKSNLGHTQA-
nysta_004_KS_003.seq      ANARLAPGDIDALEAHGTGTPLGDPIEAQALFATYGRDRDPESALLLGSVKSNIGHTQS-
nysta_004_KS_004.seq      ANSAVAPGDIDVLEGHGTGTALGDPIEAQALLATYGQDRAPERPLLLGSVKSNIGHTQM-
nysta_004_KS_005.seq      ADARLRPADIDAIEAHGTGTTLGDPIEARALITAYGRDRDAERPLLLGTVKSNIGHTQA-
nysta_004_KS_006.seq      ANARLRPADIDAVEAHGTGTRLGDPIEAQALLATYGQDR--ERPVLLGSLKSNIGHTQA-
nysta_005_KS_001.seq      ANAGLATGDIDAVEAHGTGTPLGDPIEAQSILATYGQDR--AHPVLLGSIKSNMGHTQA-
nysta_005_KS_002.seq      ANAGLTPVDVDAVEAHGTGTPLGDPIEAQALIAAYGTDRDPEHPLLLGSVKSNIGHTQS-
nysta_005_KS_003.seq      ANAGLAPGDVDAVEAHGTGTVLGDPIEAQALLATYGQDRPADRPLWLGSVKSNIGHTQA-
nysta_006_KS_001.seq      AGAGLVPGDVDAVEAHGTGTRLGDPIEAGALLATYGQDRPADRPLWLGSVKSNLGHTQA-
spino_002_KS_001.seq      ASAGLSVSDVDAVEAHGTGTRLGDPIEAQALIATYGRDRDPGRPLWLGSVKSNIGHTQA-
spino_003_KS_001.seq      TSAGLSASDVDAVEAHGTGTRLGDPIEAQALIAAYGQDRDRDRPLWLGSVKSNIGHTQA-
spino_003_KS_002.seq      ANAALSASDVDAVEAHGTGTRLGDPIEAQALIATYGQARERDRPLWLGSVKSNIGHTQA-
spino_004_KS_001.seq      ASAGLSVSDVDAVEAHGTGTTLGDPIEAQALIATYGQGREKDRPLWLGSVKSNIGHTQA-
spino_004_KS_002.seq      TSAGLSVSDVDAVEAHGTGTRLGDPIEAQALIATYGRDRDPGRPLWLGSVKSNIGHTQA-
spino_004_KS_003.seq      ANAGLSASDVDVVEAHGTGTGLGDPIEAQALIATYGQERDPERALWLGSIKSNIGHTQA-
spino_005_KS_001.seq      ANAGLSPAEVDVVEAHGTGTALGDPIEAQALIATYGANRSADHPLLLGSLKSNIGHTQA-
spino_005_KS_002.seq      ASAGLSVSDVDAVEAHGTGTRLGDPIEAQALIATYGQGRDSDRPLWLGSVKSNIGHTQA-
spino_005_KS_003.seq      ASAGLSVSDVDAVEAHGTGTRLGDPIEAQALIATYGRDRDPGRPLWLGSVKSNIGHTQA-
tylac_001_KS_002.seq      GNARLTVADVDVVEAHGTGTRLGDPIEAQALLGTYGRDRDGGRPVWLGSLKSNIGHAQA-
tylac_001_KS_003.seq      ANARLTVADVDVVEAHGTGTRLGDPIEAQALLGTYGRDRDAECPVWLGSLKSNIGHAQA-
tylac_002_KS_001.seq      RDAGVGPDQVDAVEAHGTGTELGDPIEAGALLATYGTARTAERPLWLGSLKSNIGHTQA-
tylac_003_KS_001.seq      ANARLTVADVDVVEAHGTGTRLGDPIEAQALLGTYGRDRDGGRPVWLGSLKSNIGHAQA-
tylac_003_KS_002.seq      ADAGLRPEDIDAVEAHGTGTELGDPIEAEALLATYGRTRTADRPLWLGSLKSNIGHTQA-
tylac_004_KS_001.seq      GNARLTVADVDVVEAHGTGTRLGDPIEAQALLGTYGRDRDGGRPVWLGSLKSNIGHAQA-
tylac_005_KS_001.seq      ADAGLSPEDVDAVEAHGTGTRLGDPIEAGALLAASGRNRSGDHPLWLGSLKSNIGHAQA-
nidda_001_KS_002.seq      SDAGLSADDIDAVEAHGTGTALGDPIEAGALLATYGHPK-RQTPVWLGSLKSNIGHTQA-
nidda_001_KS_003.seq      SDAGLSADDIDAVEAHGTGTPLGDPIEAEALLATYGHPE-RQTPVWLGSLKSNIGHSQA-
nidda_002_KS_001.seq      SDAGLSADDIDAVEAHGTGTPLGDPIEAGALLATYGHPE-RQTPVWLGSLKSNIGHTQA-
nidda_003_KS_001.seq      ADAHLTPDDIDAVEAHGTGTPLGDPIEAGALLATYGHPK-RQTPVWLGSLKSNIGHTQA-
nidda_003_KS_002.seq      ADAGLTAADIDAVEAHGTGTHLGDPIEAGALLATYGQAE-DRGPLWLGSLKSNIGHTQA-
nidda_004_KS_001.seq      ADAGLRADDIDAVEAHGTGTPLGDPIEAGALLATYGHPE-RRKPLWLGSLKSNIGHTQA-
nidda_005_KS_001.seq      ADAHLTPDDIDAVEAHGTGTPLGDPIEAGALLATYGHPE-RQTPVWLGSLKSNIGHTQA-
monen_001_KS_002.seq      AAARLTADEVDVVEAHGTGTTLGDPIEAQALLATYGQGRSAERPLWLGSVKSNIGHTQA-
```

```
monen_002_KS_001.seq      DNARLTPSEVDAVEAHGTGTKLGDPIEAGALLATYGQHR--ARPLLLGSLKSNIGHTHA-
monen_003_KS_001.seq      AGARLSPAEVDAVEAHGTGTRLGDPIEADALLATYGQERHGGRPLWLGSVKSNIGHTQG-
monen_003_KS_002.seq      AGAGRGPEDIDVVEAHGTGTTLGDPIEAQALLATYGQGRPEDRPLWLGSVKSNIGHTQA-
monen_004_KS_001.seq      ANAGLGAAEVDAVEAHGTGTKLGDPIEAGALLATYGRDRDEDRPLWLGSVKSNIGHPQG-
monen_004_KS_002.seq      ASGGLSSVDVDVVEGHGTGTTLGDPIEAQALLATYGQGRPEDRPLWLGSVKSNIGHTQA-
monen_005_KS_001.seq      SSAHLSTSEIDVVEAHGTGTRLGDPIEAEALIATYGKEREDDRPLWLGSVKSNIGHTQA-
monen_005_KS_002.seq      ASGGLSSVDVDVVEGHGTGTTLGDPIEAQALLATYGQGRPVDRPLWLGSVKSNIGHTQA-
monen_006_KS_001.seq      TNARVTPDAVDAVEAHGTGTTLGDPIEGNALLATYGKDRPADRPLWLGSVKSNIGHTQA-
monen_007_KS_001.seq      ANARLAPEDIDAVEAHGTGTTLGDPIEAGALISAYGRERPEDRPLWVGAVKSNIGHTQI-
monen_008_KS_001.seq      ANARLSAEDIDAVEAHGTGTTLGDPIEAQALIATYGQGRPEDRPLWLGSVKSNIGHTQA-
monen_008_KS_002.seq      ANARLSAEDVDAVEAHGTGTMLGDPIEASALVATYGKERPADRPLWLGSIKSNIGHAQA-
oligo_001_KS_002.seq      ASAGLSAADVDMVEAHGTGTSLGDPIEAQALLATYGQDRPDDRPLWLGSVKSNIGHTQA-
oligo_001_KS_003.seq      ADAGCVPSDVDAVEAHGTGTRLGDPIEAQALLATYGRDRPADRPLWLGSVKSNIGHTQA-
oligo_001_KS_004.seq      ASAGLSAADVDMVEAHGTGTSLGDPIEAQALLATYGQDRPADRPLWLGSVKSNIGHTQA-
oligo_002_KS_001.seq      ADAQLSVGQVDAVEAHGTGTRLGDPIEAGALLATYGQDRAGDEPLWLGSVKSNIGHTQA-
oligo_002_KS_002.seq      AAAGLSADDVDAVEAHGTGTTLGDPIEAQAVLATYGRDRPAGQPLWLGSLKSNIGHSGA-
oligo_003_KS_001.seq      ANAGVSPAEVDVVEAHGTGTSLGDPIEADALLATYGQAKTAGRPLWLGSLKSNIGHSGA-
oligo_003_KS_002.seq      ERSGLTAPDVDAVEAHGTGTTLGDPIEAQAILATYGRDRTEDRPLLLGSLKSNLGHSQA-
oligo_004_KS_001.seq      ASARLGVADVDVVEAHGTGTTLGDPIEAQALLATYGRERSAEQPLWLGSVKSNIGHTQA-
oligo_004_KS_002.seq      ANARLGGADVDVVEAHGTGTTLGDPIEAQALLATYGRERSAEQPLWLGSVKSNLGHTQA-
oligo_005_KS_001.seq      DDAGLATADIDVVEAHGTGTKLGDPIEARALLATYGQDRPADRPLWLGSIKSNIGHTQY-
oligo_006_KS_001.seq      ANAGLSAADVDMVEAHGTGTTLGDPIEAEALLATYGQDRPADRPLWLGTLKSNIGHTQA-
oligo_006_KS_002.seq      ANAGLSAADVDMVEAHGTGTTLGDPIEAQALLATYGQDRPADRPLWLGSVKSNFGHTGA-
oligo_006_KS_003.seq      ANAGLALADVDMVEAHGTGTTLGDPIEAQALLATYGQERPADRPLWLGSVKSNIGHTGA-
oligo_007_KS_001.seq      ANAGLALADVDMVEAHGTGTSLGDPIEAQALLATYGQDRHDGRPLWLGSLKSNIAHTQA-
oligo_007_KS_002.seq      ASAGLSAADVDMVEAHGTGTSLGDPIEAQALLATYGQDRPADRPLWLGSVKSNIGHTQY-
rifam_001_KS_001.seq      AAAGLAPSDVDVVEAHGTGTTLGDPIEAQALLATYGQER--KQPLWLGSLKSNIGHAQA-
rifam_001_KS_002.seq      AAAGLSTSDVDAVEAHGTGTTLGDPIEAEALLATYGQNRE--TPLWLGSVKSNLGHTQA-
rifam_001_KS_003.seq      ASAGLAPSDVDVVEGHGTGTALGDPIEVQALLATYGQER--EQPLWLGSLKSNLGHTQA-
rifam_002_KS_001.seq      ASAGLTPSDVDTVEGHGTGTVLGDPIEVQALLATYGQGRDPQQPLWLGSVKSVVGHTQA-
rifam_002_KS_002.seq      AGAGLQPSEVDVVEAHGTGTALGDPIEAQALLATYGKSR--ETPLWLGSLKSNIGHTQA-
rifam_002_KS_003.seq      VSAGLAPSDVDVVEAHGTGTTLGDPIEAQALLATYGKDR--ESPLWLGSLKSNIGHAQA-
rifam_003_KS_001.seq      ANAGIGPSDVDLVEAHGTGTSLGDPIEAQALLATYGQDR--ETPLWLGSLKSNIGHTQA-
rifam_004_KS_001.seq      AEAGLSPSDVDAVEGHGTGTTLGDPIEAQALLATYGKGRDPEKPLWLGSVKSNLGHTQA-
rifam_005_KS_001.seq      AGAGLVASDVDVVEAHGTGTALGDPIEAQALLATYGQGR--ERPLWLGSVKSNFGHTQA-
rifam_005_KS_002.seq      AGAGLEPSDVDIVEGHGTGTALGDPIEAQALLATYGKDRDPETPLWLGSVKSNFGHTQS-
pikro_001_KS_002.seq      ADARLTTSDVDVVEAHGTGTRLGDPIEAQALIATYGQGRDDEQPLRLGSLKSNIGHTQA-
pikro_001_KS_003.seq      ADARLAPGDIDAVETHGTGTSLGDPIEAQGLQATYGKERPAERPLAIGSVKSNIGHTQA-
pikro_002_KS_001.seq      ADARLTTSDVDVVEAHGTGTRLGDPIEAQAVIATYGQGRDGEQPLRLGSLKSNIGHTQA-
pikro_002_KS_002.seq      ADARLTTADVDVVEAHGTGTRLGDPIEAQALIATYGQGRDTEQPLRLGSLKSNIGHTQA-
pikro_003_KS_001.seq      ADARLAPGDVDVVEAHGTGTRLGDPIEAQALIATYGQEKSSEQPLRLGALKSNIGHTQA-
pikro_004_KS_001.seq      ADARLTSSDVDVVEGHGTGTRLGDPIEAQALLATYGQGRAPGQPLRLGTLKSNIGHTQA-
Stamb_001_KS_002.seq      DAAGLRPSDVDMVEAHGTGTRLGDPIEAQALLATYGQERDAERPLWLGSLKSNIGHTQS-
Stamb_001_KS_003.seq      ADAGLAPGDVDMVEAHGTGTKLGDPVEAEALLATYGQVRDAERPLWLGSLKSNIGHTQA-
Stamb_001_KS_004.seq      NSAGLWPSEVDAVEAHGTGTRLGDPIEAEALLATYGRDRS--EPLYLGSVKSNIGHTQA-
Stamb_002_KS_001.seq      ASAGLAPSDVDLVEAHGTGTRLGDPIEAQALLAAYGRNRAAGHPLWLGSVKSNIAHTQA-
Stamb_002_KS_002.seq      ANARLEASDVDVVEAHGTGTRLGDPIEAQALLATYGQGRHEDRPLWLGSVKSNIGHTQS-
Stamb_003_KS_001.seq      ANAQLTTVDVDAVEAHGTGTALGDPIELEALQDTYGRDRTDDRPLWLGSAKSNLGHTQA-
Stamb_003_KS_002.seq      ANARLEASDVDVVEAHGTGTRLGDPIEAQALLATYGQGRDEDRPLWLGSVKSNIGHTQA-
Stamb_003_KS_003.seq      ANARLAASEVDAVEAHGTGTRLGDPIEAQALLATYGQGRDEDRPLWLGSVKSNIGHTQA-
Stamb_003_KS_004.seq      ATAGLAPGEVDVVEAHGTGTRLGDPIEAQALLATYGQGRDEDRPLWLGSVKSNIGHTQA-
Stamb_003_KS_005.seq      ADARLSPSEVDAVEAHGTGTRLGDPIEAQALLATYGQDRDADRPLWLGSIKSNLGHTQA-
Stamb_004_KS_001.seq      ADARLSASDVDVVEAHGTGTRLGDPIEAQALLAAYGQDRPEGRPLLLGSVKSNIGHAQA-
Stamb_004_KS_002.seq      ADAGLDTADVDAVEAHGTGTELGDPIEVQALMATYGRGRDEEQPLWLGSLKSNIGHTQS-
Stamb_004_KS_003.seq      ANSGLDAADIDTVEAHGTGTTLGDPIEAQSILATYGQNRAEGRPLWLGSVKSNIGHTQS-
Stamb_005_KS_001.seq      ASAALVPAEVDAVEAHGTGTTLGDPIEAQALLATYGQERA--EPLYLGSIKSNIGHTQA-
Stamb_006_KS_001.seq      AAARLSAAEIDAVEAHGTGTTLGDPIEAHALLATYGQERA--QPLYLGSLKSNIGHTQA-
Stamb_006_KS_002.seq      ADADLRHSDVDVVEAHGTGTKLGDPIEAQALLATYGQERA--EPLYLGSLKSNIGHTQA-
Stamb_007_KS_001.seq      ENARLAAADIDVVEAHGTGTTLGDPIEAQALLNAYGQDRPADRPLWLGSIKSNIGHAQA-
Stamb_007_KS_002.seq      ATAGLAPHEVDLVEAHGTGTKLGDPIEAQALLDVYGQERDPERPLWLGSLKSNIGHAQA-
Stamb_008_KS_001.seq      DDARLTPDQVDVVEAHGTGTTLGDPIEAQALLATYGRGR--DRPLWLGSLKSNIGHAQS-
Stamb_008_KS_002.seq      AAARLTAADIDVVEAHGTGTTLGDPIEAQALLNAYGQDRPADRPLWLGSLKSNIGHTQA-
Stamb_009_KS_001.seq      AAARLTAADIDVVEAHGTGTTLGDPIEAQALLNAYGQDRPADRPLWLGSLKSNIGHTQA-
Stamb_009_KS_002.seq      ASAGLEATDVDLVEAHGTGTRLGDPIEAEALLATYGRDRS--EPLYLGSVKSNIGHTQA-
Stamb_009_KS_003.seq      ADARLTAAEVDAVEAHGTGTKLGDPIEAQALLATYGQGRE--EPLYLGSVKSNIGHTQA-
Stamb_009_KS_004.seq      AAAGLEAAEVDAVEAHGTGTKLGDPIEAQALLATYGQDR--EEPLYLGSVKSNIGHTQA-
                                    :  :* **:.*  ** :*  .:  .           :*: *: ..*

mycol_001_KS_001.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_001_KS_002.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_001_KS_003.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_001_KS_004.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_001_KS_005.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_001_KS_006.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_001_KS_007.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_001_KS_008.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_002_KS_001.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
```

```
mycol_003_KS_001.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_003_KS_002.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_003_KS_003.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_003_KS_004.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_003_KS_005.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_003_KS_006.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
mycol_003_KS_007.seq      AAGAAGVVKMIQAITHATLPAT-LHVDQPSPHIDWSSGTVRLLTEPIQW-PNTD---HPR
eryth_001_KS_001.seq      AAGVAGVIKMVLAMRAGTLPRT-LHASERSKEIDWSSGAISLLDEPEPW-PAG---ARPR
eryth_001_KS_002.seq      AAGVAGVIKVLLGLERGVVPPM-LCRGERSGLIDWSSGEIELADGVREWSPAAD---GVR
eryth_002_KS_001.seq      AAGVTGLLKVVLALRNGELPAT-LHVEEPTPHVDWSSGGVALLAGNQPW-RRG---ERTR
eryth_002_KS_002.seq      AAGAAGVIKVVLAMRHGMLPRS-LHADELSPHIDWESGAVEVLREEVPW-PAGE---RPR
eryth_003_KS_001.seq      AAGVAGVMKAVLALRHGEMPRT-LHFDEPSPQIEWDLGAVSVVSQARSW-PAG---ERPR
eryth_003_KS_002.seq      AAGVAGVIKVVLGLNRGLVPPM-LCRGERSPLIEWSSGGVELAEAVSPWPPAAD---GVR
ampho_002_KS_001.seq      AAGVAGLVKMVMAMHNGTLPRT-LHLTEPSTHVDWSLGAVRLLTEETAW-PETG---RVR
ampho_002_KS_002.seq      AAGVAGVIKMVMAMHAGELPGT-LHIDEPSSHVDWTSGAVTLLRERTEW-PAVG-D-RPR
ampho_003_KS_001.seq      ASGAAGVIKMIMAIQHGVLPRS-LHAERPTSNVDWTAGAVELLDRAVDW-PET---GRAR
ampho_003_KS_002.seq      AAGVAGVIKMIMSMRHGLLPQT-LHVDAPSSHVDWDAGAVELLTEQTEW-PEN---DRVR
ampho_003_KS_003.seq      AAGVAGVIKMVMSMRHGVLPCT-LHVDAPSSHVDWTEGAVELLTEQTEW-PET---DRVR
ampho_003_KS_004.seq      AAGAAGVIKMVMAMQHGVLPRS-LNISEPSSHVDWSAGAVELLTEQTEW-PET---DRVR
ampho_003_KS_005.seq      AAGVAGVIKMVMSMRHGVLPRT-LHVDAPSSHVDWTEGAVELLTEQTEW-PET---EHVR
ampho_003_KS_006.seq      AAGVAGVIKMVMSMRHGVLPRT-LHVDAPSSHVDWTEGAVELLTEQTEW-PET---EHVR
ampho_004_KS_001.seq      AAGVAGVIKMVMAMRHGVLPRT-LLADEPTRHVDWSQGAVRVLTENTEW-PATG---APR
ampho_004_KS_002.seq      AAGVAGVIKMVMSMRHGVLPRT-LHVDAPSSHVDWTEGAVELLTEQTAW-PET---GRPR
ampho_004_KS_003.seq      AAGVASIIKMVMALRHGVMPQT-LHADEPSSHVDWSPGTVRLLGENTDW-PQTG---RPR
ampho_004_KS_004.seq      ASGVASIIKLVHALREGVAPKS-LHIDQPSTHVDWSSGTIQLLTERTEW-PETG---RPR
ampho_004_KS_005.seq      AAGVAGVIKMLLAMQHGTLPRT-LHVTSPTSHVDWSSGAVSLLTEERDW-PETG---RPR
ampho_004_KS_006.seq      ASGVGGVIKMVLAMQHGLLPRS-LYTENPSSHVDWTAGNARLLTELTPW-PESE---RVR
ampho_005_KS_001.seq      ASGVAGIMKMIMAMRNEVLPKT-LHVDRPSTHVDWTAGKVELLTENRPW-PTAP--DRPR
ampho_005_KS_002.seq      AAGVAGLVKMVMAMRNGVLPRT-LHITEPSTHVDWSLGAVQVLTEETAW-PETG---RVR
ampho_005_KS_003.seq      AAGAAGLVKMVMAMHQGTLPRT-LHVTEPSTHVDWSLGAVRLLTEETAW-PETG---RVR
ampho_006_KS_001.seq      AAGAAGVIKMVLALRHGLLPRT-LHIDAPSTHVDWDAGHVSLLTEATPW-PEGE--Q-TR
averm_001_KS_001.seq      AAGVGGVIKMVMALREGVLPRT-LHVDKPSPQVDWSAGAVRLLTEAVPW-PGDAA-GRLR
averm_001_KS_002.seq      AAGVAGVIKMVMALRHGLLPRT-LHVDEPSPHVDWSAGAVQLLTETVPW-PGG-E-GRLR
averm_002_KS_001.seq      AAGVGGVIKMVMALRNGLLPRT-LHVDEPSPHVDWSAGAVQLLTETVPW-PGG-E-GRLR
averm_002_KS_002.seq      AAGVGGVIKMVMALRNGLLPRT-LHVDEPSPHVDWSAGAVQLLTETVPW-PGG-E-GRLR
averm_002_KS_003.seq      AAGVAGVIKMVMALRHRTLPPT-LHADEPSPHVDWSAGAVQLLTETVPW-PGG-E-GRPR
averm_002_KS_004.seq      AAGVGGVIKMVMALRNGLLPQT-LHVDEPTPQVDWSTGAVQLLTQPVPW-PADPA-GRPR
averm_003_KS_001.seq      AAGVGGVIKMVMALREGVLPRT-LHVDEPSPQVDWSAGAVRLLTEAVPW-PGDAA-GRLR
averm_003_KS_002.seq      AAGVAGVIKMVMALRNGLLPRT-LHVDEPSPHVDWSAGAVQLLTETVPW-PGG-E-GRLR
averm_003_KS_003.seq      AAGVGGVIKMVMALRNGLLPQT-LHVDEPTPQVDWSTGAVQLLTQPVPW-PADPA-GRPR
averm_004_KS_001.seq      AAGVAGVIKMVMALRAGVLPRT-LHVDEPSSQVDWSSGSVRVLADEVEW-PGV-E-GRLR
averm_004_KS_002.seq      AAGVAGVIKMVMALRAGVLPRT-LHVDEPSSQVDWSSGSVRVLADEVEW-PGV-E-GRLR
averm_004_KS_003.seq      AAGVAGVIKMVMALNHELLPTS-LHIDEPSPHIDWSSGGVRLLTEPVPW-QQ--N-GRPR
borre_002_KS_001.seq      AAGVAGVIKMVLAMEKGRLPRT-LHVDEPSGEVDWDSGAVRLLTEARDW-PSE-E-GRLR
borre_003_KS_001.seq      AAGVAGVIKMVKAMQAGTLPRT-LHVDEPSGEVDWDSGAVRLLTEARDW-PSE-E-GRLR
borre_003_KS_002.seq      AAGVASVIKMVQAMQAGVLPRT-LHVDEPSGEVDWDSGAVRLLTEAREW-PSG-E-GRVR
borre_004_KS_001.seq      AAGVASVIKMVQAMQAGVLPRT-LHVDEPSGEVDWDSGAVRLLTEAREW-PSG-E-GRVR
borre_005_KS_001.seq      AAGVAGVIKMVLALRKGVLPRT-LHVDEPTGEVDWDSGAVRLLTEAREW-PSG-E-GRVR
borre_006_KS_001.seq      AAGVAGVIKMVLAMEKGRLPRT-LHVDEPSGEVDWDSGAVRLLTEARDW-PSGE-G-RVR
epoth_002_KS_001.seq      AAGIAGLIKTVLALEHRQLPPS-LNFESPNPSIDFASSPFYVNTSLKDW-NTGS---TPR
epoth_003_KS_001.seq      AAGVAGLIKAALSLTHERIPRN-LNFRTLNPRIRLEGSALALATEPVPW-PRTD---RPR
epoth_003_KS_002.seq      AAGLAGLLKAVLALGQEQIPAQ-PELGELNPLLPWEALPVAVARAAVPW-PRTD---RPR
epoth_003_KS_003.seq      ASGLASLLKAVLALRHEQIPAQ-PELGELNPHLPWNTLPVAVPRKAVPW-GRGA---RPR
epoth_003_KS_004.seq      AAGVAGVIKVALALERGLIPRS-LHFDAPNPHIPWSELAVQVAAKPVEW-TRNG---APR
epoth_004_KS_001.seq      AAGVAGLIKATLSLHHERIPRN-LNFRTLNPRIRIEGTALALATEPVPW-PRTG---RTR
epoth_004_KS_002.seq      AAGLAGVLKVLLALEHEQIPAQ-PELDELNPHIPWAELPVAVVRAAVPW-PRGA---RPR
epoth_005_KS_001.seq      AAGVAGLIKAALALHHESIPRN-LHFHTLNPRIRIEGTALALATEPVPW-PRAG---RPR
lasal_001_KS_002.seq      AAGVAGVIKTVMAMRHGLLPRT-LHVDEPSPHIDWTSGAVSLLTEPVDW-HVG---DRPR
lasal_001_KS_003.seq      AAGVAGVIKMVMAMRHGLLPRT-LHVDEPTPHVDWSTGAVSLLTEDTTW-PET---GRPR
lasal_002_KS_001.seq      AAGMAGIIKMVMAMRHGILPRT-LHVDEPSPHIDWTAGAVSLLTEPVDW-QAG---DRPR
lasal_002_KS_002.seq      AAGMAGIIKTVMAMRHGLLPRT-LHLDEPTPHVDWTSGRVALLTEPVAW-PVS---DRPR
lasal_002_KS_003.seq      AAGMAGIIKMVMAMRHGILPRT-LHVDEPSPHIDWTAGAVSLLTEPVDW-QAG---DRPR
lasal_003_KS_001.seq      AAGVAGVIKMVMAMRHGVLPRS-LHIDEPTPQVDWSSGAVTLLTEPVDW-PDS---DRPR
lasal_004_KS_001.seq      AAGVAGVIKMVMAMRHGVLPRS-LHIDEPTPQVDWSSGAVTLLTEPVDW-PDS---DRPR
lasal_005_KS_001.seq      AAGMAGIIKTVMALRHGVLPRT-LHVDAPTPHVDWASGLVSLLTEQMPW-PET---GRPR
lasal_005_KS_002.seq      AAGVAGVIKMVMAMRHGVLPST-LHVDDPSPHVDWTAGQMELLTDARQW-PEG---RRPR
lasal_006_KS_001.seq      AAGMAGLIKTVLAMRHGVMPRT-LHLEEPTPHVDWTSGRVSLLTEQTAW-PAS---DRPR
lasal_007_KS_001.seq      AAGVAGVIKMVMALRHEVLPQT-LHVDAPSPHIDWASGHVSLLTEELKW-PGT---DRPR
rapam_001_KS_001.seq      AAGVSGVIKMVMALRHGFVPRT-LHVDEPSRHVDWSAGAVELVAENRSW-PAT---GRPR
rapam_001_KS_002.seq      AAGVSGVIKMVMALQRGFVPRT-LHVDEPSRHVDWSAGAVQLVTENQPW-PGT---DRPR
rapam_001_KS_003.seq      ASGVSVVIKMVMALQHNTVPRT-LHVDEPSRHVDWAAGAVELVRENQPW-PGT---DRPR
rapam_001_KS_004.seq      ASGVSGVIKMVMALRHGFVPRT-LHVDEPSRHVDWAAGAVELVRENQPW-PGT---DRPR
rapam_002_KS_001.seq      AAGVSGVIKMVMALQHVLVPRT-LHVDEPSRHVDWTDGAVALVTENQPW-PDM---GRPR
rapam_002_KS_002.seq      ASGVSGVIKMVMALQHNTVPRT-LHVDEPSRHVDWAAGAVELVRENQPW-PGT---DRPR
rapam_002_KS_003.seq      AAGVSGVIKMVMALQHSTVPRT-LHVNEPSRHVDWSAGAVELVTENQSW-PVT---GRPR
rapam_002_KS_004.seq      AAGVSGVIKMVMALQRGFVPRT-LHVDEPSRHVDWSAGAVALVTENQPW-PDM---GRAR
rapam_002_KS_005.seq      AAGVSGVIKMVMALQHGLVPRT-LHVDEPSRHVDWTDGAVELVTENQSW-PEA---GRPR
```

```
rapam_002_KS_006.seq      AAGVSGVIKMVMALQHSMVPRT-LHVDEPSRHVDWSAGAVELVAENQPW-PET---GRPR
rapam_003_KS_001.seq      AAGVSGVIKMVMALRHALVPRT-LHVDEPSRHVDWTAGAVELVTENQPW-PEI---GRPR
rapam_003_KS_002.seq      AAGVSGVIKMVMALQHGLVPRT-LHVDEPSRHVDWSAGAVQLVTENQPW-PDM---GRAR
rapam_003_KS_003.seq      ASGVSGVIKMVMALQHNTVPRT-LHVDEPSRHVDWAAGAVELVRENQPW-PGT---DRPR
rapam_003_KS_004.seq      ASGVLSVIKMVEAMAHGSLPRT-LHVDAPSPHVDWTSGSVALLTEHQPW-PDD---TKLR
nysta_002_KS_001.seq      AAGVAGVIKMVMALRHGHLPPT-LHADAPSSHVDWSAGSVRLLTEGQQW-PETG---RPR
nysta_002_KS_002.seq      AAGVAGVIKMVLAMDAGELPGT-LHLDAPSSHVDWTAGAVELLRGRTPW-PES---GRPR
nysta_003_KS_001.seq      ASGAAGVIKMIMALQRGVLPRS-LHATEPTTDVDWTAGSVDLLDETVAW-PET---GRAR
nysta_003_KS_002.seq      AAGVAGVIKMVMAMRHGVLPQT-LHVDAPSSHVDWSVGAVELLTEQTAW-PET---GRAR
nysta_003_KS_003.seq      AAGVAGVIKMVMAMRHGVLPQT-LHVDAPSSHVDWSVGAVELLTEQTAW-PET---GRVR
nysta_003_KS_004.seq      AAGVAGVIKMVMAMRHGVLPRS-LNITEPSSHVDWSAGAVELLTEQTAW-PET---GRAR
nysta_003_KS_005.seq      AAGVAGVIKMVLAMRHGVVPRT-LHVDAPSSHVDWSEGAVELLSEQAAW-PET---GRVR
nysta_003_KS_006.seq      AAGVAGVIKMVMAMRHGVLPQT-LHVDAPSSHVDWSVGAVELLTEQTVW-PET---GRVR
nysta_004_KS_001.seq      AAGVAGVIKMVMAMRHGTLPRT-LHAEEPTHHVDWSQGAVRLLTDTTDW-PATG---APR
nysta_004_KS_002.seq      AAGVAGVIKMVLAMRHGVVPRT-LHVDAPSSHVDWSEGAVELLSEQAAW-PET---GRVR
nysta_004_KS_003.seq      AAGIASVIKMVMALRHSELPPT-LHADAPSSHVDWSAGTVRLLTQARAW-PETG---RPR
nysta_004_KS_004.seq      ASGVASVIKLVRALQEGVVPKS-LHIDRPSTHVDWSSGAIGLLTERTPW-PETG---RPR
nysta_004_KS_005.seq      AAGAAGVIKMLMAMRHGTLPRT-LHVGTPSSHVDWSGGTVALLDDARPW-PRTG---QPR
nysta_004_KS_006.seq      ASGVGGVIKMVLAMQHGELPRS-LYAENPSSHVDWTAGRAHLLTARTPW-PDSG---RPR
nysta_005_KS_001.seq      ASGVAGVIKMIMAMRHGVLPRT-LHVDRPSTHVDWTTGSVELLTDAHPW-PET---GRPR
nysta_005_KS_002.seq      AAGAAGLVKMVMAMRHGILPQT-LHLTEPSSHVDWSAGTVRLLTERTAW-PRTD---RPR
nysta_005_KS_003.seq      AAGAAGLMKMVLALQHGTLPRT-LHVTEPSTRVDWSAGAVRLLTERTVW-PRTD---RPR
nysta_006_KS_001.seq      AAGVAGVIKMVLALRHGVLPQT-LHVDAPSSHVDWESGAVRLLTAPVAW-SEGD--DRVR
spino_002_KS_001.seq      AAGVAGVIKMVMAMRHGQLPRT-LHVESPSPEVDWSAGTVQLLTENTPW-PRS---GRVR
spino_003_KS_001.seq      AAGVAGVIKMVMAMRHGELPRT-LHVDEPNSHVDWSAGAVRLLTENIRW-PGT---GT-R
spino_003_KS_002.seq      AAGVAGVIKMVMAMRHGQLPAS-LHADEPTSEVDWSSGAVRLLAEQVPW-PES---DRVR
spino_004_KS_001.seq      AAGVAGVIKMVLAMRHGQLPAT-LHVDEPTSAVDWSAGSVRLLTENTPW-PDS---GRPC
spino_004_KS_002.seq      AAGVAGVIKMVMAMRQGELPRT-LHVDEPSAQVDWSAGTVQLLTENTPW-PDS---GRLR
spino_004_KS_003.seq      AAGVAGVIKMVQAMRHGELPAT-LHVDKPTPQVDWSAGAVRLLTGNTPW-PES---GRPR
spino_005_KS_001.seq      AAGVAGVIKSVLAIRHREMPRS-LHIDQPSQHVDWSAGAVRLLTDSVDW-PDL---GRPR
spino_005_KS_002.seq      AAGVAGVIKMVMAMRHGQLPAT-LHVDEPTSEVDWSAGDVQLLTENTPW-PGN---SHPR
spino_005_KS_003.seq      AAGVAGVIKMVMAMRHGQLPRT-LHVDAPSSQVDWSAGRVQLLTENTPW-PDS---GRPC
tylac_001_KS_002.seq      AAGVAGVIKMVLAMRYGWLPRT-LHVDEPSRHVDWSAGGVWLLTEAREW-PGV---DRPR
tylac_001_KS_003.seq      AAGVAGVIKMVLAMRYGWLPRT-LHVDEPSRHVDWSAGGVRLLTEAREW-PGV---DRPR
tylac_002_KS_001.seq      AAGVAGVIKMVLAMRHGRLPRT-LHVDRPTTRVDWEKGGVRLLTEPVPW-PGEA-G-EPR
tylac_003_KS_001.seq      AAGVAGVIKMVLAMRYGWLPRT-LHVDEPSRHVDWSAGGVWLLTEAREW-PGV---DRPR
tylac_003_KS_002.seq      AAGVAGVIKMVLALGNETLPRT-LHVDEPTPRVDWSSGAVSLLTEPVDW-PAGP-S-APR
tylac_004_KS_001.seq      AAGVAGVIKMVLAMRYGWLPRT-LHVDEPSRHVDWSAGGVRLLTEAREW-PGV---DRPR
tylac_005_KS_001.seq      AAGVGGVIKMLQALRHGLLPRT-LHADEPTPHADWSSGRVRLLTSEVPW-QR-T-G-RPR
nidda_001_KS_002.seq      AAGIAGIIKMVQALRHDTLPRT-LHADHPSSKVDWDAGPLQLLTDARPW-PADP-D-RPR
nidda_001_KS_003.seq      AAGIAGIIKMIQAMHHGTLPRT-LHVDAPTSQVDWEAGSLQLLAEARPW-PADP-D-RPR
nidda_002_KS_001.seq      AAGIAGIIKMIQAMHHGTLPRT-LHVDHPSSKVDWKAGSLQLLAQARPW-PADP-D-RPR
nidda_003_KS_001.seq      AAGIAGIIKMVQALRHDTLPRT-LHADHPSSKVDWEAGPLQLLAQARPW-PADP-D-RPR
nidda_003_KS_002.seq      AAGVAGVIKMVQALRHDTLPRT-LHVDRPTSKVDWETGRVRLLTDARPW-PAGP-D-RPR
nidda_004_KS_001.seq      AAGVAGIIKMVQALRHDTLPRT-LHADHPSSKVDWDAGPLRLLTDARPW-PADP-D-RPR
nidda_005_KS_001.seq      AAGIAGIIKMIQAMHHGTLPRT-LHVDHPSSKVDWKAGSLQLLAQARPW-PADP-D-RPR
monen_001_KS_002.seq      AAGVAGVIKMVMAMRHDLLPAT-LHVDEPSGHVDWSTGAVRLLTEPVVW-PRG---ERPR
monen_002_KS_001.seq      TAGVAGVIKTVMAIRNGLLPAT-LHVEELSPHVDWDAGAVEVVTEPTPW-PET---GHPR
monen_003_KS_001.seq      AAGAAGLIKMVQALRHETLPAT-LYADEPTPHADWESGAVRLLSAPVAW-PRGEHGEHTR
monen_003_KS_002.seq      AAGVAGVIKMVMALRHEQLPTT-LHADEPTPHVQWDGGGVRLLTEPVPW-SRG---ERTR
monen_004_KS_001.seq      AAGVAGVIKMVMALQRELLPAT-LYVDEPTPHVDWSSGSVRLLTEPVPW-TRGE---RPR
monen_004_KS_002.seq      AAGVAGVIKMVMAMRHGVVPAS-LHVDVPSPHVEWDSGAVRLAVESVPW-PQVE--GRPR
monen_005_KS_001.seq      AAGVAGVIKMVMALQRELLPAT-LNVDEPTPHVQWEGGGVRLLTEPVPW-SRGE---RPR
monen_005_KS_002.seq      AAGVAGVIKMVMAMRHGVVPAS-LHVDVPSPHVEWDSGAVRLAVESVPW-PEV-E-GRPR
monen_006_KS_001.seq      AAGVAGVIKMVMAMRHGELPAS-LHIDRPTPHVDWEGGGVRLLTDPVPW-PRA---DRPR
monen_007_KS_001.seq      AAGVAGVIKMVLALRHDLLPAI-LHVDAPSPHVEWDGSGLRLLTDPVKW-PRG---ERPR
monen_008_KS_001.seq      AAGVAGVIKMVMAMRNGLLPTS-LHIDAPSPHVQWEQGSVRLLSEPVDW-PA----ERTR
monen_008_KS_002.seq      SAGVAGVIKMVMALRNEQLPAS-LHIDAPTPHVDWDGSGVRLLSEPVSW-PRG---ERPR
oligo_001_KS_002.seq      AAGVAGVIKSVLALRNGLLPKT-LHVDEPTPEVDWSAGAVELLTEGREW-PET---DGPR
oligo_001_KS_003.seq      AAGMAGVLKTLSAMRHGQLPRT-LHAARPTPKVDWTEGAVELLTETRPW-PAT-G-ARPR
oligo_001_KS_004.seq      AAGVAGVIKSVLALRNGLLPAT-LHVDEPSREVDWSAGAVELLTEGREW-PETD---GPR
oligo_002_KS_001.seq      AAGMAGVIKMVLAMRHGLLPRT-LHVDEPTPQVDWSAGAVRLLTEEQTW-PSTG---RPR
oligo_002_KS_002.seq      AAGVGGVIKMVQALRHGVLPRT-LHVDEPTPEVDWSAGTVGLLTENREW-PHTG---KPR
oligo_003_KS_001.seq      AAGVGGVIKMVQALRHGLLPRT-LHAEEPTPNVDWSSGAVELLTEARPW-PA-S-D-TPR
oligo_003_KS_002.seq      AAGVGGVIKMVQALHHGVLPRT-LHIDEPTPHVDWTAGAVDLLTETRPW-PETG---RPR
oligo_004_KS_001.seq      AAGAAGIIKMVQAMRHEQLPRT-LHVDRPTPEVDWSEGAVELLTENRPW-PRG---DRPR
oligo_004_KS_002.seq      AAGAAGIIKMIMAMRYGVLPRT-LHVDRPSPEVDWSPGTVELLTEEREW-PDA---GRPR
oligo_005_KS_001.seq      AAGVSGVIKTVMALRHGVMPKT-LHVDEPTPHVDWSSGAVRLLTENRPW-PGHG---RPR
oligo_006_KS_001.seq      AAGVSGVIKAVLSLQHGLLPKT-LHVDEPTPQVDWSAGAVELLTQAREW-PETD---GPR
oligo_006_KS_002.seq      AAGVAGVIKAVLALRNGLLPKT-LHVDEPTPQVDWSAGAVELLTQAREW-PDAE---RPR
oligo_006_KS_003.seq      AAGVAGVIKSVLALRNGLLPKT-LNVDEPTPKVDWSAGAVELLTQAREW-PQGE---RPR
oligo_007_KS_001.seq      VSGVAGVIKTVLALRNGVLPKT-LHVGKPSTQVDWSAGAVELLTEGREW-PETG---GPR
oligo_007_KS_002.seq      AAGVAGVIKSVLALRNGLLPKT-LHVDEPTPEVDWSAGAVELLTEGREW-PETD---GPR
rifam_001_KS_001.seq      AAGVAGVIKMVQALRHETLPPT-LHVDKPTLEVDWSAGAIELLTEARAW-PRN---GRPR
rifam_001_KS_002.seq      AAGVAGVIKMVMAMRHGVLPRT-LHVDRPSSYVDWSAGAVELLTEARDW-VSN---GHPR
rifam_001_KS_003.seq      AAGVVGVIKMIMAMRHGVMPAT-LHVDERTSQVDWSAGAIEVLTEAREW-PRT---GRPR
rifam_002_KS_001.seq      ASGVAGVIKMVQSLRHGQLPAT-QHVDAPTPQVDWSAGAIELLAEGREW-PR-N-G-HPR
```

```
rifam_002_KS_002.seq      AAGVAAVIKMVQALRHDTLPPT-LHVQEPTKQVDWSAGAVELLTEGREW-AR-N-G-HPR
rifam_002_KS_003.seq      AAGVAGVIKMVQALRHEVLPPT-LHVDRPTPEVDWSAGAVELLTEAREW-PRN---GRPR
rifam_003_KS_001.seq      AAGVASVIKVVQALRHGVMPPT-LHVDEPSSQVDWSEGAVELLTGSRDW-PR-G-D-RPR
rifam_004_KS_001.seq      AAGVASVIKMVQALRHGVLPPT-LHVDRPSTEVDWSAGAVSLLTEAREW-PRE---GRPR
rifam_005_KS_001.seq      AAGVAGVIKMVQALRHGAMPPT-LHVAEPTPEVDWSAGAVELLTEPREW-PAG---DRPR
rifam_005_KS_002.seq      AAGVAGVIKMVQALRHGVMPPT-LHVDRPTSQVDWSAGAVEVLTEAREW-PR-N-G-RPR
pikro_001_KS_002.seq      AAGVSGVIKMVQAMRHGLLPKT-LHVDEPSDQIDWSAGAVELLTEAVDW-PEKQD-GGLR
pikro_001_KS_003.seq      AAGAAGIIKMVLAMRHGTLPKT-LHADEPSPHVDWANSGLALVTEPIDW-PAGT---GPR
pikro_002_KS_001.seq      AAGVSGVIKMVQAMRHGVLPKT-LHVEKPTDQVDWSAGAVELLTEAMDW-PDKGD-GGLR
pikro_002_KS_002.seq      AAGVSGIIKMVQAMRHGVLPKT-LHVDRPSDQIDWSAGTVELLTEAMDW-PRKQE-GGLR
pikro_003_KS_001.seq      AAGVAGVIKMVQAMRHGLLPKT-LHVDEPSDQIDWSAGTVELLTEAVDW-PEKQ-DGGLR
pikro_004_KS_001.seq      ASGVAGVIKMVQALRHGVLPKT-LHVDEPTDQVDWSAGSVELLTEAVDW-PERP-G-RLR
Stamb_001_KS_002.seq      AAGVAGVIKTVLALRHGVLPRT-LHADEPTPHVDWSAGAVRLLTENVTL-PPAD---RPY
Stamb_001_KS_003.seq      AAGVGGIIKTVLALRHGVLPRT-LHADEPSSHVDWSAGAVELLTDEREW-PSV---GRPR
Stamb_001_KS_004.seq      AAGIAGVIKMLLAMDHGELPRT-LHAAAPSPSVDWDSGSLSLLTETTPWTPRVD---RPR
Stamb_002_KS_001.seq      AAGVAGVIKTVMALRHGVLPRT-LHADEPTRQVDWSTGAVRLLTENRPW-PGP---GRTR
Stamb_002_KS_002.seq      AAGVAGVIKMVQALRSGVVPAT-LHVDEPTSHVDWSAGGIRLAVESRAW-PETG---RAR
Stamb_003_KS_001.seq      ASGVVGVIKSVLALRNGLLPRT-LHVAVPTPHVVWEGRGVRLLADARPW-PRTD---VPR
Stamb_003_KS_002.seq      AAGVAGVIKMVQALRSGTVPAT-LHVDEPTPHVDWTTGAVRLAVASRAW-PETG---RAR
Stamb_003_KS_003.seq      AAGVAGVIKMVLAMRNGVLPPTLLHADEPTPHVDWSAGAVELLTEARPW-PLSD---RPH
Stamb_003_KS_004.seq      ASGVAGVIKMVMAMRHGVLPAT-LHVDEPTPHVDWSTGAVKLLTDARPW-TVTD---RPR
Stamb_003_KS_005.seq      AAGVAGVIKMVLALRHGTLPPT-LHADEPTEQVDWSAGAVELLTRARPW-PEAD---RPR
Stamb_004_KS_001.seq      AAGVAGVIKMVQALRHERLPRT-LHVDEPSTQVDWSAGAVELLTEGRSW-PRG---ERPR
Stamb_004_KS_002.seq      AAGVAGIIKMVMALRHGVLPRS-LHITEPTRHVDWDGSGVRLLTRQVEW-ADTE---RPR
Stamb_004_KS_003.seq      AAGVGGVIKTVMALREGVLPGT-LHIDAPTPHVDWSTGAVELLTETIPW-PDTD--GRPR
Stamb_005_KS_001.seq      AAGMAGVIKMVEGLRRGVLPKT-LHVDAPTPQVDWSAGAVELLTETRPW-PETG---RPR
Stamb_006_KS_001.seq      ASGVAGVIKMVQAMRHGVLPKT-LHVDAPSPHVDWSSGSVELLTENRSW-PDS---GGPR
Stamb_006_KS_002.seq      AAGVAGVIKMVEALRRGVLPGT-LHVDEPTPQVDWSAGAVELLTEARPW-PET---DRPR
Stamb_007_KS_001.seq      ASGVAGLIKMVMALRAGELPPT-LHVTEPTPHVDWSQGHVRLLTDSASW-PQDP--ERPR
Stamb_007_KS_002.seq      AAGVGGVIKTVMALRHGVLPRT-LHVTEPTPEVDWTTGAVELLTEARPW-PEPD---RPR
Stamb_008_KS_001.seq      AAGAAGVIKMVLAMSHGVLPRT-LHADVPTPHVDWSSGRIELLTAEQPW-EPGE---QPR
Stamb_008_KS_002.seq      AAGAGGVIKMVQALRHGVLPRT-LHADEPTPHVDWSSGAVELLTQALEW-PSG---EQPR
Stamb_009_KS_001.seq      AAGAGGVIKMVQALRHGVLPRT-LHADEPTPHVDWSSGAVELLTQALEW-PSGE---QPR
Stamb_009_KS_002.seq      AAGVVGVIKAVMALRHGVLPAT-LHVDEPTPEVDWSSGAVELLTENRTW-PEA---GRPR
Stamb_009_KS_003.seq      AAGVAGVIKMVEALRHGVLPAT-LHAREPSAQVDWSAGAVELLTDARRW-PET---DRPR
Stamb_009_KS_004.seq      ASGIAGVIKTVEAIRRGVLPRT-LHVDEPTPEVDWSAGAVELLTDDRPW-PDTG---RAR
                           :*   ::*   .:     *         .           :

mycol_001_KS_001.seq      TAAVSSFGISGTNAHLIL-------
mycol_001_KS_002.seq      TAAVSSFGISGTNAHLIL-------
mycol_001_KS_003.seq      TAAVSSFGISGTNAHLIL-------
mycol_001_KS_004.seq      TAAVSSFGISGTNAHLIL-------
mycol_001_KS_005.seq      TAAVSSFGISGTNAHLIL-------
mycol_001_KS_006.seq      TAAVSSFGISGTNAHLIL-------
mycol_001_KS_007.seq      TAAVSSFGISGTNAHLIL-------
mycol_001_KS_008.seq      TAAVSSFGISGTNAHLIL-------
mycol_002_KS_001.seq      TAAVSSFGISGTNAHLIL-------
mycol_003_KS_001.seq      TAAVSSFGISGTNAHLIL-------
mycol_003_KS_002.seq      TAAVSSFGISGTNAHLIL-------
mycol_003_KS_003.seq      TAAVSSFGISGTNAHLIL-------
mycol_003_KS_004.seq      TAAVSSFGISGTNAHLIL-------
mycol_003_KS_005.seq      TAAVSSFGISGTNAHLIL-------
mycol_003_KS_006.seq      TAAVSSFGISGTNAHLIL-------
mycol_003_KS_007.seq      TAAVSSFGISGTNAHLIL-------
eryth_001_KS_001.seq      RAGVSSFGISGTNAHAIIEEAP---
eryth_001_KS_002.seq      RAGVSAFGVSGTNAHVIIAEPP---
eryth_002_KS_001.seq      RARVSAFGISGTNAHVIVEEAP---
eryth_002_KS_002.seq      RAGVSSFGVSGTNAHVIVEEAP---
eryth_003_KS_001.seq      RAGVSSFGISGTNAHVIVEEAP---
eryth_003_KS_002.seq      RAGVSAFGVSGTNAHVIIAEPP---
ampho_002_KS_001.seq      RAAVSSFGISGTNAHAVLEQAP---
ampho_002_KS_002.seq      RAGISSFGISGTNAHLLLEQAP---
ampho_003_KS_001.seq      RAGVSSFGISGTNAHLILEQAP---
ampho_003_KS_002.seq      RAAVSSFGISGTNAHVIIEQP----
ampho_003_KS_003.seq      RAAVSSFGISGTNAHVILEQP----
ampho_003_KS_004.seq      RAGVSSFGISGTNAHVILEQP----
ampho_003_KS_005.seq      RAGVSSFGISGTNAHVILEQP----
ampho_003_KS_006.seq      RAGVSSFGISGTNAHVILEQP----
ampho_004_KS_001.seq      RAAVSSFGISGTNAHTIVEQAP---
ampho_004_KS_002.seq      RAAVSSFGISGTNVHTVLEQAP---
ampho_004_KS_003.seq      RAAVSSFGISGTNAHVILEQ-----
ampho_004_KS_004.seq      RAAVSSFGLSGTNVHTVLEQAP---
ampho_004_KS_005.seq      RAGVSAFGVSGTNAHVIVEQAP---
ampho_004_KS_006.seq      RAAVSSFGASGTNAHLILEQP----
ampho_005_KS_001.seq      RSGVSSFGVSGTNAHVIVEQAP---
ampho_005_KS_002.seq      RAGVSSFGISGTNAHVILEGAP---
ampho_005_KS_003.seq      RAGVSSFGISGTNAHVILEGAP---
```

```
ampho_006_KS_001.seq      RAGVSSFGISGTNAHVILEEAP---
averm_001_KS_001.seq      RAGVSSFGIGGTNAHVILEEAP---
averm_001_KS_002.seq      RAGVSSFGVSGTNAHVILEEAP---
averm_002_KS_001.seq      RAGVSSFGVSGTNAHVILEEAP---
averm_002_KS_002.seq      RAGVSSFGVSGTNAHVILEEAP---
averm_002_KS_003.seq      RAGVSSFGVSGTNAHVILEEAP---
averm_002_KS_004.seq      HAGVSSFGVSGTNAHIILEEAP---
averm_003_KS_001.seq      RAGVSSFGVSGTNAHVILEEAP---
averm_003_KS_002.seq      RAGVSSFGVSGTNAHVILEEAP---
averm_003_KS_003.seq      HAGVSSFGVSGTNAHVILEEAP---
averm_004_KS_001.seq      RAGVSAFGVSGTNAHVILEE-----
averm_004_KS_002.seq      RAGVSAFGVSGTNAHVILEE-----
averm_004_KS_003.seq      RAGVSAFGVSGTNAHVIIEQAP---
borre_002_KS_001.seq      RAGVSSFGISGTNAHVIIEEAP---
borre_003_KS_001.seq      RAGVSSFGISGTNAHVILEEPP---
borre_003_KS_002.seq      RAGVSSFGISGTNAHVILEEPP---
borre_004_KS_001.seq      RAGVSSFGISGTNAHVILEEPP---
borre_005_KS_001.seq      RAGVSSFGISGTNAHVIVEEAP---
borre_006_KS_001.seq      RAGVSSFGISGTNAHVIIEEP----
epoth_002_KS_001.seq      RAGVSSFGIGGTNAHVVLEEAP---
epoth_003_KS_001.seq      FAGVSSFGMSGTNAHVVLEEAP---
epoth_003_KS_002.seq      FAGVSSFGMSGTNAHVVLEEAP---
epoth_003_KS_003.seq      RAGVSAFGLSGTNVHVVLEEAP---
epoth_003_KS_004.seq      RAGVSSFGVSGTNAHVVLEEAP---
epoth_004_KS_001.seq      FAGVSSFGMSGTNAHVVLEEAP---
epoth_004_KS_002.seq      RAGVSAFGLSGTNAHVVLEEAP---
epoth_005_KS_001.seq      FAGVSAFGLSGTNVHVVLEEAP---
lasal_001_KS_002.seq      RAGVSSFGISGTNAHVILED-----
lasal_001_KS_003.seq      RAAVSSFGISGTNAHVVLEAP----
lasal_002_KS_001.seq      RAGVSSFGISGTNAHVILEQAP---
lasal_002_KS_002.seq      RAGVSSFGASGTNAHVILEQ-----
lasal_002_KS_003.seq      RAGVSSFGISGTNAHVILEHVAEVP
lasal_003_KS_001.seq      RAGVSAFGISGTNAHVILEQAP---
lasal_004_KS_001.seq      RAGVSAFGISGTNAHVILEQAP---
lasal_005_KS_001.seq      RAGVSSFGVSGTNAHVILEQ-----
lasal_005_KS_002.seq      RAAVSSFGISGTNAHVILEAAP---
lasal_006_KS_001.seq      RAGVSSFGMSGTNAHVILEEAP---
lasal_007_KS_001.seq      RAGVSSFGISGTNAHVILEAAPP--
rapam_001_KS_001.seq      RAGVSAFGVSGTNAHVILEGAP---
rapam_001_KS_002.seq      RAGVSSFGISGTNAHVIL-------
rapam_001_KS_003.seq      RAGVSSFGVSGTNAHVILESAP---
rapam_001_KS_004.seq      RAGVSSFGVSGTNAHVVLESAP---
rapam_002_KS_001.seq      RAGVSSFGISGTNAHVILESAP---
rapam_002_KS_002.seq      RAGVSSFGVSGTNAHVILESAP---
rapam_002_KS_003.seq      RAGVSAFGVSGTNAHVVLESAP---
rapam_002_KS_004.seq      RAGVSSFGISGTNAHVILESAP---
rapam_002_KS_005.seq      RAGVSSFGVSGTNAHVILESAP---
rapam_002_KS_006.seq      RAGVSSFGISGTNAHVILESAP---
rapam_003_KS_001.seq      RAGVSSFGVSGTNAHVILESAP---
rapam_003_KS_002.seq      RAGVSSFGISGTNAHVILESAP---
rapam_003_KS_003.seq      RAGVSSFGVSGTNAHVILESAP---
rapam_003_KS_004.seq      RAGVSSFGLSGTNAHVVLEQ-----
nysta_002_KS_001.seq      RAAVSSFGISGTNAHALLEQAP---
nysta_002_KS_002.seq      RAGVSSFGISGTNAHLILEQAP---
nysta_003_KS_001.seq      RAGVSSFGISGTNAHVILEQAP---
nysta_003_KS_002.seq      RAGVSSFGISGTNAHVVIEQSP---
nysta_003_KS_003.seq      RAGVSSFGISGTNAHVIVEQP----
nysta_003_KS_004.seq      RAGISSFGISGTNAHVILEQP----
nysta_003_KS_005.seq      RAGVSSFGISGTNAHVILERP----
nysta_003_KS_006.seq      RAGVSSFGISGTNAHVILEQP----
nysta_004_KS_001.seq      RAAVSSFGISGTNAHTIIEQAP---
nysta_004_KS_002.seq      RAGVSSFGISGTNVHVIVEQAP---
nysta_004_KS_003.seq      RAAVSSFGISGTNAHVLLEQAP---
nysta_004_KS_004.seq      RAAVSSFGISGTNVHTILEQAP---
nysta_004_KS_005.seq      RAGVSAFGVSGTNAHVVVEQAP---
nysta_004_KS_006.seq      RAAVSSFGASGTNAHAILEQPP---
nysta_005_KS_001.seq      RTGISSFGVSGTNAHVIVEQAP---
nysta_005_KS_002.seq      RAGVSSFGISGTNAHVILEQPP---
nysta_005_KS_003.seq      RAGVSSFGISGTNAHVILEQPP---
nysta_006_KS_001.seq      RAGVSSFGISGTNAHVILEQAP---
spino_002_KS_001.seq      RVGVSSFGISGTNAHVILE------
spino_003_KS_001.seq      RAGVSSFGVSGTNAHVILEHDP---
spino_003_KS_002.seq      RVGVSSFGISGTNAHVILE------
spino_004_KS_001.seq      RVGVSSFGISGTNAHVILEQSP---
spino_004_KS_002.seq      RAGVSSFGISGTNAHLILEQPP---
spino_004_KS_003.seq      RAGVSSFGISGTNAHLILE------
spino_005_KS_001.seq      RAGVSSFGMSGTNAHLIVEE-----
```

```
spino_005_KS_002.seq      RVGVSSFGISGTNAHVILEQA----
spino_005_KS_003.seq      RVGVSSFGISGTNAHVILE------
tylac_001_KS_002.seq      RAAVSAFGVSGTNAHLILE------
tylac_001_KS_003.seq      RAAVSAFGVSGTNAHLILE------
tylac_002_KS_001.seq      RAGVSSFGASGTNAHVVLE------
tylac_003_KS_001.seq      RAAVSAFGVSGTNAHLILE------
tylac_003_KS_002.seq      RAAVSSFGISGTNAHTILE------
tylac_004_KS_001.seq      RAAVSAFGVSGTNAHLILE------
tylac_005_KS_001.seq      RTGVSAFGVGGTNAHVVLE------
nidda_001_KS_002.seq      RAGISAFGVSGTNAHVVLEEPP---
nidda_001_KS_003.seq      RAGISAFGVSGTNAHVVLEEPP---
nidda_002_KS_001.seq      RAAVSAFGVSGTNAHAIIEEP----
nidda_003_KS_001.seq      RAGISAFGVSGTNAHAIIEEPP---
nidda_003_KS_002.seq      RAGISAFGISGTNAHVVIEEPP---
nidda_004_KS_001.seq      RAGISSFGVSGTNAHVVLEEPP---
nidda_005_KS_001.seq      RAAVSAFGVSGTNAHTIIEEPP---
monen_001_KS_002.seq      RAAVSSFGISGTNAHLVLEE-----
monen_002_KS_001.seq      RAGVSAFGISGTNAHLILEEAP---
monen_003_KS_001.seq      RAGISSFGISGTNAHLILEEAP---
monen_003_KS_002.seq      RAGVSSFGISGTNAHLILEEPP---
monen_004_KS_001.seq      RAGVSAFGMSGTNAHVILEEAP---
monen_004_KS_002.seq      RAGVSSFGASGTNAHVIVESVP---
monen_005_KS_001.seq      RAGISSFGISGTNAHVVLEEAP---
monen_005_KS_002.seq      RAGVSSFGASGTNAHVIVESVP---
monen_006_KS_001.seq      RAGVSSFGISGTNAHLIVEQAP---
monen_007_KS_001.seq      RAGVSSFGFSGTNAHLILEEAP---
monen_008_KS_001.seq      RAGISAFGISGTNAHLILEEAP---
monen_008_KS_002.seq      RAGVSAFGISGTNAHLILEQAP---
oligo_001_KS_002.seq      RAGVSAFGISGTNAHVILE------
oligo_001_KS_003.seq      RAGVSAFGVSGTNAHVILE------
oligo_001_KS_004.seq      RAGVSAFGISGTNAHVILE------
oligo_002_KS_001.seq      RSAVSSFGISGTNAHVVLE------
oligo_002_KS_002.seq      RVGVSAFGVSGTNAHVVLE------
oligo_003_KS_001.seq      LAAVSSFGISGTNAHVILE------
oligo_003_KS_002.seq      RAGVSSFGVSGTNAHVILE------
oligo_004_KS_001.seq      RAGVSSFGISGTNVHVILE------
oligo_004_KS_002.seq      RAAVSSFGISGTNAHVILE------
oligo_005_KS_001.seq      RAAVSSFGVSGTNAHLILE------
oligo_006_KS_001.seq      RAGVSSFGISGTNAHVILE------
oligo_006_KS_002.seq      RAGVSSFGISGTNAHVILE------
oligo_006_KS_003.seq      RAGVSSFGISGTNAHVIIE------
oligo_007_KS_001.seq      RAGVSSFGISGTNAHVILE------
oligo_007_KS_002.seq      RAGVSSFGISGTNAHVILE------
rifam_001_KS_001.seq      RAGVSSFGVSGTNAHLILEEAP---
rifam_001_KS_002.seq      RAGVSSFGIGGTNAHVVLEE-----
rifam_001_KS_003.seq      RAGVSSFGASGTNAHLIIEEGP---
rifam_002_KS_001.seq      RGGISSFGASGTNAHMILEEAP---
rifam_002_KS_002.seq      RAGVSSFGISGTNAHLILEEAP---
rifam_002_KS_003.seq      RAGVSAFGVSGTNAHLILEEAP---
rifam_003_KS_001.seq      RAGVSSFGVSGTNVHLIIEEAP---
rifam_004_KS_001.seq      RAGVSSFGISGTNAHLILEEAP---
rifam_005_KS_001.seq      RAGVSAFGISGTNAHLILEEAP---
rifam_005_KS_002.seq      RAGVSSFGISGTNAHLIIEEAP---
pikro_001_KS_002.seq      RAAVSSFGISGTNAHVVLEEAP---
pikro_001_KS_003.seq      RAAVSSFGISGTNAHVVLEQAP---
pikro_002_KS_001.seq      RAAVSSFGVSGTNAHVVLEEAP---
pikro_002_KS_002.seq      RAAVSSFGISGTNAHIVLEEAP---
pikro_003_KS_001.seq      RAAVSSFGISGTNAHVVLEEAP---
pikro_004_KS_001.seq      RAGVSAFGVGGTNAHVVLEEAP---
Stamb_001_KS_002.seq      RAGVSSFGISGTNAHAVIEAAP---
Stamb_001_KS_003.seq      RAAVSSFGISGTNAHTVIEQAP---
Stamb_001_KS_004.seq      RAAVSSFGISGTNAHVVLEQ-----
Stamb_002_KS_001.seq      RAAVSSFGVSGTNAHTVLEQAP---
Stamb_002_KS_002.seq      RAGVSSFGMSGTNAHVIIEQAEP--
Stamb_003_KS_001.seq      RIGVSAFGVSGTNAHLILEEAP---
Stamb_003_KS_002.seq      RAGVSSFGMSGTNAHVIIEQAEP--
Stamb_003_KS_003.seq      RAGVSSFGISGTNAHVIL-------
Stamb_003_KS_004.seq      RAGVSSFGISGTNAHILLESAP---
Stamb_003_KS_005.seq      RAGVSSFGISGTNAHVILEAEP---
Stamb_004_KS_001.seq      RAGVSSFGVSGTNGHVILEEAP---
Stamb_004_KS_002.seq      RAAVSSFGMSGTNAHTILEQ-----
Stamb_004_KS_003.seq      RAGVSSFGGSGTNAHLVLEQ-----
Stamb_005_KS_001.seq      RAGVSSFGVSGTNAHVIL-------
Stamb_006_KS_001.seq      RAGVSSFGVSGTNAHVIL-------
Stamb_006_KS_002.seq      RAGVSSFGVSGTNAHVVLEAAP---
Stamb_007_KS_001.seq      RAGVSSFSISGTNAHLVL-------
Stamb_007_KS_002.seq      RAGVSAFGISGTNAHVVLEHAP---
```

```
Stamb_008_KS_001.seq      RAAVSSFGISGTNAHVILEEAP---
Stamb_008_KS_002.seq      RAAVSSFGISGTNAHVVLEE-----
Stamb_009_KS_001.seq      RAAVSSFGVSGTNAHVVLEE-----
Stamb_009_KS_002.seq      RAGVSSFGISGTNAHTILEQ-----
Stamb_009_KS_003.seq      RAGVSSFGMSGTNAHVILESAP---
Stamb_009_KS_004.seq      RAGVSSFGISGTNAHVLIEEAP---
                             :*:*. .*** * ::
```

References

[1] D. G. Gibson, *Methods Enzymol,.* 2011, **498**, 349.

[2] L. A. Kelley and M. J. E. Sternberg, *Nat. Protoc.,* 2009, **4**, 363.

[3] A. Roy, A. Kucukural and Y. Zhang, *Nat. Protoc.,* 2010, **5**, 725.

[4] Y. Li, N. M. Llewellyn, R. Giri, F. Huang and J. B. Spencer, *Chem. Biol.,* 2005, **12**, 665.

[5] H. B. Dixon and K. F. Geoghegan, *Anal. Biochem*., 1999, **267**, 169.

[6] M. M. Bradford, *Anal. Biochem.,* 1976, **72**, 248.

[7] I. H. Gilbert, M. Ginty, J. A. O'Neill, T. J. Simpson, J. Staunton and C. L. Willis, *Bioorg. Med. Chem. Lett.,* 1995, **5**, 1587.

[8] S. K. Piasecki, C. A. Taylor, J. F. Detelich, J. Liu, J. Zheng, A. Komsoukaniants, D. R. Siegel and A. T. Keatinge-Clay, *Chem. Biol*., 2011, **18**, 1331.

[9] S. Mo *et al., J. Am. Chem. Soc.,* 2011, **133**, 976.

[10] O. Vergnolle, F. Hahn, A. Baerga-Ortiz, P. F. Leadlay and J. N. Andexer, *ChemBioChem,* 2011, **12**, 1011.

[11] G. Prasad, L. S. Borketey, T.-Y. Lin and N. A. Schnarr, *Org. Biomol. Chem*., 2012, **10**, 6717.

[12] B. D. Ames, *et al. Proc. Natl. Acad. Sci.,* 2012, **109**, 11144.

[13] R. C. Harris, A. L. Cutter, K. J. Weissman, U. Hanefeld, M. C. Timoney and J. Staunton, *J. Chem. Res. Synop.* 1998, **6**, 283.

[14] M. Jenner, S. Frank, A. Kampa, C. Kohlhaas, P. Pöplau, G. S. Briggs, J. Piel and N. J. Oldham, *Angew. Chem. Int. Ed*. 2013, **52**, 1143.

[15] M. Jenner, J. P. Afonso, H. R. Bailey, S. Frank, A. Kampa, J. Piel and N. J. Oldham, *Angew. Chem. Int. Ed Engl.,* 2015, **54**, 1817.

[16] W. Cao and E. M. De La Cruz, *Sci. Rep.,* 2013, **3**, 2658.

[17] S. Dutta *et al., Nature,* 2014, **510**, 512.

[18] F. Sievers *et al., Mol. Syst. Biol.,* 2011, **7**, 539.