Supplementary Figure 7: COBALT sequence alignment of DNA-PKcs from selected metazoan plus choanoflagellates, platyhelminths, paramecium, slime mould, amoeba, fungi, green algae, liverwort, moss and oomycetes. (All phyla, expanded list). See Supplementary Table 1 for additional information on sequences. The human sequence is highlight in yellow. Conserved amino acids are shown in red and blue.

Panel A: Key to identification of sequences. See Supplementary Table 1 for details.

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

Accession	**Description**	
lcl|Query_CAPIT	CAPITELLA	Annelid worm
lcl|Query_LINGU	LINGULA	Brachiopod
lcl|Query_LIMUL	LIMULUS2	Horseshoe crab
lcl|Query_SPIDE	SPIDER	Spider
lcl|Query_SCORP	SCORPION	Scorpion
lcl|Query_TICK_	TICK	Tick
lcl|Query_MONOS	MONOSIGA	Choanoflagellate
lcl|Query_ROSET	ROSETTA	Choanoflagellate
lcl|Query_BELCH	BELCHER	Lancelet
lcl|Query_CIONA	CIONA	Tunicate
lcl|Query_CAECIL	CAECIL	Amphibian
lcl|Query_CAPI	HUMAN	Human
lcl|Query_MOUSE	MOUSE	Mouse
lcl|Query_XENOPUS	XENOPUS	Frog
lcl|Query_PLATY	PLATYPUS	Platypus
lcl|Query_TASDE	TASDEVIL	Tasmanian Devil
lcl|Query_CANAR	CANARY	Canary
lcl|Query_OSTRICH	OSTRICH	Ostrich
lcl|Query_TURTLE	TURTLE	Turtle
lcl|Query_GECKO	GECKO	Gecko
lcl|Query_DANIO	DANIO	Zebrafish
lcl|Query_SNAKE	SNAKE	SNAKE
lcl|Query_STERL	STERLET	Fish
lcl|Query_MILII	MILII	Shark
lcl|Query_BAMBO	BAMBOO	Shark
lcl|Query_ALLIG	ALLIGATOR	Alligator
lcl|Query_GAR__	GAR	Fish
lcl|Query_TETRA	TETRAHYMENA	Tetrahymena
lcl|Query_PARAM	PARAMECIUM	Paramecium
lcl|Query_SLEMN	SLEMNAE	Ciliate
lcl|Query_STENT	STENTOR	Stentor
lcl|Query_NEMAT	NEMATOSTELLA	Sea Anemone
lcl|Query_PISTI	PISTILLATA	Coral
lcl|Query_CORAL	SOFTCORAL	Coral
lcl|Query_HYDRA	HYDRA	Fresh water polyp
lcl|Query_LATIM	LATIMERIA	Coelacanth
lcl|Query_SHRIM	SHRIMP	Shrimp
lcl|Query_MCRAB	MUDCRAB	Mud crab
lcl|Query_AMPHI	AMPHIPOD	Amphipod
lcl|Query_BARNA	BARNACLE	Barnacle
lcl|Query_COPEP	COPEPOD	Copepod
lcl|Query_DAPHN	DAPHNIA	Daphnia
lcl|Query_STARF	STARFISH	Starfish
lcl|Query_STRON	STRONG	Sea urchin
lcl|Query_SEACU	SEACUKE	Sea cucumber
lcl|Query_FUNGI	FUNGIVORUM	Amoeba
lcl|Query_DPURP	DPURPUREUM	Amoeba
lcl|Query_HALBU	HALBUM	Slime mould
lcl|Query_GLOBO	GLOBOSUM	Fungi
lcl|Query_PALUS	PALUSTRIS	Fungi
lcl|Query_CALIF	CALIFORN	Fungi
lcl|Query_GLOMU	GLOMUS	Fungi
lcl|Query_LOBOS	LOBOSPOR	Fungi
lcl|Query_RACEM	RACEMOS	Fungi
lcl|Query_PVIOL	PVIOLACEUM	Slime Mould
lcl|Query_BOMBU	BOMBUS	Bumble Bee
lcl|Query_AEDES	AEDES	Mosquito
lcl|Query_ANT__	ANT	Ant
lcl|Query_WHITE	WHITEFLY	Fly
lcl|Query_BEETL	BEETLE	Beetle
lcl|Query_THRIP	THRIP	Insect
lcl|Query_TERMI	TERMITE	Termite
lcl|Query_FLEA_	FLEA	Flea
lcl|Query_MIDGE	MIDGE	Insect
lcl|Query_LAODE	LAODELPHAX	Planthopper/bug
lcl|Query_LOTTI	LOTTIA	Limpet
lcl|Query_APLYS	APLYSIA	Sea slug
lcl|Query_SCALL	SCALLOP	Scallop
lcl|Query_CRASS	CRASSO2	Oyster
lcl|Query_OCTOP	OCTOPUS	Octopus
lcl|Query_POMAC	POMACEA	Snail
lcl|Query_WHIPW	WHIPWORM	Worm
lcl|Query_TRICH	TRICHINELLA	Parasitic Worm
lcl|Query_AASTA	AASTACI2	Mold
lcl|Query_AEUTE	AEUTEICHES	Mold
lcl|Query_PLASM	PLASMOPARA	Plant pathogen
lcl|Query_BREMI	BREMIA	Mildew
lcl|Query_OLIGA	OLIGAND	Oomycete/fungal parasite
lcl|Query_LIGNO	LIGNANO	Flatworm
lcl|Query_CESTO	CESTODA	Tapeworm

lcl|Query_OVI2_	OVI2	Liverfluke
lcl|Query_PRIAP	PRIAPULA	Worm
lcl|Query_ROTIF	ROTIFER	Rotifer
lcl|Query_SPONG	SPONGE	Sponge
lcl|Query_TARDI	TARDIGRADE	Tardigrade
lcl|Query_TRICH	TRICHOPLAX	Placazoa
lcl|Query_OSTRE	OSTREOCOCCUS	Green algae
lcl|Query_NITEN	NITENS	Green algae
lcl|Query_CLUBM	CLUBMOSS	Moss
lcl|Query_MOSS_	MOSS	Moss
lcl|Query_LIVER	LIVERWO	Liverwort

Panel B: Sequence alignment from COBALT

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1     M[ 5]--EQLAGLVSDTR-GQ    PGVVNARISHES   IGNIGAI   CLNDV-NASDV--DLAYSVIFNHET   G-  61
Query_LINGU  1     M[ 5]--DLLCTLHSCVGgEE    EDARGGRKTFDL   VRDIANL   CLKET-SDRET--DYCSSLLFNKDT   G-  62
Query_LIMUL  1     M[ 6]LYDQLISLQECLD---    INSSDYEKADNI   IKDII--[2]CMDSRMSEKET--AYMCSLVFNEDL   G-  63
Query_SPIDE  1     M[ 6]IQSHLDALK-FAQ---    DSDSISEKEVSI   LETIQKL[5]NFD--LKEQER--EIGNARIFCND-   G-  64
Query_SCORP  1     M[ 6]LHENILKLNDFIK---    YQKDEGHSSIYV   IKDIS--[2]CLRGTLSEKDR--SYVCSLILRKED   G-  63
Query_TICK_  1     -[ 2]-----EQIREALQ-RL    NASDHLAEALPL   AEEVERT   CVEQSFSVKDL--GYVTSLLAGDAQ   S-  55
Query_MONOS  1     -    ----------------[24]GGSTRETLAYEL   VNDCHGH[2]----ACTDARH-lDLLATQLLSEQD   G-  66
Query_ROSET  1     M[ 7]-ERLCRNLEEQRLsAS[ 5]-------AAAAI   VADIAAY[2]---TLTTPPEL--ALAASILLGSDQ   G-  63
Query_BELCH  1     M[ 7]--ARLSELHALLA-ES    DHGHGAVEAHNL   VADIAQI   CLRHT-AERDV--AYCCSVLFQEST   G-  63
Query_CIONA  1     M[ 2]LRECLTKLQQSLL---    --DSDYIQVEVQ   IKEISQL   CNQRF-SKQEL--GLYCSLLLNKSS   G-  56
Query_CAECI  1     M[12]VQGSLLRLHASLS-SK    DPGLEAPVAQSL   IRSLGEE   CVTSS-GDSVL--ALQTSLIFSKEH   G-  70
Query_HUMAN  1     M[ 6]VRCSLLRLQETLS-AA    DRCGAALAGHQL   IRGLGQE   CVLSS-SPAVL--ALQTSLVFSRDF   G-  64
Query_MOUSE  1     M[ 6]VRCWLLQLQEFLS-AA    DRCSAAGASYQL   IRSLGQE   CVLST-SSAVQ--ALQISLVFSRDF   G-  64
Query_XENOP  1     M[26]ILGRLHQLHGLLS-LE    S-GVGAEGAHSL   LHNLAEE   CLVSALGSSAL--DLNTSLIFSKEF   G-  84
Query_PLATY  1     M[ 6]PQGSLLRLQRFLE-AE    DAGVSAVAAYQL   IRGLGQE   CVLSA-GPAVQ--ALHTSLVFSKEF   G-  64
Query_TASDE  1     M[ 7]LQGSLLQLQELLS-GG    DPGSAALAAYHL   IRAVGQE   GVLCT-SPAVL--ALQTSLIFSKEF   G-  65
Query_CANAR  1     M[ 2]VQQCLLQLHRCLQ-PG    DAGGSALHGYSL   LRSLAET   CLASL-ASGEQ--ALHLSLVFSQEY   G-  60
Query_OSTRI  1     M[ 2]ER--------------    ------------   -KDVQAD   CVFAL-PA-----ALHISLVFSQEY   G-  31
Query_TURTL  1     M[11]IQGSLHQLSCALQ-DP    DPGTAALRGYHL   IRGVGEA   CVTST-GDGAR--ALQTSLVFSKDN   G-  69
Query_GECKO  1     M[10]LQGLLLQLHGTLQ-SP    DPASAALHGHSL   IRSMAET   CVTSS-GDVAR--ALQTSLVFSKED   G-  68
Query_DANIO  1     M[ 9]IQGYLLKLHSSLE-DT    VSTNVAIVCHDI   IGDLGQE   CMITK-NENEL--VLQTSLLFAKEE   G-  67
Query_SNAKE  1     M[10]LQGFLLQLHDALR-SS    DASSAALRGCSL   IRSVAES   CVTSS-GDDIL--ALQISLVFSKEN   G-  68
Query_STERL  1     M[12]IQGNLHKLHSFLN-AT    DAKSAAVESHNL   IGNLGQE   CMVSK-NENEL--VLQTSLIFSKEN   G-  70
Query_MILII  1     M[ 9]TLSAFSKLHHLLS-LS    DVKSAAAESHSA   VGSISRD[2]CSSSS-D-SQL--ALQSSMVFSKDT   G-  68
Query_BAMBO  1     M[14]VERRLQQLEQLLS-AQ    DSRSAAVESHSI   VANLTHD[1]LHNTH-E-SRL--ALHSSLVFSKDE   G-  72
Query_ALLIG  1     M[ 6]----------------    ------------   -------   -VIRT-GDSAQ--ALQSSLVFSKDQ   G-  29
Query_GAR__  1     M[ 6]IQGYLQKLHSLCG-AG    ETKQTALECHSV   IGDLGQE   CLVVR-NENEL--VLRSSLLFSKED   G-  64
Query_TETRA  1     -    ----------------[12]LYQGNFTDSKQI   VQAISDI[1]-LDETLDDSQI--SLCNSKIFNSDL   N-  55
Query_PARAM  1     -    ----------------[11]LAAGNYQRASDI[1]-DSIENI[2]--ENQLNHAEI--GLCSSKIMASEK   S-  54
Query_SLEMN  1     -[ 4]IDELLRQLHKFLF---    ------ADSEQV   VNELRIY[1]-LDESLKEADL--AFISSKIFNSNS   S-  54
Query_STENT  1     -[ 4]VEELLNTLHSLTF---    ------TDPESI   IDEIRAI[1]-LSPNLSMPSM--AFFSSKLLNSEV   S-  54
Query_NEMAT  1     M[ 5]--SRLLQLNELVA--E    ESPLRGQESAEI   ARDLRAH   CSQLR-DESELg-ARCSSILFDDKT   G-  61
Query_PISTI  1     M[ 5]--RYLTHLHSLPS--L    DAPHCGLDASNT   VVDLGSF   CLENV-SSSEQ--AYYSSLLFSGET   S-  60
Query_CORAL  1     M[ 5]--RYLNQLNSCLK--T    LSSSRGREAADI   CRDLRLY   CIENI-NETQL--AYCSSLLFNQDD   G-  60
Query_HYDRA  1     M[ 2]E---LTRLLETIFlLR    TKNSDSIEVSST   LNNIFSL   A-LNI-ADNEI--GYFSETLFNENN   G-  57
Query_LATIM  1     M[12]VSGYLLKIHSSLT-VT    DTKSAAIEISNI   FGNLRQE   CLVTS-DENEL--ALHISLIFSKEI   G-  70
Query_SHRIM  1     M[ 8]LDLLLGKLQSFGP-GN    GQ-GDPDAAKDV   IVDIAEL   -LAVDLTPAQY--DLYISQVLGEK-   G-  64
Query_MCRAB  1     M[ 8]QDLL-EKLHHYSS-GN    EL-RNARSARDI   VVDISEH   -LTKDLSPVQY--DLFLSHVFGEK-   G-  63
Query_AMPHI  1     M[ 7]VQILLDKLENLVY-ND    ND-SNGGFAKDI   IVDLSEL   -LSSDTTSAHY--DVHVSCSLGKN-   G-  63
Query_BARNA  1     -    ----------------[20]ELHGDPERAIGL   ILNIKRY[2]---EQFTDTDF--QFILSVLFPPND[3]G-  65
Query_COPEP  1     -    ----------------[ 6]DVGELCELLSHL[6]IQTLNSA[5]SVYGLLSELRTlnDDFPWLSFIGTD   G-  62
Query_DAPHN  1     -    ----------------[11]MHSATCSNVAEQ   VNDIKEI[2]---KEKNESVIp------VLLNRQN   G-  49
Query_STARF  1     M[ 6]--VKLSELNGLIS-TK    SGLHVAEKAQDL   VVDLGHL   CLRQL-SENEL--AYCCSQLFNKDN   G-  62
Query_STRON  1     M[ 6]--RHLAELHEVID-SG    GTAHGPETAQDL   VVDIAQL   CLNQT-RENEL--AFCRSELFQKET   G-  62
Query_SEACU  1     M[ 6]--AKLSDLYELLK-SS    RDTQAAEEAKNV   IVDLGHV   LLRPL-SDTEQ--AFATSEVFHKER   G-  62
Query_FUNGI  1     M[29]LEILLAGIFHLLGpNG[30]SSKDFEGDVELI   -------   --------SSA--SMISSLLFVGNH   S-  104
Query_DPURP  1     M[20]IENYLNKFFQTITdNQ[17]NSFKFEND-LLL   INELSVI[1]-LKELVLDEEI--GLVASLLFTGDH   S-  96
Query_HALBU  1     M[13]IDGYLVKLHQSV----[26]NFSAIDKDGMLI   -EQLSNV[1]-LKELA-EDEI--GVAASLLFHGDH   N-  93
Query_GLOBO  1     M[ 1]---------DLITqLE[ 3]SRPDIA---EAA   VVDLQNQ[1]-V-RMH-ADRK--PLAMSALFNKQQ   G-  50
Query_PALUS  1     M[ 8]LRRCLVELNECVSkSA[ 3]STKALAEGIQAA   LGDLDHQ   ---------DD--ALATALLFDAER   G-  62
Query_CALIF  1     M[10]LNELSIINRDVKSiSV[ 3]NFPYKM------   ---LNNK[1]-I-----------VYAMSLLFNNEN   G-  55
Query_GLOMU  1     M[ 9]LQELA----NCRQnIV[ 3]SFGEAVGEALRY   SQDIDHI[1]-SNDIE-DQEL--GLATQLLFDPQI   G-  67
Query_LOBOS  1     M[10]LVELL----D-EMnGE[ 3]NVDSQV-----I   ATNICSM[1]-M-ELK-EDEI--DMATSLLFNPER   G-  61
Query_RACEM  1     M[ 5]IEACL----QDLSeAL[ 3]LSLKSI------   ASDLATR[1]-MADVD-DIDR--DYVCSLLFDPEN   G-  57
Query_PVIOL  1     M[27]LSNYLRKLYDIV----[ 8]ASNSLESDSVII   IEELKNV[1]-LKELVMEEEI--ALASSLLFSGDH   S-  91
Query_BOMBU  1     -[ 3]FKTFMDIFERAVT-ER    ----NSTNLTEI   LKDSKKY[1]--Q-NIAKDD--sELMLSILFDKKK   G-  55
Query_AEDES  1     -[ 1]FKRF-EHVRHW-S-TP    ----SPSRAPPS   LPKKTVY[1]--Q-----------YYLDTLINSPN   G-  43
Query_ANT__  1     -[ 3]YQTFIDIFKQAAS-HQ    ----YFKDLPNI   LKDGQQY[1]--K-LVSLEH--fESVVCSIFDERD   G-  55
Query_WHITE  1     -[ 5]EI--LQDLSHIPS-EP    -DAQAAENYKQR   LIALFSH[1]--KAQVDNLD--yDLS--ALFNEET   G-  57
Query_BEETL  1     -[ 4]PRDFLKLLYKYLD-EE[ 4]-ASKQCNQLLTN   FSDALTH   ---EQINSTY--lDQSLHYFFSTSS   G-  62
Query_THRIP  1     -[ 4]----IESLWLCLP-A-    ----KEHDAKIL   IHRIACH[1]--QNYFTSED--iDLSLHLLFQEEK   G-  52
Query_TERMI  1     -[ 5]LIGFVEVINSCIS-GN    -HLNKYIEAKSN   ISNIISY[1]--EDKRKRTD--lDGAFTVLFDKKT   G-  61
Query_FLEA_  1     -[ 4]VEDTFKDLKLCIQ-TK    ----QGTKAKLC   ISNILRF[1]--KGEVTSSE--iDFALSIIFNEDD   G-  57
Query_MIDGE  1     -[ 6]VKRLLEKLRQQLK-NK    ----NQSKCLEI   IGILTEH[1]--EGLASPSE--v-YYAVELLLEWN[4]D-  62
Query_LAODE  1     -    ----------------    -MENELESIHDL[8]VGNLKTL[3]IKENLMASDKHglDKNIAILFGHDD[4]Gk  60
Query_LOTTI  1     M[ 6]--ESLRDLHANL--KD    GGQSELDQIDSI   VPTLSQI   CLHEI-TEKDI--DYCSSVLFDKEI   G-  61
Query_APLYS  1     M[ 5]--NFLQSLHDFI--DA    SISSNREQIEVI   IDDIRQV   CQHEI-KEEET--EYVCSLVFAKDT   G-  60
Query_SCALL  1     M[ 6]--NNLRSLHKCL--KA    GDASAVEQSNGI   ISDISQL   CLHEI-DDKEI--DLCCSALFDKDK   G-  61
Query_CRASS  1     M[ 6]--DNLRSLNECL--TT    ADNSSKEQADSI   VADISQI   CINES-TDKDI--DLCCSVLFDKRK   G-  61
Query_OCTOP  1     -    MAREIKRIEQCLN---    DLWNGKDDRKQI   VEDIA--[1]ICPNQLTVKNI--ESVCTLLFDKEI   G-  55
Query_POMAC  1     M[ 5]--ETLRSLHD-----Q    SGASAYNLANGL   IADIRQI   CLHEI-SEKDR--DYCCSALFDKDI   G-  57
Query_WHIPW  1     M[ 7]LWSKLQSLKTSAD---    KGDGAALQAIED   LQDIKHL[4]----SGQEGDF--DFVVSHFAAKTD[1]A-  65
Query_TRICH  1     M[ 7]IRDALNSLKDFIE---    GADSD--QAEDC   VQQIFNH[4]----NGSERNF--GFCFNSLLSNDS[1]S-  63
Query_AASTA  1     -    ---MSQKCHHLDQvDR[12]DVEVDKDTIESF   VAALFRR[3]CLLESRAAAEL----CCFRLFQTTP   Gq  70
Query_AEUTE  1     -    ---MNDKSNRFKElDR[12]DVEVDKDTIESA   VTTLFRR[3]NSLESRAAAEK----CCSRLFHGSS   Gp  70
Query_PLASM  1     -    ---MTTKLNE------[ 9]LHAQAVESIDET   LSHLRLE[3]AMILSRASAQQ----CTILLFQSAE   Pp  61
Query_BREMI  1     -    ---M-TKLSE------[ 9]AHEYAVDSVEET   LAHLQTE[3]AMALSRASAQQ----CTILLFQSTD   Pp  60
Query_OLIGA  1     -    ---MAEKVRD------[ 9]ESEHAVDSVEES   VEHLQLA[3]AMARSRGLAQQ----CTILLFQSHE   Pp  61
Query_LIGNO  1     M[ 8]-------IDKFLK---    --DPELKESKAL   FNVIFNE[1]---ASSNENFKh-NKYLAILFDEPE   G-  55
Query_CESTO        -    ----------------    ------------   -------   -------------------------   --
Query_OVI2_        -    ----------------    ------------   -------   -------------------------   --
Query_PRIAP  1     M[ 7]--DHVQTLLESLP-SD    DREADPETAHSA   VADLAHK   CLQDI-SKRDV--DLTCSVLFSEDR   G-  63
Query_ROTIF  1     M[ 4]--KHIQNLFDSLS---    --NSNVDAAKSI   VLDIVYL   CGENL-NPQTI--TFCLSLFFKREK   N-  56
Query_SPONG  1     -[ 2]--SRIESLWESLH--S    SSPQRYSYATEV   ASDLYSL   CVEQI-TESEL--AYSSSVLFNPDN   G-  56
Query_TARDI  1     M[ 4]VLSAVEELSAIVS-AT    PSLSTDDNACDA   VEQIGRI   CLAET-GLRDI--KAIIPIIAEGSD   C-  62
Query_TRICH  1     M[ 5]--QLLIDLHALLN--L    QSLHRGSNACDI   ASDITVL   CLQTA-SDAEK--AYCSSALFGQDY   G-  60
Query_OSTRE  1     -    ----------------[16]SGDGGATSTSGT   ISVIRER[4]TIDGLLASGAIh------AIFDAER   G-  59
Query_NITEN  1     -[ 4]LVFYLTELREDA----[ 2]RSLGSAEELVRR   VEDIAD-[1]-CTNLTSEDEL--AFASSVLFDEDS   G-  60
Query_CLUBM  1     -[ 4]VSQYLATLNSEA----[ 3]-LKGSETDLGVT   VDDLSG-[1]-CSQLVSDEDF--AFAASAVFDRDT   G-  60
Query_MOSS_  1     -[ 4]LGSCLASLRDAA----[ 3]RSSGGFVDLAPY   VVDFSE-[1]-CKDLYSDEDI--SYAIAIVFDEDT   G-  61
Query_LIVER  1     -[ 4]LVRCMTALKDEA----[ 3]RMGASASDLNSL   ISDVAD-[1]-CKQALSDEDV--SFALALVFDRED   G-  61
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  62    -LIKFLKKII---Q   NDEFEG   CKCE--ILELLKTFISKF-PG-KVHP   YAVDVKDACLQLIT---RD-RRARV   120
Query_LINGU  63    -LVSFVKKVV---K   QAEYDS   TKTA--ILETLCLFMERA-EK-KVLP   YAVDIKDVCITLFI---RE-EKAAV   121
Query_LIMUL  64    -LTVFLKKTA---S   LSQFQN   GKLE--ILDFLPKLLDNV-KK-KAVS   YGVRIKDACFSVFL---RD-QFARV   122
Query_SPIDE  65    -ILSLLRNLL---I   VKELRN   VKLN--CFKLLNDIIESS-GS-LIGE   YGFDILEVCKKYMK--qRD-ATAEV   124
Query_SCORP  64    -FLHFLEEAV---N   TPSLKD   SLIE--ALQFLNLFLINI-GN-KVVQ   YGITVKDVVFSIFL---RN-QLAKV   122
Query_TICK_  56    -LVSFLKKHV---A   TDQLKD   AKAK--SLNILTGWFERV-PA-SQLL[1]HALAIKDVSCQLLF---RD-KAAKV   115
Query_MONOS  67    -LLGFFATLR---N   RQKLPQ[8]AKAH--ALACLGLLVARL-STeTLTR   YLQPLMRHAQALLQ-ikSE--KSAV   135
Query_ROSET  64    -VLSFLRQLGl---   --ELYE[1]CKVA--ALEFLAGFIQRLpAD-RLTP   YAVRIKDTCLTCFR----D-KRAKV   121
Query_BELCH  64    -ITAFLRKTV---T   LDQYLP   AKVE--TLSFLLAFLEKI-GR-KIQP   HAVEVKEVCMAVFS---RD-RLSRV   122
Query_CIONA  57    -ILLVLNKNV---S   LAGLKP   SRDL--VLTFLVTFVPRV-GS-YVLQ   YVDDLRSSCLNVFR---SD-GYSRN   115
Query_CAECI  71    -LLAFIHKSL---G   IEEF--[2]CREE--TLKFLCIFLEKI-GQ-KVQP   YACDIKNTCVVVYT---KE-KAAKC   129
Query_HUMAN  65    -LLVFVRKSL---N   SIEF--[2]CREE--ILKFLCIFLEKM-GQ-KIAP   YSVEIKNTCTSVYT---KD-RAAKC   123
Query_MOUSE  65    -LLVFIRKSL---S   IEDF--[2]CREE--ALKFLCVFLEKI-DQ-KVMH   YSLDIKNTCTSVYT---KD-RTAKC   123
Query_XENOP  85    -LLAFVRKSL---S   SDEF--[2]CREE--ALKFLYTFLEKI-GS-NVQP   YAMDIKTLCVIVYT---KD-RAAKC   143
Query_PLATY  65    -LPVFIRKSL---D   IEEF--[2]CREE--ALKFLCIFLEKI-GQ-KIEP   YSYDIKSVCISVYS---KD-RAAKC   123
Query_TASDE  66    -LLVFIRESL---S   IEEF--[2]CREE--ALKFLCIFLEKI-GQ-KIEP   YSLGIKSVCTSVYT---KD-KAAKC   124
Query_CANAR  61    -LLVFIHRSL---T   IEEF--[2]CREE--ALKFLCVFMEKV-GD-KIHP   YACNIKQTCLSVYI---KE-RTAKC   119
Query_OSTRI  32    -LLVFIHKSL---G   IEEF--[2]CREE--ALKFLCAFLEKI-GE-KVHP   YACNVKQTCISVYT---KE-RAAKC   90
Query_TURTL  70    -LLVFVRKSL---S   MEEF--[2]CREE--ALKFLCAFLEKI-GQ-QVHP   YAHDIKQTCISVYT---KE-KAAKC   128
Query_GECKO  69    -LLAFIHKSL---S   TEEF--[2]CREE--ILKFLCAFLEKI-GE-KIHL   YAHDVKRTCIIVYT---KD-KSAKC   127
Query_DANIO  68    -LLSFLRRSL---S   TEKLGT[8]TRVE--IMNFLGAFLQRM-SA-TVRG[4]YAVELKDTCIVVYT---KD-KSAKC   138
Query_SNAKE  69    -LLSFIYKSL---S   VEDF--[2]CREE--ALKFILAFVEKI-GP-KIQP   YAQDVKRICVTAYT---KD-RSAKC   127
Query_STERL  71    -LLAFVRKSL---N   IEDF--[2]TREE--ALKLLTTFLDRI-AH-KVNG[4]YACDIKDTCLVVYT---KD-RAAKC   133
Query_MILII  69    -LLVFLRKSL---R   FEEF--[2]SRED--VLKLLSEVVERI-GE-AVGR   YACDIKGTCVVVYT---KE-KAAKC   127
Query_BAMBO  73    -LLIFLRKSL---S   NEEF--[2]CREE--VLKLITAFVEKI-KE-TVGP   YALDIKDVCLVIYT---KE-KAAKC   131
Query_ALLIG  30    -LLVFIRKSL---S   IEEF--[2]CREE--ALKFLCTFLEKI-GQ-GVHP   YACDIKQTCISVYT---KD-KMAKC   88
Query_GAR__  65    -LLSFLRKSL---S   SEDF--[2]VREE--VLKLLSTFLQRI-TQ-SVKG[4]YAVDIKETCLVVYT---KD-KAAKC   127
Query_TETRA  56    -LFTYLENNR---T[1]KDNLQA[1]IRKG--LFRTIAAYIKLR--KlKVFE   YIELICKNCMSLFK---KE-ESNQA   116
Query_PARAM  55    -FLEFLTNHV--lD   KDTYYA[1]TKKS--LYQILAKYIKNYvQY--VSE   YLESIISTCIQNFK---KE-ESGVA   115
Query_SLEMN  55    -LLEFLRENS-lhN   -DKLIS[1]VKKK--VYELIIDYLKLR-SQ-EATP   YLKLLKDTCLRSFR---SD-HNSLV   115
Query_STENT  55    -LLQFILKNR-tdS   -DRLIW[1]TMSN--ALSLVSDYIRER-GN-RINE   YVIEIKEVCLQLMA---SStKSSLV   116
Query_NEMAT  62    -IIRFLHKTI---A   NQEFV-[1]CKTE--LLKFLQSYLGQL-NT-RVNP   YAVAIKDICMKLLF---QE-SSSKV   120
Query_PISTI  61    -IVSFLRKVI---T   LDEFQ-[1]AKIE--LLKFLQSYVMKM-GK-KINP   YVVEIKDICLKIFS---QD-HSNKA   119
Query_CORAL  61    -LITFVKAVV---Q   LDEFA-[1]CKEV--ILSFLAEYIAKV-KT-RISP   YATDIKDICLKLFT---KD-KNARV   119
Query_HYDRA  58    -ILKFLKLSI---G   SNELEN   AKGD--ILRFLSEFCTK---K-LVKF   YIKGIKETCLALFF---CD-TSVKV   114
Query_LATIM  71    -LLIFLRKSL---S   SEEF--[2]NREE--ILKFLTAFVEKI-GQ-KLNP   YACNVKLILFPVFF---F------F   124
Query_SHRIM  65    -IITFGRSVS---K   DQSFVE   AKKS--TLDVLKTLLDKE-AQ-KINK   YVDDIRKYTLYLY---nVE-KSTKV   123
Query_MCRAB  64    -IIAFGTRVT---K   EQAFVD   VRNS--TLDLIITLLMKD-TK-KIIK   YVGDITNYTLFLY---nAD-KSTKV   122
Query_AMPHI  64    -LLQFCRQTV---S   RKHYGD   AKKS--ALEVIRILLEKQ-AS-KVAQ   YTDEIFVSSVLLY---rGD-PAAKV   122
Query_BARNA  66    -LMGLLQDVR---E   IKEFGS[1]-KVEg--CRVTADLVSRA-GAaSLDC[1]-AAGIRSTACLLSG-rpQE--TSAV   126
Query_COPEP  63    -LFPKLKKIA--gE   -KDFAN[2]--QD--VLRLLVLLLEKNpEQ-----[3]HLADFQVLSDSLYF---MN-PAAKV   121
Query_DAPHN  50    -LVPLLSKFR----   --DFRK[6]IIKN--VLEIFGHLLEQF-----IEA[3]HLVVIKNSTLQYMN--sKD---AKL   110
Query_STARF  63    -ILVFLNKAI---P   KDELKG   SKEQ--LLEFLHAFLLKA-GS-RLLA   RTVEIKKICISAFM---RD-KSAKV   121
Query_STRON  63    -IITFLQKAV---S   KDEYGG   SKVQ--TLEFLLGFVQKA-KE-LVLD   HAKDIKDVCMSLYT---RD-KLAKV   121
Query_SEACU  63    -LLSLLKSIA---A   KEEFTG   CKEE--ALRLIGGFLKKR-HA-HLVE   HAVDIKDTCTALYL---RD-KSSKV   121
Query_FUNGI  105   -LLKIVSLSY---G   DQEKHI[2]TKES--LFKLLKNWLKTI-GI-RARD   YAQDIREVCYQSF---rRE-QSNSV   165
Query_DPURP  97    -LLKYLEKTS---T   LNTKEI[2]IKVA--ILNLITEFIEIL-GG-RTDD   YAIAIKNTCISIF---rKD-QSNSV   157
Query_HALBU  94    -LLKFLERSV---L   INTKDM[2]LKTS--ILNLISKLIKSF-KL-QVAD   YVVPIKSICISLF---rRD-QSPTV   154
Query_GLOBO  51    -ILQLLR-------   ----TN[2]A------------FTSKA-KY-MMLD   FIAELTT-CWLFFS--nSK-ELSAV   93
Query_PALUS  63    -ALQFLRLS-----   LGNKVH[2]AKAV--ICKALAGFMYET-SS-SLRE   YAIDIQSICQSLFS--sFK-ESSKV   122
Query_CALIF  56    -ILTFITRT-----   MKYKEH[2]SRQI--LLNLIHDLLENF-PN-YLTE   YIDKIRRISENFF---fTE-KAART   114
Query_GLOMU  68    -IIKFLNDTS---R   VTDSVL[2]AKIA--LLEFLFLYIRKI-QS-DIEP   YTEIIKNVHFGIL---nVD-DEAFV   128
Query_LOBOS  62    -IMLFLRKA-----   LDKTAH[2]AKVT--FLEFTAEFISRS-RN-AITP   YIVDIKRTCLNLF---lND-STARV   120
Query_RACEM  58    -ILRFMRYNG---S   TKDSTA[2]AVTT--LLDFLVSFIQKA-GT-SLQR   YAPSLKQCCVALI---nAP-S-SRV   117
Query_PVIOL  92    -ILKLIEKTI---T   LNSKEV[2]VKTT--LLSLISEFIVIL-QK-RTEN   YALAIKNLCISVF---rRD-QSASV   152
Query_BOMBU  56    -LLSFLNNEL--gK   YKPEKG[4]-VKE--AFHLLEFIIEQF--CdIFIP   YIVGTKNVCQQAL--ktR--CSALI   118
Query_AEDES  44    -ILSFLQRS--ikA   SGRFQK[1]-NEE--IFDLLRRIIQK-h-SqLVRK   SVESVVRECVRFV--qsSS-VSARE   104
Query_ANT__  56    -LLWFLNHVI--kK   YQVIKR[5]-LKE--AFNFLEFFIIHFyITpCFVQ   YIQKIKYVCQIVL--kkKN-CNYYI   122
Query_WHITE  58    -IITLLHKLL---R   KTKYKS[1]-VSY--TLSEISETILKDh-KnRVLQ   HVEKIAGFCVKCA--qeDS-QSSDS   118
Query_BEETL  63    -LHTFVKKII---N   KKPFKT[1]-VVK--FFEIIYALINN-f-PfKLKK   YVSVIFETCFKVI--ksLL-VSSDE   122
Query_THRIP  53    -LKSFLD-AL---S   NSNMKK[1]-SSD--VLNFLEFLIVK-l-PfSLEN   YVVIIRDIAMSYV--qrTS-SSAEE   111
Query_TERMI  62    -ILQFIQKAL---K   QDAFTE[1]-IKE--SFTLLKQLIDV-f-KeRMSC   HVVQIKDLCKTCI--l-LE-KKAEE   120
Query_FLEA_  58    -IIPFIKQHI---S   EVYLQE[1]-IKF--SLDLLLLLVEE-l-PiKIRE   HIKEISKISIKTL--ytDK-M-AKE   116
Query_MIDGE  63    -VLRFLAKSLhilK   SEKFNE[1]-NKK--VFELIQTLTKK-h-PdKIKI   YSKNIVDICTIYI--qnVS-TSAIE   125
Query_LAODE  61    nVVAFIKSCLssrS[2]VVEFLK[5]CVLYp---EALINYVEKI--------   --IDISKLCVLSKS-vtTN---TSV   121
Query_LOTTI  62    -VTTFLQKIS---K   KNEYQG   SNAKygLLELLSDFIHKV-GK-KALP   YLVEIKEASLSNYM---TD-RFTKI   122
Query_APLYS  61    -LISFLQKTQ---N   SEEIQN   GKSA--ALDIISGLIQKV-GR-NVLP   YAVDIKEATLAVYS---CD-RFAKV   119
Query_SCALL  62    -IITFLNKTV---N   KDEFLG   CKQK--LLELLSGFIQKV-EK-KILP   YTVDIKDVCVSLFF---RD-RFAKV   120
Query_CRASS  62    -LVSFLSETV---T   KEEFQT   CKQK--ILEFLSEFVQRI-GK-KIIP   YVFDIKDVCVSLFF---RE-RLAKI   120
Query_OCTOP  56    -VTTFLRSNA---N   QPT--E   IRHD--ILAMLINVQQKF-GE-YLQP   YAADLKGICKLQYN---QD-DSAKV   112
Query_POMAC  58    -IIAFLQKVA---A   KDEFQS   TKTN--LLDFIGDFLQKV-GK-KLLP   YAIDIKEACVSAFS---RD-KYAKV   116
Query_WHIPW  66    -LLGILYRFL---P   LPSFAD[4]----g--VDILRTVICKA--S-GIRS[1]QRKTMICFALKCAL---VD-HLKGK   124
Query_TRICH  64    -LLSILSSSI---G   KTAFAQ[4]----g--IDIIRLAIAKT--T-SLSP[1]HRKAIMRFTLRLLK---QD-YLKGR   122
Query_AASTA  71    qHAPSLLSIMqlqD[4]KSTLGV[1]-LGI--ALEHLLTFIKDMqDTtLRHA[3]HVGSITQACCILLL--sSQ-LPSKT   143
Query_AEUTE  71    --TPSLLDILqlqD[4]K-TYAT[1]-LAF--TLEHLVTFLKDMhDT-IRQS[3]HVETLVPTCCNVLT--sSQ-MASKC   139
Query_PLASM  62    sLLRFLSTSAeftD[3]KREISA[1]-RIG--VFELLSAFLKVYgAH---SA[4]HVIDLYKVC-QQTA--rSD-PFNRV   130
Query_BREMI  61    nLLRFLSSSAnfaD[3]KREISA[1]-RGG--VLDLLAKFLKVYgTH---PA[4]HVVDLFKAC-QQTA--hAD-SYSRV   129
Query_OLIGA  62    sLMEFLDQSGrmsE[3]KREISH[1]-RIK--VLELLVGFVKTYsQH---AA[4]HIVNTMKRC-QAIA--rSD-WSNKV   130
Query_LIGNO  56    -LVYRLQS------   -KEVSE[1]AKKDg--LDFLTKLIAQS--P-ALLA[2]-LPKVKCLAESLLF--sSI-SSAKV   112
Query_CESTO  1     --------------   ------   --------------------------[3]RVVCWVETSYSGFVc---S-KSAKL   24
Query_OVI2_  1     --------------   ------   --ME------LNAGIPYTfQRaQTNK[1]-TSDSRINCKDWFL--hRR------   35
Query_PRIAP  64    -VLIFLRRAI---P   LDEFKL   AKKE--LLQLLHSFINKG-QT-RILT   YAVDIKDTCLAVFS---RD-RQADV   122
Query_ROTIF  57    -LLTFLRKTI---T   RDEFRG   CKVD--LLQFLEKFITSA-EK-KILP   YSPDIKDTCILLFS---SD-KYSDV   115
Query_SPONG  57    -ILKFLHDVT---T   LEEFAS   CKEE--LLKLLSNFVKNL-SR-RVLS   HIVPITDVCVFIFM---RE-KSAKV   115
Query_TARDI  63    -FCAILADAA---G   ASLVAT[5]---Qf-LIELMEAVVQQE-AL-RKCV[3]-LPRIWDTTYAIWR---IC-SAAKI   126
Query_TRICH  61    -VVTFLRKSI---G   LDDLD-[1]CKED--LLKLITAYMEKI-GG-KISP   YAVQIKDVCISLFT---RD-KAARV   119
Query_OSTRE  60    -LTAYVRETEgraG   ------[5]ARVK--ALETMYECAWRLgERdGIGE[2]----IRATWDVVSGtvrRE-GSNAV   123
Query_NITEN  61    -IFAFLEAAL----   -PEKAL[2]PREK--ALAFLAKYIEKV-KD-RIAP   YALFIRKKCYDLY---rRE-TSNSP   119
Query_CLUBM  61    -IFCVVDRLR---I   SRERSV[2]AREK--CLVFLGEFVRSV-GPeRASL   YCSLIKENALRIL---rID-ETRSV   122
Query_MOSS_  62    -IFSILESLL---R   VRDKSV[2]ARES--CWSFLALFIKKI-GP-RALH   YAALLKEKSLTFF---rRE-DSKSV   122
Query_LIVER  62    -IFSIVQLLL---S   IKERSV[2]ARET--SLNFVSEFIEKV-GE-RILP   HVSFLKENCLLIF---rRD-DSKAV   122
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  121   KNNATPLLIELLKQTSGLPV   AGD-LQ---VEGMIER-LFKELIKS-PK--PAASVKAHI--FHLLG    VIAELFP  183
Query_LINGU  122   KKVVIPVLIKLLELTEGSQM   SEE-LN---MREMIEL-FFMELSRK-TK--HQASVKREL--YKLLG    IFAEVYP  184
Query_LIMUL  123   RSGALVCVAKILETCRGTSQ   VQD-FD---IPDLVKK-MFMQLIQAs-K--LPPTVKQGI--LCLLG    SLSEYYP  185
Query_SPIDE  125   KKASMITISKILETCHGYNG   EKR-ID---VKSLIED-LFFQLTTGa-K--LSPKAKEEV--LCLIG    IIAHYYP  187
Query_SCORP  123   KILALSALIKVLELCSGAIS   SEQ-FN---LQKIVNK-IIEYLAQAsSK--ITPTVKQQA--YYLLG    ISSRHYP  186
Query_TICK   116   RLQALVLLSKVLEACKGASV   CTE-LD---VPGLVER-MFMLLSQP-SK--LTPTVKEEI--LCVLG    VIAECHQ  178
Query_MONOS  136   KTAAATLMIQILTTDSLITP[2]SDAlFDelqIVNLMSM-LQGIFQSNvSAt--PSTVKERV--LHLMG    VLIRHFP  205
Query_ROSET  122   KSLSFKPLIALFSHEQLFTD[2]RGQlLRelkADDLLKK-YLDPFHSSaSKl--TQTVKRDM--LHAIG    LIISLSP  191
Query_BELCH  123   KCATFPVLKKVLQLTIHSQL   GDE-LR---VSDMVDR-MFLELTMK-SQ--TTATVKSGA--LSLLG    VISEVYP  185
Query_CIONA  116   REKALQIFNKLI---ENRKV   GIN-FDgdyIRELVSNnLFKELAKSaSK--TTSSVCCQV--FQVLG    SLARYFP  180
Query_CAECI  130   KVPALELLIKVLQILKNSYE   IEE-LK---VKEIFNK-FYSELGVK-SK--VPDTVLEKV--YELLG    VLGEVQP  192
Query_HUMAN  124   KIPALDLLIKLLQTFRSSRL   MDE-FK---IGELFSK-FYGELALK-KK--IPDTVLEKV--YELLG    LLGEVHP  186
Query_MOUSE  124   KIPALDLLIKLLQILRSTRL   MDE-FK---IGELFNK-FYGELASK-SK--LPDTVLEKV--YELLG    VLGEVHP  186
Query_XENOP  144   KIPSLELLIKLLQLLKNSII   IEE-FK---IGEIFNK-FYGELATK-SK--LSDTVLEKV--YELLG    ILGEVQP  206
Query_PLATY  124   KIPALELLIKLLQTLRSSRL   MEE-LK---VGGIFNK-FYEELALK-TK--ITDTVLEKI--YELLG    VLGEVHP  186
Query_TASDE  125   KIPALELLIELLKTLRSSRY   MEE-FQ---LGELFSK-FYGELALK-TK--IADTVLEKI--YELLG    VLGEVHP  187
Query_CANAR  120   KISALELLIKLLHTLQRSCL   MEE-MK---VGEIFNK-FYEELAAR-SR--IPDTVLEKI--YELLG    VLGEVQP  182
Query_OSTRI   91   KIPALELLIKLLQNLRRSCL   MEE-MK---VGEIFNK-FYGELATR-SK--VPDTVLEKI--YELLG    VLGEVQP  153
Query_TURTL  129   KIPALELLIKLLQNLRSSFL   MEE-LK---VGEIFNK-FYGELAAR-SK--VPDTVLEKL--YELLG    VLGEVQP  191
Query_GECKO  128   RIPALELLIKLLQNVRSACV   MED-MK---VGEIFNK-FYGELAVK-SK--VPDTVLERI--YELLG    VLGEVQP  190
Query_DANIO  139   RNPALDLLIKILYLTKDSSI   TQN-LR---IGDMFNK-FYGELCQK-HK--IPDTVLGCI--YELLG    VLGEVHP  201
Query_SNAKE  128   GIPALELLIKLLQKLQSSYA   MVD-MK---VGEIFNK-FYGEIAIK-SK--VPDTVLEKI--YELLG    VLGEVQP  190
Query_STERL  134   KVSALELLIKLLHIT-SPCL   LQE-FK---VGDMFNK-FYGDLAQK-SK--IPDTVLEKI--YELLG    VLGEVQP  195
Query_MILII  128   KAAALDLLITLLHSKLSSSC   LQQ-FK---IGELFNK-FYGDLSQK-SK--FPDTVLEKV--YELLG    VLAEVQP  190
Query_BAMBO  132   KVAALDLLLVLLHSKV---C   LQE-FK---IGDMFNK-FYGELAQK-SK--LPDTVLEKV--YELLG    VLAEVQP  191
Query_ALLIG   89   KIPALELLIKLLQNVQCSSL   MGE-LK---VGEIFNK-FYGELAGK-SR--VPDTVLEKI--YELLG    VLGEVQP  151
Query_GAR__  128   RVPALSLLIQLLHGTKSSGV   AQD-FR---VGDMFNK-FYGELSQK-SR--IPDTVLGNI--YELLG    VLAEVQP  190
Query_TETRA  117   KETSLSPIIKIIKCIDNSEL[1]-EKiIDp---KNTISA-LLNEIKFC--Kp--TQKVKGSI--WNLIG    IIFQFFT  179
Query_PARAM  116   KEASLLPIMKIYKHSDQQIL[1]---kLQv-------HKnLYHILLQEcLEtkLTAGVKGRI--FQTLG    VLIKRFN  177
Query_SLEMN  116   KESSLQLLIKIIETFNSKEI   -EEmIRp---NDMMVMlLD---EIKlRKp--SATVKGTI--WTLIG    LLHEKYE  177
Query_STENT  117   KEAALRPLIKLIEVFSPNKL   -KLiIDp---TELFIVlLESKLKYEeRTm--GSTLKGAI--FNMLG    MLVSSFE  181
Query_NEMAT  121   KIETFLPLIQLMQTKLESSI   IDE-MD---IPKLIDR-YFKACLQS-SK--QQSTVRFWL--YTILG    VFAEVFP  183
Query_PISTI  120   KAETLPLMAEILDVKLESSV   VEK-LD---VPKIFDK-YFGACIQP-VK--HSSTVKYWL--FSLLG    TIAKVFP  182
Query_CORAL  120   KNETFEVLLQLFELKFEPTI   VEK-LD---VKNIVEK-YFSACSQS-TK--KKSSVSYGI--YSLLG    TLAEYFP  182
Query_HYDRA  115   KIEVIHVIEKIMLLPDINQF   AVD-LD---IKSLISK-FWSECLKSa-K--YSSTMRGEL--YNLMG    SFAKVVP  177
Query_LATIM  125   FIFIFFTYLQLLQSTKSPHC   IEE-FK---VSEMFNK-FYGELAQR-SK--LPDTVLEKV--YELLG    VLGEVQP  187
Query_SHRIM  124   QAAALEVFCTLLAKCGSHID   IQE-IN---VDQFIER-L-CIDIRS-AK--GSTVLQHQL---QTLG    MLMYSYP  184
Query_MCRAB  123   CAHALEAFCLLLEKFSSHVS   VQD-IN---LHQLIEG-I-CIQTQR-YK--GSTILHHQL---RTLG    VLAKVYP  183
Query_AMPHI  123   RCAALELLSVLLLRCVCYLS   ADI-LN---VEQLVVD-L-SMGIRG-AK--VPSVVHHEL---QLLG    LLTFALP  183
Query_BARNA  127   RSEALQLLITVLAAGGGRST[1]-AHqMAa----DLLSK--------DlTRa-APPTVRRQQ--LHLLG    ALCQHCP  184
Query_COPEP  122   KVEGLKLLQFCIPGLKDQDV   ----FDa---EKVADR-YFRELART-SKl--SSTVSSQLr--QMLA[ 1]-LCQFYP  181
Query_DAPHN  111   RETSYSVLSQVVIAAQQSSG[1]VIDeLDff----LLSNkLFEKMAASaEDn----ASKSKMefLMLMG[ 1]-LAQYFP  176
Query_STARF  122   KTKTFPVLIELLQRSASAQM   EAD-LD---IPKMIEL-YFEQLMQP-SK--LSQTVKQGI--YCLLG    VLAEIYP  184
Query_STRON  122   KAKTFPLVIELLKLTSGTRM   DDEaMG---VDAMLKK-FFGELGQP-SK--LSQSVKLGI--YQTLG    AFGEYYP  185
Query_SEACU  122   KASALEVLTQVLQLASSHES   KVS-LN---VSSIIKK-FVEELMKS-TR--APQGVKSGL--CQLLG    VISEVYP  184
Query_FUNGI  166   RIAALKALVKLFTLHLKTLE   DGEkSAv---VDMSST-IYNEYVIG--K--SSQTVKGLM--LEALG    LFARQFP  228
Query_DPURP  158   QAAAFGPIKKILHLMSKILK[4]VAEsFGv---LDMSQL-FLLQFTCG--K--LTQTVRGEI--IVVLG    LFTEYFS  224
Query_HALBU  155   QAATFEPIQKILHM--KLTD[2]NPEvFGv---REMSDL-LLLQFTCG--K--LTQTVKAEL--LETVG    LFTEYFT  217
Query_GLOBO  94    KKSAVHVLVALVEEKL--LD   -AAtLEa-----MVKR-FMTHFVQNqSK--LVSSVKCAI--LELFG    ACVRYYP  153
Query_PALUS  123   KAESFNVLSALLDPTLNILD   -PKkIDi---QKLFQF-YLQTYVQQqTK--LTSSVKGAL--LRLLG    NISRYYP  186
Query_CALIF  115   VCASFQVLNILYNPEKKFLI   -RDnIKp---GELFER-YMNSYVQKyTK--LSTSVKGDL--LQFLG    KIIQYYS  178
Query_GLOMU  129   KAASFKPMMAMLDSSKRIID   -PEkLNv---RDIISR-YLEIFFVR-TK--GTSGVKAKAgaLDMLG    IIAHRFP  193
Query_LOBOS  121   KKSSMLPIKILLSSCSKMID   -PEvLEi---GGLFER-IFEAYGTQqSK--MVATVKAEV--LEVLG    IISRYFP  184
Query_RACEM  118   QSSALDVLATLLNPDISHLD   -TSvFRi---DSLYEQ-NKRTYASY-TK--ISASVRSKT--LILMG    TIARYYP  180
Query_PVIOL  153   QAMAFAPIKKILHMIPHYLK[2]RAEvFGv---REMSDL-LLLQFTCG--K--FSQTVRGEI--VVTLG    LFTEIFP  217
Query_BOMBU  119   QNAACSTFNMLI--ANFEEY   --EiELdktVKEFVS---LFRFVNIkERs-L-------L--FSVLG    TIAKHCP  174
Query_AEDES  105   RELGTAVIQDLL--VYKCLD   --EqYD---MSQLLGE-LLVVFDQRsKLt-N--RFQQCL--FELIG    LMAKDFP  164
Query_ANT__  123   KKAACNVFTKLI--ELFKDH   --DlKLvdtINEFKTD-EQFALYRTeEKr-F-------I--FCIVS    KIMKYNK  180
Query_WHITE  119   KEAALKILIEFI--SHNPSV   --HsLDgrqFSPFITK-LL-AMKTKfSNt-----VMKQV--YILLG    HLGNSYF  178
Query_BEETL  123   KVKSLEVVIVIL--EKEIFL   --DnS------EIYLE-VYNNLLKSiSIt-KTDSVEEKT--FIALG    TFAKHF-  180
Query_THRIP  112   KDRAFDILIALV--QNGPWN   --AdL---qITPIVTK-L---FSMAsHSt-SKPTVLRKV--YKLIG    ALAEHLP  170
Query_TERMI  121   KDKAFDVIQSVL--ENGPWN   --DnL---nVDELFDK-LINPFKSNkSNt-MPAKVVQKH--LHILG    IIVHKYP  182
Query_FLEA_  117   KSIALQILVAIV--NNKEIV   --DlLD---INYVIAQ-VVHHLEYGkLSs-L--VYEQ-L--FPLLG    ILARDFP  175
Query_MIDGE  126   KESAVQTIHEVL--SNNAFG   --DdVD---LSKVITD-VASIFKQKnP-p-M--RFQQHS--YQLLG    LLSKNHP  184
Query_LAODE  122   TEHGNNLVRTAVETFHGSGY[1]----FSg-----IYDS-FKCQIELGcIKg---PKVCLAF--FKLVG    VLCKFYP  180
Query_LOTTI  123   KSSALPVLIKVLELSVGSNM   GED-LK---IQKFIEK-FFMELTKA-SK--LTATVKGKI--YYMLG    VICEIYP  185
Query_APLYS  120   KNAAIPVLIKLLELTAGSSL   GEE-LK---VDKMIQK-FFNELTKA-SK--LSATVKANI--YYLLG    VIAEVFP  182
Query_SCALL  121   KNGAVPVLAKLLELTAGTNI   GKD-LN---IPKLIEK-FFMELTKG-SK--VSSTVKENI--FVLLG    VVAEVYP  183
Query_CRASS  121   KNLSVPVLAKVLELSVGSVN   VQD-LG---IDKMINK-FFMELTKSvSK--LPASVKESI--YVILG    LLAEIYP  184
Query_OCTOP  113   KNVALDLLMHVLEASSGYYQ   ISE-LD---IPGLIRK-FLMELANPs-K--TKLTVRAKI--YLLFG    TLAHHYP  175
Query_POMAC  117   KNSALGVLVKVLELTVGSPI   REE-MK---IDNLVEK-LFRELTKT-SK--LVPTVKGNI--YFLLG    VIAEVYP  179
Query_WHIPW  125   RELVLYLLDELLQVRNPVNT[2]ATE------QDGILHC--CLVLLDSlVK--LQSGVKKAL--FILMG    SVARHLP  187
Query_TRICH  123   REVVLYLANECIKYSSVLST[1]SEE-LNf-cYDELLHC--CLALIDSlKK--LQNGVKKGL--FIIIG    SLCQEIP  188
Query_AASTA  144   RSAAGELVTYFIKHHKHVSA   -SAhFDv---ASLPDR-FLNELNSS--Kc--TQTVKGVI--LDVLG    GLFNKYP  205
Query_AEUTE  140   RSEAGYLVKYSVKHYAPPSA   -S--FQl---PRLIGR-LLDELKSS--Kc--TQTVKGVI--LELLG    ALLQKYP  199
Query_PLASM  131   KAQALNVVVNALKYAHKHTT   -SDeIEp---RAYVEK-LFYDIQYS--Ka--AQTVKGQM--LKVIG    YLVEKFS  192
Query_BREMI  130   KAQAVTVVTNVLKYTANQVG   -NDdIDp---RAYVEK-LFNDIKFS--Ka--MQTAKGQM--LELIG    HLVEKFP  191
Query_OLIGA  131   RSDALKVVIHVLKYATSHIS   -SAdVEp---VAYMEK-IVYDLKFG--Ka--SQTAKGTM--LEVVG    HLTAVFP  192
Query_LIGNO  113   RIASLNCLSICFKHAVGSSV   DGNqLN---VGQLCER-LRLEVGQRqTS----EGVTQGL--YRALG    SICKHYP  175
Query_CESTO  25    KEAWLRLFKKVITNSTNDPL[1]-------ltVPEIYAKqLFDQLATKhPQs-----VTSVLc--STLA    TIVALTP  84
Query_OVI2_  36    ---VFNIAIFVLDYHLAYSG[1]---rISv----DMDKRvYMIDEQSIsSKs-ILTFINNYI---HIRN[ 8]MMVALWQ  103
Query_PRIAP  123   KNATFPLLIKVLQLCRGSQT   QAK-LD---VPKLIDR-YFNELTQS-SR--LTSTVKQGV--YHLLG    VIAESYP  185
Query_ROTIF  116   RCATFSIISKILELTKGEYE   ITAkMN---IPKLADD-YFLSLANQ-SK--LSSSIKANI--LSILG    RICWYHP  179
Query_SPONG  116   KNATLNLLINTLELCGKTEY   SHD-LD---IPKIADK-FFQAALLP-TK--HSATIRGGI--YHLLG    VLCEWFP  178
Query_TARDI  127   RACCADLLSQVAIVAARADR   IND-LA---VAEVLSS-LLADARRNvGKaiVGITIKQAL--YGFFG    TMAHKAP  192
Query_TRICH  120   KNATFPPLLQLLSSKLHSSV   VEE-LA---ISKLVDK-YMSACKQP-TK--YSSTVKSGM--YSLLG    TIAEKFP  182
Query_OSTRE  124   KAAGLRVMLALARWSARYLG[1]--EtLWah--GRALRR----DFERT-TKn--SKTVRGLI--LRVIG[14]VLGEKVA  198
Query_NITEN  120   KANTFDPLFLIIRLQLPLVT   -AEnFAa---ADMGRQ-YFGDYRRM-----TTATLKSNT--FVMLG    LLVEYFP  180
Query_CLUBM  123   QAAALAPIISLFEIGVLRAG   -NAtSDglmDWKVAEM-LLAEFRKA--K--TSGSLRANI--LITLG    YSIELFP  187
Query_MOSS_  123   QAAALTPLIVLLELNMARLK   -PAmFA--vDQELSTL-LMLEFRKA--K--TNSTLKANI--IQLLG    DLIEYFR  185
Query_LIVER  123   QAAALSPLLSLLETEITRKH   -ITvAN--lDWTFGEV-LLIEYRKA--K--TSGTLRSSI--LILLG    YLVEYFP  185
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  184   EK--MVR--GSDKLLNLYMTTLKQE   IESK   THEPEYVLIAGCLQGLTHFMVN--FT---QSADEGA-EHA---YS  244
Query_LINGU  185   EG--MVN--YAERLTDLCVRTLQAE   MNKK   TGKPEIPVMTGCLEALSSILTN--FT---QSAEEGK--HS---YD  244
Query_LIMUL  186   EY--MIS--YSEKLCDLYLNELKQQ   MTSK   TRKLERSIVSGCLQGLTGYLHS--FT---VSVEEDP-QKS---YK  246
Query_SPIDE  188   ED--FIP--YQERVLSMFLQELKSQ   MSAK   VKAFNVSTVAGCLLGLKEYLYN--FS---VMCSEDT-ENS---YF  248
Query_SCORP  187   DC--MQL--CSERLLNWYLKDLKFH   TTSN   -QKSESSIIVGCLQGLNEFMYQ--FP---LQDTEDT-EKA---YQ  246
Query_TICK   179   VA--TEP--YRDKLLALYLAELQRV[4]----   ----EKNVVAGCLQGLSGFLFH--FG---TCFEDDP-DNC---FK  235
Query_MONOS  206   SLagVTGA----DGRALLISVLHAN   MTGK[1]----DWPIIGGCFRGLAVTSVQm--------------------RQ  252
Query_ROSET  192   AF-aASSL--STSVLNVYLSELRKQ[3]----   TKAPELPTVEGCFRGLTAYLKT---Np--VSCDPSA-RQAtpgVQ  255
Query_BELCH  186   DL--MVQKdHPDRLLKVYISMLKEQ   MESR   TRKPEMLVISGCLEGLTHFLVN--FT---QSVEEGS-AHS---KD  248
Query_CIONA  181   GH--MTTW--SERLYEIFLRRIKLE[2]-STK   TA--DLKIIEGCLLGLHGLMFT--FT---SAHDEEP-VRC---AE  240
Query_CAECI  193   SD--MLD--NSEKLFRAYLGELKKQ   MTST   TREPKLSVVAGCLKGLTSLMIN--FT---KSMDEDP-KVS---KE  253
Query_HUMAN  187   SE--MIN--NAENLFRAFLGELKTQ   MTSA   VREPKLPVLAGCLKGLSSLLCN--FT---KSMEEDP-QTS---RE  247
Query_MOUSE  187   SE--MIN--HSENLFRAFLGELKTQ   MTST   VREPKFPVLAGCLKGLSSLLCN--FT---KSMEEDP-QTS---KE  247
Query_XENOP  207   CE--MTY--NSEKLFKAFLGELKAQ   MNSS   TRNPKFPVIAGCLKGLSALMIN--FT---KTMEEDP-RTS---KE  267
Query_PLATY  187   SE--MLC--NSEKLFRAYLGELKMQ   MTSS   TREPKLPVVAGCLKGLSALMCN--FT---KSMDEDP-QTS---KE  247
Query_TASDE  188   SV--MVH--NSEKLFRAFLGELKTQ   MTST   TREPKLPVVAGCLKGLSSLMYN--FT---KSMDEDP-QMS---KE  248
Query_CANAR  183   SD--MIN--NSEKLFRAYLRELKTQ   MTSA   TRGPKLPLVAGCLRGLVALMYN--FT---KSVDEDP-QTA---KE  243
Query_OSTRI  154   SD--MIN--NSEKLFRAYLGELKTQ   MTST   TRVPKLPIVAGCLRGLAALMYN--FT---KSEDEDP-QTS---KE  214
Query_TURTL  192   SD--MVD--NSEKLFRAFLGELKTQ   MTSA   TREPKLPVVAGCLKGLTALMYN--FT---KSMDEDP-HTS---KE  252
Query_GECKO  191   SD--MLE--NSEKLFRAYLGELKTQ   MTSS   TREPKFPVVAGCLKGLTALMYN--FT---KSMDEDP-QTS---KE  251
Query_DANIO  202   SE--MVN--NSDKLYKAYLGELKGQ   MTST   TKEPKLPVVAGCLKGIAALMVN--FT---KSVEEDP-AAS---KE  262
Query_SNAKE  191   SE--MID--NSEKLFRAYLLELKVQ   MTSA   TRQTKFPIVAGCLKGLTALMYN--FT---KSMDEDP-RTS---KE  251
Query_STERL  196   SD--MLN--NSEKLYRAYLSELKSQ   MTSA   TREPKLPVVAGCLKGLTALMIN--FT---KSMDEDP-KSS---KE  256
Query_MILII  191   TE--MLH--NSEKLFCAYLNELRVQ   MTSA   TREPKFLVVSGCLRGLTALMIN--FT---KSMEEDP-KAS---KD  251
Query_BAMBO  192   TE--MLD--NSEKLYRAYLNELKIQ   MTSA   TRQPKFLVVSGCLRGLIALMVN--FT---KSIEEDP-RTS---KE  252
Query_ALLIG  152   SD--MLD--NSEKLFRAYLGELKTQ   MTSA   TREPKLPVVAGCLKGLTALLYN--FT---KSMEEDP-RTS---KE  212
Query_GAR__  191   SD--MTD--NSEKLYRAYLGELKSQ   MTST   TREPKLPLVAGCLKGLTALLVN--FT---KSMSEDP-KTS---KE  251
Query_TETRA  180   DQ--LEDD-LKAEIQTVCFKELQVQ[2]--SQ   --KPEVSYVSGMFKSFRYSLKEihLSn----------KEL---ED  235
Query_PARAM  178   ND--LGEQ-------GIYLHAFQSI[3]IKSS   --KPEMVIIEGILKFFKQCLMYhqFNq-----EQIS---L-----  230
Query_SLEMN  178   -Q--LNDF--KVESQDVGFQELVAQ[1]-KSE   --KPEIKAIIGMLKGFTHSLVGe-CTl-----DQDQ---L---DA  232
Query_STENT  182   NE--LQDK--IEELQILMIIKLREE[2]-KGP   --KTEEKTLAGLLKGMASGLK-v-FRy-----DKNQ---L---KD  237
Query_NEMAT  184   EH--MLG--NADRLISIYINLLNSE[1]KSKG   NKSPDKQVIAGCLKGLSSYLVN--FT---QSAAEGS-KFA---RD  245
Query_PISTI  183   EC--VIS--HSDHLVKHYVSVLKLE[6]MTRK   TKKPDMLIIAGCLRGLTSYLTN--FT---QSVSEGS-KFA---KD  249
Query_CORAL  183   EY--MVD--KADRLVQIYVGVLKAE   -MRK   SGKPDMPVVAGCIRGLDSTLVN--FT---QSVNEGS-IYA---KN  242
Query_HYDRA  178   EL--MTEI--NERLLSLFLSFLKQE[2]-STK   KKKPELLVIAGCLNGLSGYLVH--FT---QSCAESS-KYA---SI  239
Query_LATIM  188   SE--MLH--NSEKLFRAFLGELKTQ   MRST   TKEPKLSVVAGCLRGLTALMIN--FT---KSVDEDP-KTS---KE  248
Query_SHRIM  185   QN--LRHQ--SPLILKIFKGQLVKH   TGSK   P---DLNALAGVLRGLALYLHH--F---pEGIDDDP-QLF---TE  242
Query_MCRAB  184   EN--VRCH--SAMILKIYKDHIARH   MGQK   T---DLIALAGVLRGLTLYLFN--F---pEGIDDDP-LLY---KT  241
Query_AMPHI  184   EK--AQSQ--AVFILKVLRDQLNRH   TGAR   P---DLTILSGCVRGLTYYLHH--F---pDGAHDTP-ELF---AD  241
Query_BARNA  185   QQ--TAEL--AERACTALLDELKRQ[2]--SR   SAKLEMTIVSGAIQGLTCLLHS--FSq---SADDNP--VVcg--R  245
Query_COPEP  182   VR--LKPL--ADRILVNYIQFLKGC[3]-KSG[2]----DLTVLEGTLIGLDFLLNGgiLDp------EDV-ENL---ET  241
Query_DAPHN  177   NT--LKEK--ASRMINQILYEMEKE[5]YKDR   -------VVGGCLEALRGFLYS--FQs---IVSQKS---L---EK  233
Query_STARF  185   VL--MSG--YSERLLSVYIATLNQQ   MTSK   TRKPELTVIAGCLTGLGHYLVN--FT---QSVEEGA-KHS---KD  245
Query_STRON  186   QL--MMH--YSDRLFNVLVSTLKQQ   MTTK   TKKPEMTTVAGCLMGLSHLLVN--FT---QSVEENS-KHC---KD  246
Query_SEACU  185   EE--MLP--YSEKLTGIYVSMLRQQ   MTSK   NKKPEMTTIAGCLIGLSHVLVN--FT---QSVEEDG-KHT---LD  245
Query_FUNGI  229   KH--MNDK-I-GQMTHIFLENLQRQ   FKSH   K--PDMQLISGSLSGLRHFIFQ--FSd--HFVK-DT-KNV---QQ  287
Query_DPURP  225   NH--MFDL-N-QQLSSIFMETLGGQ   LKSK   S--PESTLIQSCLKGINSLLVH--FSg--DFVS-NP-KNI---QL  283
Query_HALBU  218   GI--MFDK-S-AQLISIFMETLGSQ   LKSK   T--PETFLIKSSLLGLNSLLLN--FSs--DFISGSS-KNV---QL  277
Query_GLOBO  154   DSftIDKQ-DPDQIKKRLLTTISSC   I-N-   SANPDMELLAGAVLGLSSYLHS--FEa--SAAEST---EI-----  212
Query_PALUS  187   SS--IDAG-QISHLHHILMQCVSTL   F-E-   K-NPDLPQVEGAIRALDDLM----FRa--DFDRDL---GI-----  240
Query_CALIF  179   SD--IPDN-QKKIYTRHCIQTINNF   LQNP   KSFDDYALITSIIYGLNSYLHS--NNd--IIIEDKP--EY---VT  240
Query_GLOMU  194   E---YVDF-EKNKILRVFLANLKEK   ----   ------SLTDAAFTGLYSFLFA--YSd--SIKKEN---QD---AY  243
Query_LOBOS  185   K---VAEK-RQSMVVRWCLTTIQEQ   L-KP   NAKQELTLVAGALIGLDNCLYS--FAd--KVTKDV---PT-----  241
Query_RACEM  181   DA--ITDN-DVQTLKRWTLSALETE   L-KK   SVAEH-QVLSGAL-GAMDCLLS--RDd--VTLDQQ---EL---RQ  238
Query_PVIOL  218   TH--MSEK-N-NQICALLMETLGAQ   LKSK   A--PETYLIVCCLKGLDSLLFH--FSg--DFVA-VP-KNV---QM  276
Query_BOMBU  175   EV--SEA-----QEYHIIFQQLRND[3]QYNP[1]TMAHSLDIFQVYLDAFSNILTR--LT---EKIEKK--YYK----E  234
Query_AEDES  165   EC--VSE----SRQRKILDNMMRVA[3]LLEE   NYP-SLISLAGALQGLTYFLVN--FAp-dDAEEGQE-QQLnlkRR  230
Query_ANT__  181   NI--YSQ-----TFLNITLKRLNNS[3]EYND[4]PKTHKND-FHLYFDALADMLDS--VEs-pDTVRKMT-VFH----N  245
Query_WHITE  179   EV--FNSN--PERFFNAFVCEIEES[3]LDSD   TAG-VSAVLEGCFYGLVEYMKQ--FEi-pPQSKSYG-EKL---YK  243
Query_BEETL  181   QY--LVPN--NESLRIVFLNGLETQ[3]LSLS   NKQ-----LSGLFTGLDRFCAS--FPf-dMNSKEDV-FVV---NK  241
Query_THRIP  171   HD--INFL--SKRLLDALM--M-DL[3]IFHS   RKV-EFTIIEGCLDGLTGFLHS--F---pLDGSTDS-ERC---NK  230
Query_TERMI  183   EC--LRPE--ETRLLLRIY--LSEL[3]TSHA   TAV-KYQIIAGCLNGLTEFLYV--F---pL--KPQE-KEA---RR  241
Query_FLEA_  176   YE--MQDH--VSRVQVLLYSQFETQ[3]LVNT   KLR----IIHGILTAFENMFCT--FAp-dPEIDTEE-CL-----R  235
Query_MIDGE  185   EK--FLD----SQAIELRNQMIATV[3]FKDD   KASISMVVISGAVEGLKSHLIH--FTp-tIDEDPQ------fsEK  246
Query_LAODE  181   DS--MQQE--KCKLFKEAAYMLAKQ[2]-KNE[1]--ERAYLVIEGSFVAINGLLFL--FTt-------DElKKLeelNT  241
Query_LOTTI  186   EY--MTT--DSTKLLDIYKSSLKSQ   MMSK   TNKPEMTIISGCLEGLTSYLVN--FT---QQSGDDS---C---YY  244
Query_APLYS  183   EL--MTP--YSDRLVGLYTSSLRME   MTSK   TKKPELAIISGCLEGLTAFMLN--FT---HSAEEGS-KYS---YE  243
Query_SCALL  184   EH--MTS--YAERFTGLCVTSLKTE   MTAK   TRKAELSIIAGSLAGLTSCLVN--FT---QSADDRS-KQS---YE  244
Query_CRASS  185   EH--MTQ--YAGKLVDIYLRALKAE   MTSK   TKKPELSVISGCLEGLASCLVN--FT---E--DEDS-QNS---YE  243
Query_OCTOP  176   EF--MSD--YAKKLFDIYLSAMVAE   FRSL   -KKPEFLIIAACLEGCNKLLSK--HTp--YDLEEGV-KKY---YE  236
Query_POMAC  180   ED--ITQ--DSDRLVGLYVSALKAE   MTSK   TKKPEFPVIVGCLQGLTAYLVN--FT---QSSDEGS-SHS---YE  240
Query_WHIPW  188   EL--LDKR--GKKYALQFANALSEE[4]MHSG   YRVPAYNVIEGCFVGLNGCMEH--FAnl--FDGRCL-MKI---EN  253
Query_TRICH  189   EK--LITA--DYNYALLFANIMSEE[4]MCSS   IHSPSYSVIEGCFLGLNDLLAH--FSnl--FDGACR-MSC---SE  254
Query_AASTA  206   DA---MAV-HRAAVGRWIDQALDKQ[2]--SN   --APEMQIIHGCFVCLSEIL-----De--ATY--DQ---S-krDT  260
Query_AEUTE  200   QE---MDS-SRENIRSWIEAALDKQ[2]--SN   --SPELQIINGSFTCLAEII-----Ta--NTY--DE---G-frDK  254
Query_PLASM  193   AE---LKK-TVSPLLGWIEHELEKQ[2]--ST   --SPEMMIVNGLLFALARLLKC---Da--GRYKQND---A-lrKK  251
Query_BREMI  192   RD---TSG-NVPPLLSWIEIELEKQ[2]--SN   --GPEMMLISGLLFALARLLEC---Nk--ERYKQDE---A-rkKK  250
Query_OLIGA  193   DE---VKT-VALPLIAWIEDQLEKQ[2]--AA   --SPETLLVAGLLIALARLVVF---Dp--SRYKEDA---T-krQQ  251
Query_LIGNO  176   EE--FQPV-----AVKISVNVVKSL[3]IKSD   HTKVQFAVIASLFEFLSNYFCH--FSl-------EE-PDL---KQ  232
Query_CESTO  85    TT-dLRDR------FLTFLTKILAE[2]-SSS   TKRAEFVIISAVSSSIAKLLRSa-ETe--SKYESDA-------NR  142
Query_OVI2_  104   RG----QY---SSLISLATSAYYQA[6]LFSP[2]--------------GIPALLDQ--ISh------EIE---------  147
Query_PRIAP  186   EL--MVA--YADKLISIYVNELKNQ   MKSK   TKKPEYPVITGCLRGLTCYLNT--FS---QSADEGS-GYA---RD  246
Query_ROTIF  180   EV--ISS--KSNKCLELFLTTLKNE   INAK   FRKSDLSLIAGALDGLNSYLNN--F-----SVSETS-DHL---RA  238
Query_SPONG  179   EH--LVKI--SSRLIDIYMSNLKSE[1]TVKK   TKKPELPIISGCLYGLTAYLNN--FS---QSASEGS-VYA---QD  240
Query_TARDI  193   MV--ANEF--ADSMFGLLRSELQKE[1]---N   AKEPKYGIIAGCLDGFRGLLHN--FA----DSDKGS-VDL---DM  250
Query_TRICH  183   EY--IMP--HAQQLLTIYLNTLKSQ   MTGK   NTKLDMPVIYGCLRGLTSYLYN--FT---QSVEEES-KYA---AD  243
Query_OSTRE  199   AE---APT--AEWLLKASCRVLDAA[2]--DE[1]TLKKEIILIASALDALTSALTAh-------GQENDRlNNAc----  257
Query_NITEN  181   EA--LHDQ--IAPIMNLALTNLMAQ   AKSD   K--PEMLVMAGILDCVTSLLTH--CDe--GEFPGGSqENA---QL  241
Query_CLUBM  188   SL--YDDG-NVSTVLRLCLNTLEEQ   LKQS   GKDVTFVLIAGAVKCLDSLLQW--FDd--TVPAGGTiENS---KR  251
Query_MOSS_  186   ES--FEEE-NVLAILKLCLDTLEEQ   LQKS   ---PQLVLMAGAIKCLDSLLMW--FNd--SLPAGGSvENS---RR  246
Query_LIVER  186   HT--FSEE-NVNTLLKICLDTLEEQ   LRQS   NKDPLFVLVSGVIKCLDSLLSW--FDd--SLPAGGSvEAS---RK  249
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  245   IYKYCVKTMDQ   NIKK---KRYEMPKASLGLMAKHSAQMS--DYLLD-DFKD    LYETLS-YWT-NHNNREMAHIG  309
Query_LINGU  245   IFKCTRMALNL   E-ET---ARYDLTRAALTLLKRHASQFS--QFIYE-NHER    LYSVLL-AWS-QHNNKDVSSLG  308
Query_LIMUL  247   IFCYARSSILS   SVEH---SRFEMPRAGLMLFARHAPQFN--QFLFD-DYEN    MYKNFV-HWC-HQSNKESRAIG  311
Query_SPIDE  249   IFDTARK-LIS   NTEKfssKSSAITRAGLQLLAAHAIQFN--DFIFE-DCID    LYEVLL-QWV-KHHNREMQKLG  315
Query_SCORP  247   LFKFIRI-ILS   KRDV---KTYDMAKASLTLLSSHASLLA--QFLYE-DSMN    IYKIIE-IWT-NHSNREVRSLG  310
Query_TICK   236   LFKCVKRSLTP[1]--DRpg--RYDIPRAALILLSKHAALFG--QLLLD-DHRT    LYQELG-RWS-RHGNRELQSLG  300
Query_MONOS  253   IYEHVANMTKA[2]EQAL---KRFEPIRAALDLFAADAAHLQp---FFTaSQNT[ 7]-LDRLA-----KNSNHDLSVAA  322
Query_ROSET  256   IYNYVKMAINP[1]-VTVkm------------------TRFQ-----------    -----A----------------  277
Query_BELCH  249   IYNFTRRAVSP   QVEM---ARYDVPKAALWLLAKHAAQFN--EYLMN-DYED    MYRKLY-EWS-QHPNRDLKQAG  313
Query_CIONA  241   IYKYAKTSIMQ[1]--ESl--TRYAVPLAAMKIFRDHSSQFS--NLLID-DYSA    FFSAIE-GWF-HHKNWQLKHAA  305
Query_CAECI  254   IFDYSVKAISP   QVEL---KRYAVPLAGLHMLAIHASQFG--SCLMD-NYRS    LYEVMS-KWC-GHTNADLKKAG  318
Query_HUMAN  248   IFNFVLKAIRP   QIDL---KRYAVPSAGLRLFALHASQFS--TCLLD-NYVS    LFEVLL-KWC-AHTNVELKKAA  312
Query_MOUSE  248   IFGFTFKAIRP   QIEM---KRYAVPLAGLRLLTLHASQFT--ACLLD-NYIT    LFEVLS-KWC-SHTNVELKKAA  312
Query_XENOP  268   IFDYTVKAISP   QVEM---KRYAVPSAGLNLLALHASQFS--SYLMD-DYQS    LFEVIS-KWC-GHTNGEMKKLA  332
Query_PLATY  248   IFDFAVKAISP   QVDL---KRYAVPLAGLHLFTVHASQFG--TNLLD-NYVY    LFEVMA-KWC-GHTNGELKKVA  312
Query_TASDE  249   IFDFALKAISP   QVDL---KRYAVPLAGLRLFTLHASQFS--TCLLD-NYIS    LFEIMS-KWC-GHTNGELKKIA  313
Query_CANAR  244   IFDFAMKAINP   QVDQ---KRYAVPLAGLQLFIWHSAQFG--ALLLD-NYVS    LYEVMC-KWC-GHTHPELKKAS  308
Query_OSTRI  215   IFDFAMKAIRP   QIDQ---KRYAVQLAGLQLFGRHAAQFG--TFLLD-SYVS    LFETMC-KWC-GHTNQELKKAG  279
Query_TURTL  253   IFDFAVKAINP   QVDL---KRYAVPLAGLHLFSMHAAQFG--TFLLD-NYIS    LFEVLS-KWC-GHTNAELKKGG  317
Query_GECKO  252   IFDFAVKAINP   QVDL---KRYAVPLAGLHLFSVHAVQFS--TCLLD-NYVS    LFEVMS-KWC-GHQNAELRKVS  316
Query_DANIO  263   IFDYALKAISP   QTDI---KRYAVIFAGLKLFAKHSSQFG--SCLMD-HYIS    IFDVMS-KHC-GHINAELKKSS  327
Query_SNAKE  252   IFDFAVKAISP   KIDL---KRYAVPLAGLHLFSKHAEQFS--TCLLD-NYVS    LFHTMA-KWC-GHQNAELKKAG  316
Query_STERL  257   IFDFALKAISP   QTEF---KRYAVLLAGLRLFAKHAGQFS--SCLMD-NYRS    LYEVMS-KWC-GHTNIEMKKAG  321
Query_MILII  252   IFDYALKAINP   QIDL---KRYAVPLAGLRLFTLHSSQFG--TCLMD-NYRT    VFEKMK-KSC-GHTNSEMKKAS  316
Query_BAMBO  253   IFDCAVKAINL   KIDI---KRYAVPLAGLRLFTLHANQFK--TCLLD-SYQV    LFETMS-KWC-GHTNIEMKKAG  317
Query_ALLIG  213   VFDFSVKAINP   QVDL---KRYAVPLAGLRMFGLHAAQFG--TCLLD-NYVM    LFEVMS-KWC-GHTNVDLKKAG  277
Query_GAR__  252   IFDYALKSISP   QTEF---KRYAVLFAGLKLFARHAEQFS--SCLID-NYRS    IFEVMS-KWC-GHTNAELKRAG  316
Query_TETRA  236   LYKMIKIGITE[2]--DV---KTYSVINNALKVLTAYSHIFKv--QILD-DSKE[ 4]VLGLIS------HNNRDVKENA  300
Query_PARAM  231   LYQALVTFCKP[2]--DL---STFKVPIAACKCMKQNLDIFS--HYLII---KT[ 4]LFDQYI-KIL-QHRNRDLKECV  297
Query_SLEMN  233   LYVCLKTAIQP[1]-LDM---QHKGVMKSAMKLLGTHINLFK--KHVT--RHAI[ 1]LVALTL-KLC-VNENMETRDAA  297
Query_STENT  238   IFIYVKALLTP   -LKV---GNYVVSKASLKVLSENAEMFK--EFIR--EYSV[ 1]VIRTLF-DIN-KLSNHKVKILV  301
Query_NEMAT  246   IYEFTRMAIDV   QGTL---QRYDVPRAGLNLLSRHAGQFQ--EYICT-EHNV[ 3]VYPAIY-RWC-CHQNRETKYAA  313
Query_PISTI  250   IYTYGRMAINP   QVSL---SRYEVPKAGLYLFAKHSAQFK--EYIAK-DYEV    MYGDLY-RWC-KHHNKETSYAA  314
Query_CORAL  243   IYSFLRSAIDS   KINY---ARYDVPRAGLNFLARHASQYK--EYLTK-DYES    VYNTLF-DWC-KRKNKELRVKG  307
Query_HYDRA  240   IMKYVKMCILL[1]--D-y--SRYDVPKAALTLLALHSSQFN--ELLTL-QAEQ    IYDALE-LWI-NHQNKEVKKIA  303
Query_LATIM  249   IFDYAVMAISP   QVDL---KRYAVPLAGLRMFALHANQFN--TCLMD-NYRS    LFETMS-KCC-GHANGELKRAG  313
Query_SHRIM  243   LYKHIHKVLNP   DLNL---PRRDAQRAALELIHTHGAKFS--SQLFE-DCTQ    VFQWFV-KWC-GSRNRDDHKAA  307
Query_MCRAB  242   LFEHVRKALNP   NLDL---HRRDAQRAALELISSHGDKFS--AQLYD-HYDT    LFPWFI-QWC-SSKNRDDYKAA  306
Query_AMPHI  242   LYERLHKIINP   ALNL---PKREAPRVTLELLHQHAARFS--SELLK-QHAL    VYAWLL-RWS-CASNTDDLKAG  306
Query_BARNA  246   IFDYVARVINP[2]--EL---SRREAPRAALALLATHAAQFS--KHMYD-RYVT    LYSHLV-RWA-LSNNRDDARAG  310
Query_COPEP  242   LYEVLKKICRM[3]EAQV---SRRGSMRAGLELFSAHSSLFS--RIILG-DVKF[ 1]-FLELR-KWS-NSKNPDDHKIG  309
Query_DAPHN  234   IHGYLLMLISR[4]EKKAis----FTPKAVLNMLAEHAALFN--KWLLR-QYKM[ 1]-FDSLK-TYV-LLDKPGIREIS  301
Query_STARF  246   IYRFARMAIDP   TVDH---TRYELNKAGLNMFAKHSAQFG--PYLYS-DHQD    LHEAFN-KWS-AHNNPAVKAVG  310
Query_STRON  247   IYQYVTMSVDP[4]SGTK---PRYEIPKAGLNLIRRHADQFN--EYLYR-DHEM    VYNRLK-YWS-HSHNSEMKWAG  315
Query_SEACU  246   IYKFASLAMDP   STQL---LRYGVPEAGLRLFGRHAPQFG--SLLYR-DYQT    IYEKFL-HWS-EHHNKDMKSAG  310
Query_FUNGI  288   LYHYVTMALNP   PPQL---VRYDIPKAALDFISSHGSLFR--QYLTQ-DAEK    MFQRLF-ALC-KHQNKKLRSKA  352
Query_DPURP  284   LYQYIYICMDP   QSST---QRFEIPRAAMKIVARHSPLLA--QYLAE-HSEK    FYQRIE-HWC-NHGNKLNRDIA  348
Query_HALBU  278   LYQYLYICLDP   VSSS---QRFDIPRAAMMVVQRHTSLFK--QYLTE-QSEK    FFARLE-EWC-KHINKKNRDLA  342
Query_GLOBO  213   LST-IRKMIIP   LENQ---TRYAVPRAGLIFIRDHAENFR--NQLCS-DRDT[ 3]IHSCLK-SMC-THGNRDLQN--  277
Query_PALUS  241   IGGYVRRMIEP   LGEL---TQYSIPKAALQIVIDHAAKLQ--DPFF----EN[ 3]IYVDLR-NAS-QHQNIDVSRLG  305
Query_CALIF  241   ITNLVKVILSI[1]VNTL---NRYEQIKACLQFVIDNGHYIS--KVLFK-N---[ 3]FFESLL-TIC-KHQNRDLHKLG  306
Query_GLOMU  244   IFKCIMNSITM   VEND---RTYSIVKTGLQYIIDHSNIFG--EFFLR-E---[ 3]IWQCLW-KLR-VHRNPEVRNAA  308
Query_LOBOS  242   ILQLVKTLVNV   PEDL---SRFATPIAAMDLFAHHIHLFR--SKLID-I---[ 3]MYQRIA-GYC-EHSHSGMSKCG  306
Query_RACEM  239   LYEADKTILDI   PEKL---SRYAAPMAALQLFCNHASLFE--RHLIV-D---[ 3]LYSRLS-SIR-FHRNHDLSKLG  303
Query_PVIOL  277   LYQYIYLCMDP   VSSA---TRYEVPRAAMKIILHHSLLFK--QYLTE-QAEN    FYTRIE-HWC-NHVNKPNRDIA  341
Query_BOMBU  235   LYTWV-KELSN[2]KYSVk----KVSMRSAINLLSRHIDLFG--DFVYC-DYKY[ 1]-YDLLT--NVaQDKNVKCSQCG  299
Query_AEDES  231   IYLIV-KKLSD[2]EAVKe----RQAFRNALILLERHSALFS--PFLYV-DYVH[ 1]-QNILTtKWL-RSSV-EDRKVA  295
Query_ANT__  246   SYEWI-KYFSQ[2]SQQKk----RTIMKSAIDLLCRHMDLFR--DLLYP-DYKY[ 1]-HNFLQ--RLsSETTMY-GNSG  309
Query_WHITE  244   LAKTI--CIFD[2]RNR------KTYQRAALKMMEIHSGLFK--NFLVS-DYKW[ 1]-HTNLR-KWLtNNTNKDNQTAV  306
Query_BEETL  242   VYKYI-ETLAI[2]GTV------KIANKAALTFLANNISLFI--NLMVD-KHKY[ 1]-HDTLL-KWL--ESGTEDQKIG  303
Query_THRIP  231   IVDCLWKTLLL[2]VHR------KTAPRAALNLLANHGDQFR--ASLVL-KFED[ 1]-HKTLC-RWL--PIGKDERRVG  293
Query_TERMI  242   IYDCV-TAVAR[2]KEM------RVAPRCALGLIQAHGGQLA--KFLYD-EYKF[ 1]-HEFLL-KWI--NSAGEDHKAG  303
Query_FLEA_  236   LYNML-KTMSY[2]DV-----KFRPAFRAALSLLTKHGNLVS--EHLYK-DCLW[ 1]-HNTLQ-SWL-KLSW-EDISIA  298
Query_MIDGE  247   LYECM-FQLSD[2]RLPNp--TTRVAFRNNLQLIENYGGLHEipTFFFR-DYKS[ 1]-HKTLK-KWT-NAKLYEDKSVG  315
Query_LAODE  242   IYDSMKLVISV[2]------rQRHTPHRAAFSIIENHSEKISd---LLKrDYVD[ 1]-HEKLI-KWI-DHNNKDDKGIS  305
Query_LOTTI  245   IYKYCKMALKQ   STTL---TRYEVPKAGLHLLSKHASQFN--EFLID-DYED    MYDRLY-FWS-HHHNRDILHVG  309
Query_APLYS  244   IFKYMRMAIDP   NIEH---TRYDMPKAGLKLLTRHAGQFQ--QYLMD-DYQN    LFDKLL-KWS-RHHNREILHAG  308
Query_SCALL  245   IFKYARMAIDP   KVDY---TRYQVPSAGLKLFTKHAAQFS--LYLIE-DYQG    MYDKLR-HWS-HHHNRDMLHLG  309
Query_CRASS  244   TFKFARMAIDP   DINY---TRYDVPRAGLKLFSRHSAQFK--SQLAD-DYQG    MYQKLK-SWS-KHNNRELLHLG  308
Query_OCTOP  237   LFQAARSAAKP   DLDL---SRYEMPKASLILFENHSAFFI--KFLQD-CYED    MYDVFL-KWF-HHKNRELANHG  301
Query_POMAC  241   IFNYARMAIDH   KVEF---SKYDVPKAGLQLFARHAGQFS--QFLVD-DYEN    MYTKLQ-AWS-KHQNRDVLHHG  305
Query_WHIPW  254   MYSMIKMIINP   ESRF---NRLGPVKAALSILTSHPNLFR--GKLFR-DYEL    LNNYFS-HWS-RHYDRELRSLS  318
Query_TRICH  255   LFNLIKKACNP   KNDV---HRLGHLKAALSLFISRVQFFR--HHFLD-EYSL    LNECFS-HWA-KHYDREIKTYS  319
Query_AASTA  261   LFQFIHVTLAT[2]SGNL---SRLAIVKACLGLLGKHMHLFA--ANLVEaDPYQ[ 1]-YLLML--HCcASSNKKIRKPA  328
Query_AEUTE  255   LFKYMHVTLAT[2]SGNL---SRLAIVKSCLLLLGRHIQVFS--ENIVRsDTYQ[ 1]-YSMLL--FCcTSSNKKVRKPA  322
Query_PLASM  252   VYSYLLTIFAT[2]NGNL---SRYQVTNSAEHFLEKHAQVFH--EEIGP-NSLA[ 1]-FSYMK--SCcLSDNKTVKQHA  318
Query_BREMI  251   IYSYLLTVFAT[2]SGNL---SRFQVTNSAETFLAKHAQIFQ--QEIGE-NGEL[ 1]-FSYVQ--YCcESENKTIKEHA  317
Query_OLIGA  252   VYSYLTRILAT[2]EGQL---SRYNVTKASLQFLEEHGQTFI--DELGE-NGYT[ 1]-FTYMR--YCaASESKTIKENA  318
Query_LIGNO  233   VYHLAVKVIAG[1]--NY---QRYEAPKQALLLFANHCPVFS--AFLAD-NYEV    LFEQFF-RWS-THHNRDLMHVG  296
Query_CESTO  143   FFEYALQILLA   --DLdniSRYKPIKSACKLLSASCKQIS------TcDWSR[ 3]VITYLQ-NLT-NHKNSSLASTA  208
Query_OVI2_  148   -VRFSLTFVIP[1]SAALlv----GIVQLLTKLIAKCPQRLA--KYLLL-----[ 3]LITNIR--QCmIHRNPELSRKA  210
Query_PRIAP  247   VYDCVMRCTGQ[1]LSEI---RRYDVAKSALWLLAKHAAQLD--RFLLA-DYKA    VYSCLR-SWS-RHHNRDVQIAG  312
Query_ROTIF  239   VFNFIKLSLLS[1]SEDL---SRYAVPKSALNILIKHCRQFD--ELIYS-DYKE    IFDRIS-QWS-EHKNYDMKKLA  304
Query_SPONG  241   IYKYTRMAIDP[1]-ANL---TRYEVPRAGLSLISRHAAQFR--EYLVK-DSEH    MYESLS-SWS-NHKNKECRVVA  305
Query_TARDI  251   IFSTCRKLIHI[1]--DA---ARYAMIFAALKLLTEHMDKFL--KHLLA-DKLT[ 3]MYGDLK-RWA-VHTNSDARKLG  317
Query_TRICH  244   IYKYARLAIDP   KVEL---PRYDLPKAGLNLIAKHAAQFC--DYISR-EAEA    LYEVIY-RWC-RHRNRDIRAAA  308
Query_OSTRE  258   VVRIVTRAVAY[1]---Vag-QRHDMTRSALELISEHVRTLR-------hEFVK[ 4]VFQILL-NLR-VSKTSDVARLA  322
Query_NITEN  242   LYKCVTDIV-Q   KDDL---TRYAVLRSALKLLSRHASLFR--RRLLK-DIMR[12]LFEYFR-RYS-NHVNHQIRDAA  317
Query_CLUBM  252   AYAYLEQVLYY   REDM---KRYEMIRAGLRLLARHAWRFR--LALFK-NSGD    LLKWLQ-IYC-LHGNASLRSAA  316
Query_MOSS_  247   VFTCIEQVLWP   REEL---KRYEPLKASLRLLSRHAGLFR--MRLIE-SGIR    NMHWLQ-EHC-THSNAKIRDAA  311
Query_LIVER  250   IYIFLLQVVYY   REDI---KRYEVVRAGLRLLSRHAGRFA--LRLVE-DSKD    ILSWLQ-VYC-IHTNAPLRDAG  314
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  310    MLALEDFLVQISS-ALVTKVKEGnR-K-E   GAVFKFFIQKFRDI    MDNNSSGS--KE---IAIAIKGYGYFAAPC   374
Query_LINGU  309    MEALEAFFKQMAE-MLATKAEEG-H-P-D   RAAFAFFIRQFRET    LDNPHAKQ--NE---IASAVKGYGAFAAPC   372
Query_LIMUL  312    YAALESFLKEITD-ILIAKAREG-KSE-G   SSVFKFFITAFQQC    MRDNKSDS--RD---LSVAVRGYGLFAGPC   376
Query_SPIDE  316    KEAVVAVLKVTSE-MMMSQYDER-KDV-C   SNIFKYFMNCFRDI    FLDKSTTS--KE---LQIAVMGYGLFAGPC   380
Query_SCORP  311    YNALDAVFKQISD-MLEQNYED---SC-K   AEIFKFFIEKFQLM    IKCSENIR--R----VSTAIRGFGYFAKVC   372
Query_TICK   301    TAALESFLREVAS-ALVEHADRQ---D-Y   LDIFQFFIKDFQRT    LK-METSN--HG---LAAALKGYGLFALPC   362
Query_MONOS  323    QGALEALFQWTAHFLAQDPARDSeAVT-E[4]--MYQHLVNVFSNM    LAAEVTPPf--Sl-------------SPV   380
Query_ROSET  278    -----------------------------   --------------    -----ASH--RA---VAIAIRGYGYLAAPC   297
Query_BELCH  314    YAALESFLKQISD-MLVARHEKEnNKA-D   LAVFKFFIREFRTI    IDTTEAGT--RE---MSIAVRGYGFFAAPC   379
Query_CIONA  306    HMVIISFLRQVTR--NISDGGTN-NNS-S   IHLFNFFMDKFRAM    LDSSRTSN--KS---LSVAIEGYGLFAGPC   369
Query_CAECI  319    HLALESFLKQIAS-IVTHDAET--HKS-K   L---QFFMDQFYRI    IRKLDSSN--KE---LSIAIRGYGLFAAPC   379
Query_HUMAN  313    LSALESFLKQVSN-MVAKNAEM--HKN-K   L---QYFMEQFYGI    IRNVDSNN--KE---LSIAIRGYGLFAGPC   373
Query_MOUSE  313    HSALESFLRQISF-TVAEDAEL--HKS-R   L---KYFMEQFYGI    IRNTDSNN--KE---LAIAIRGYGLFAGPC   373
Query_XENOP  333    FAALDSFLKQIAH-LVASDAET--HKN-K   L---HFFMEQFYEI    IRKMDSSN--KE---LSIAIRGYGLFAAPC   393
Query_PLATY  313    HSALDSFLKQISL-MVAKDVEL--HRS-K   L---QYFMEQFYEI    IRKIDSTN--KE---LSIAIRGYGLFAAPC   373
Query_TASDE  314    HSALDSFLKQVSL-MVAKDAEA--HTS-T   L---RYFTEQFYGI    IRNVDSSS--KE---LSIAIRGYGMFAGPC   374
Query_CANAR  309    YSALDSFLKQISL-MVAKDAEL--HKS-K   L---RFFMEQFYGT    IRRMDSSS--KE---LSIAIRGYGLFASPC   369
Query_OSTRI  280    HSALDSFLKQISL-MVAKDAEL--HKS-K   L---QFFMDQFYGI    IRRMDSSN--RE---LSIAIHGYGLFAAPC   340
Query_TURTL  318    HSALDSFLRQISL-MVAKDAEM--HKS-K   L---QFFMEKFYDI    IRKMDSSN--KE---LSIAIRGYGLFAAPC   378
Query_GECKO  317    NSALDAFLKQVSE-MVAKDAEM--HKS-K   L---NYFMKQFYSI    IRKMDSSN--KE---LSIAIRGYGLFAAPC   377
Query_DANIO  328    YTALESFLKQVAT-LVAENIEL--HKS-K   L---KFFMQKFCAI    IRTMDSTN--KE---LSIAIRGYGLFAAPC   388
Query_SNAKE  317    HSALDSFLKQVSS-MVAKDVEM--HKS-K   L---RFFMEEFYGI    IRNMDASN--KE---LSIAIRGYGLFAAPC   377
Query_STERL  322    YHALEAFLKQVAL-LVAENAEL--HKS-K   L---EFFMRQFYTV    IRTLDSST--KE---LSIAIRGYGLFAAPC   382
Query_MILII  317    YSALESFLKQVAI-QVAENAHL--HKD-K   L---EYIMQQFYDI    IRTPESSM--KE---LSIAIRGYGLFAAPC   377
Query_BAMBO  318    YAALESFLKEIAL-QVAEDAHL--HKD-K   L---QYFMQQFYSI    IRTMESST--KE---LSIAIRGYGLFAAPC   378
Query_ALLIG  278    HSALDSVLKQISL-MVAKDAEM--HKS-K   L---KFFMEQFYSI    IRKMDSSN--KE---LSIAIRGYGLFAAPC   338
Query_GAR__  317    HSALESFLKQVAV-LVAENAEL--HKS-K   L---KFFMQQFCSI    IKTTDSTN--KE---LSIAIRGYGLFAAPC   377
Query_TETRA  301    SDCLESLMALLAIELDSSYKNHE--LS--   ---FNYIIQKMSFL    INQERQQIy---tasITVIIKSIGIFSKAI   363
Query_PARAM  298    QELLELFIQKISQ---QLDHSINdHKI--   --IFNVILNKFIEI    LNDNKADMf-----lMTVIIRSVGHFAKAI   358
Query_SLEMN  298    NEMLVHIMQEISDGITIDSSI---HDD--   --IFRKIMKEFETI    LNNETNQTn---vl-LISTIKAVGIFSKAI   359
Query_STENT  302    QETTERCLSVLSD--SLQENY---DKY--   --TFDYIIKMLNMQ    INSKSQE------v-LQSTIRYLGVFSSAI   358
Query_NEMAT  314    LDAMDSFLKQVSEYIVLDEAEAVnPKE-D   ------FIQRFRAI    IDSDSSTT--QQ---MSIAIRGYGFFAKPC   374
Query_PISTI  315    YAAMEAFFQQISLYFVGRHGD----KS-D   LAMFMYFIKTFRRI    IDSVLSDT--RQ---MAIAVRGYGYFSKPC   377
Query_CORAL  308    FAALEAFLKQVAECLVTEDFETTvECG-D   I--FKYFIKEFRRI    INSNESLT--RE---ISIAIRGYGYFAKPC   372
Query_HYDRA  304    ITATESFLYEVSC--YISS--VK-NDQ-T   TKILKVFLFRFHKT    LNSPTASC--KE---ISMAIKGYGKFSLAF   365
Query_LATIM  314    YAALENFLKQVAA-MVAENAEL--HKN-K   L---EYFMKQFYGI    IRTLDSSN--KE---LSIAIRGYGLFAAPC   374
Query_SHRIM  308    LPAMDTFIVVIANMLELQAMEDD--KR-G   VKIFKYLLQEYNKM    LETSESTS--K----LSLAVKGYGLLSGAC   371
Query_MCRAB  307    IPALDTFITVLSNILELQSSED---KK-T   VEIFKYLLQEYNKM    LETGKSTN--Q----VSLAVKGYGLLSGAC   369
Query_AMPHI  307    RRAMESFLSVLASELELNPSDV------A   HTAFSHLVKEFNAL    LESDKSEG--NSsycVSLAVRGYGLFAGAC   371
Query_BARNA  311    VGALDAFINQVCEWLTTAPTDDGrPAK--   --VFAFLLARLREL    LDDRAAPPa-----vLSVAVHGYGALAAPC   374
Query_COPEP  310    WKSLSALFRSVREGLSGIDASSK---NiP[1]--VLQFLTDQFQTI[ 1]-RSKYSSH--KD---ISLAIQGYGEFAEPL   373
Query_DAPHN  302    LNAFGKFIEESIKSLKSEADVAK--KS--   -TVLEFFVRSLCEV    LESDRSSL---Qfr--TVGLKALGECSLSV   364
Query_STARF  311    FAAQEAFLQQIAE-ILVERAAKRdA-S-D   MAVFRFFVKRFRDI    IDNPQSDT--RA---LSSAVRGYGYFAAPC   375
Query_STRON  316    YDALQAFFKVVSH-MLVENADKKrS-S-D   VAVLKYFIKSFRST    IDDASSSS--KS---LSTAIRGYGVFAAPC   380
Query_SEACU  311    LSALETFLKLIAE-QLVTNAAQKsN-T-D   AAVFQFFIRQFRSS    IDNTDSGL--KS---WATAIRGYGFFASPC   375
Query_FUNGI  353    FAAYEVFIIQVSFEITSGSRNSS---S-D   EATFRFFIRSFYEM    LESSSSRI--YE---VSTAIRGLGYMADPI   416
Query_DPURP  349    FAAVDSFFQQISKELTSGNRNFE---A-D   QQTFKFFIRKFYSI    FENNDSSR--FE---LSIAIRGCGRFASPV   412
Query_HALBU  343    FNTIDSFLQQIANELTVGNRSIE---S-D   QATFKYFIRKFYSI    FENNNSTP--FD---ISIAIRGCGRFAAPV   406
Query_GLOBO  278    DTAMTSFLQQVAQALTD-RKLSP---S-E   IEIFWYFVKEFINN    LNKPETDQ-yAE---LSLAVRGLGLFATPC   341
Query_PALUS  306    YRAIDQFLKTIAKVLKS-KSAEQ---D-A   KKCFWWFMRQFTDI    LQQHKA-A-yKD---VSLAIRGYGFFAGPC   368
Query_CALIF  307    FQAMDHLLNEIVSYSENILDEI----E-A   KKIFWYFIEFFTNS    LRKEDADT--KD---ISLAIRGYGLFSGPC   369
Query_GLOMU  309    FKAIWAFLKQIAKVLNSATEF-----K-E   FDCVDYFMLIFDDI[11]LNKEEQDQktKD---ISIAVRGVGYFACSV   383
Query_LOBOS  307    YNSLDVFLQEMANVLTS-TEITE---R-E   YNCFLFFTANFTQI[11]Y---------KA---MSVAIRGYGYFAEPC   373
Query_RACEM  304    FRAMDVLIEQTANVLAKLTPGN----T-E   KGVFRFFFLKFFD-[ 1]MNRESEDPdyRR---IAAAVKGFGHFSKVF   368
Query_PVIOL  342    FNTVDVFFSQIAKELISGNRSFE---N-D   QSTFKFFIRKFYSI    FENNTSSR--YE---LSIAIRGCGRFASPV   405
Query_BOMBU  300    QYALRKFYRVIGHILKHKSSEGD--KA--   --IFLYFKNLFEEQ    LKSSHHVN--SNmlc--YIIYGFSQIAAPC   362
Query_AEDES  296    IYALHSFHREIARQLLAREDLEEsTSTaE[7]ITVLNFFMKYFKTV    LMAPGSKP--YEi---RVAIRGFGSMAEAC   370
Query_ANT__  310    RSALKKFYQVIGQILTNQNSEENdYND--   --ILVYFLQYFEKK    FRDS---N--LDlitPRLIAYGFSQMAAPC   373
Query_WHITE  307    IETLKSVCRVIAKEIVDKETDQGrQI---   ---LEYFSKEITNW    S-NTSCIF--QSa---DLSLDLLIIFVEPY   367
Query_BEETL  304    ILVLEAFYLGLGTFLLNDGSCHYkHV---   ---TEYFKEYFIAN    MTQSKVFN--YDk---RLCLQGLSSFSGVY   365
Query_THRIP  294    KQALLTFYYIMANHI---TGQRGpKDT--   ---LDFFCNHFKNQ    F----QMS--MDv---DIAVKGFGLFAKPM   349
Query_TERMI  304    VSALDVFYNQVASILKDNTSETGhHVF--   ---L-YFINFFKNK    MLSSSTSS--YDm---KLVIKAFGLFAAPC   365
Query_FLEA_  299    VPAIEVFHIQIAKVLLIKKELK-------   --VLLFFINFFKTQ    LQTEKIKT--YEl---RIVIRGFGTFSSCC   357
Query_MIDGE  316    IFATNAFHQFIATTIENRNSIEDkQ----   --ILLFFMKYFQET    LESPNSQP--HEi---RIAIRGFGIMAGAC   377
Query_LAODE  306    MRAYNAFINHMAKLLSIEVLEEQeEEAiA[2]----KQMVQFFKAL[ 8]LSGENASN-------VAAAVESIGLFAPIL   377
Query_LOTTI  310    MCAMESFLKEVNM-SYLSFGD--------   -SVF-FFIKQFRQI    MNSSESSD--KE---VSLAIKGYGLLAGPC   366
Query_APLYS  309    AAALESFVKEICD-ALVQRAQEG-Q-T-E   GSVFKFLLIQFRTI    MNAKNSTD--KH---VSLAIKGYGQLAAPC   372
Query_SCALL  310    MSAMEAFLGQVSE-MLVTRAEEG-H-K-E   GAVFKFFIQEFRTL    MNDTSTSA--KE---VSLAIRGYGLLAAPC   373
Query_CRASS  309    TAAMEAFLTQISE-VLVTKAREG-K-K-E   AAMFKFFVTRFKEI    MENSSASA--KE---MSLAIRGYGLLAAPC   372
Query_OCTOP  302    QWAMEAFIYQISQ-AIV---EDE-DAL-Q   KNMFEFFLEKSNVI    IKNKDSTK--KE---IDVAIKCFGLLAGPC   363
Query_POMAC  306    VAAMDAFLRQIAE-AITEKAKRG-I-K-E   GGIFKFFVQVFQTT    MQNSSASS--KE---VSMAVRGYGLFAAPC   369
Query_WHIPW  319    LSAQDLLMHEVSA-ELLRTVEFD-RSN-A   RSIYKYFSDIFDAS    LLKERDIR--RV----SMIAKSYGRLFAVS   382
Query_TRICH  320    LNAEDLLIRELAT-ELITKVDSD-PNH-V   RSVYKYFCNQFDEI    FSNTSGIR--EI----SMISKSYGRLFQIG   383
Query_AASTA  329    FACLDSMFAIIADGFA-ADIPKH--KK--   --QFNRFLKEFLTC    LTDKLSDD--KA----SIALAGLGNFAAIV   388
Query_AEUTE  323    FSCLDGAFDIIAQGLV-CDVAQN--KK--   --HFNKFLKEFLSC    LTEKMSDD--KV----SIALAGLGSFAAII   382
Query_PLASM  319    FECSNSILQVLNSYLAEAKDDMR--KK--   --CLNKVLKEVLPV    ISDFGAST--LA---ITFALQCLGWFASSI   380
Query_BREMI  318    FACSNAVLQVVNTYLSEVHDDSR--KK--   --CLNKILKEVLPD    VSNSTAGT--ST---MSFAVQCIGWLASSI   379
Query_OLIGA  319    LACANNVLAVIGEFLVQTQEESR--KK--   --TLNRVLKEILPV    LNDKSADT--ST---MAFALRCLARIAPAI   380
Query_LIGNO  297    MLAFDAFVKSVSTAATSAASAESaTPA-Q   KKLFLFFLQQFRQI    LDSDGSSW--KT---VSVAVRGFGYFAEAA   363
Query_CESTO  209    RCTYIALLKLVSKSDSISDDEIKlILS-K[3]------FLTDKSAM[ 2]LPTRIHSF---RlkcLSILLHLYAKHSSNH   276
Query_OVI2_  211    TKAYLAVLLRTGQLLLEQISVGAaLTTtS[7]ATHTNLFVAIIQDV    LQSISNPSl--AlskVVISITAFTYVCEVC   288
Query_PRIAP  313    YAALEAVLKQIAS-VLVEGGEETrQ-R-D   TIVFTFFVREFRTI    IE-AGTDT--RD---VAIAVRGYGSFATPC   376
Query_ROTIF  305    YLALDSYYKSLAE-MIRVKSNTEtDKC-R   L-IFKYFIQKFYKN    LTDGNKDL--KE---MVISIKGYGAFAGPC   369
Query_SPONG  306    VNAIGSFFREVST--QISDPSIS-REL-G   ISNYKFFMKCFRSL    MNDPDINH--RL---LSIAVKGYGHFAKPC   369
Query_TARDI  318    FAALSVYLKQVSRSVIMHLEQGNnEGV-C   TSAFSFFMNEFNEK    IKHAESP---RE---MTVAVRGFGAFVAPS   383
Query_TRICH  309    FHALEEIMKQIAEYLVNENDAA------N   INIFKFFLKKFRSI    IDSGSSDA--KE---LAIAIRGYGYFAAPC   370
Query_OSTRE  323    GEAYEKFVLAMSDAFDSDASPSReART-K   --LLEKIMESIIHI    LDDDSASH--RDr---TACIRALGKFAVPA   387
Query_NITEN  318    NHALEETLTQVSKAMLESDPATE---PlV   RHTYTELMKAVFGI    LNDPASSP--RD---VSLAIRAQGKLAGAI   382
Query_CLUBM  317    LPAYDQFVSQISSCLISDDRRIE---D-P   DRTYKALMKHMLDT    LESRQDNP--KL---VSLAIRGIGKLAGPI   380
Query_MOSS_  312    LPAYDQFITHVSSALV-SDSRIR---N-A   QALYTAFVKDALDV    LNAQDADP--KR---ISLAIRALGKLAACI   374
Query_LIVER  315    ILAFNEFISQVSNCLIHEESRLQ---N-V   QERYKHFMTTFFDI    LDAPSVDK--ML---LSLAIRGVGKLSPLI   378
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  375   -KLY-LTPKDVLFMFN-EMLHRSE   QLFL    SGA-EQNDDRI    IGLHS-FLGALASIVEELDQ    FSETFL  434
Query_LINGU  373   -KFH-MTAEDVKYMFN-EMMQRSE   QIFF    SKN-ELKVDNV    YNLPS-FIESLARIVEQMEE    VSDTFL  432
Query_LIMUL  377   -KLY-LGEEDVKFMQN-DMLQRSE   HFYL    SEE-ILLEEKI    QTLPS-YLEALASIISQLNE    VSETHL  436
Query_SPIDE  381   -KLC-LSEVQVKKMLQ-NVIQKSE   HEFF    SGG-KDIDDKI    VGLNY-YIEALSSIVSVTEC    MTDNIL  440
Query_SCORP  373   -KLY-LTEKDVKYMLN-DVIKKSE   QHCL    NEE-KILDERI    QILPS-FLDSLSSIVLQLQE    ISESSL  432
Query_TICK   363   -KRY-LKEADVLYMLN-DVLLRSE   HVFQ[ 3]SHA--VLEDSL    PLLSS-FQEATASILRCTGQ    VPASQL  424
Query_MONOS  381   ---L------IPVVFF--------[2]QALL[ 9]----------I    --LPT---------------    ------  406
Query_ROSET  298   -KLLc-GVDDLRLMSD-DLARRCQ   QLYF[ 2]--GdDEQEESA[ 1]-HLAS-YIDAIARLIQQLEE[ 2]--DVFL  358
Query_BELCH  380   -KMF-LSESDVKFMFS-EMMQRSE   QLYF    GTA-DSMEDKV    HHLPS-FLEALASIVQHLDQ    LSDMFV  439
Query_CIONA  370   -KRF-LPKSEVRYMFT-QMMNQCE   NIFL[ 1]-EKvDFDDNRT    HQLPT-FVQSLASITKQIES[10]LPESYR  440
Query_CAECI  380   -KV--INAKDVDFMYI-ELIQRCK   QMYL    TEA-ETEDDNM    YQLPH-FLQSIASVIVHMDS    IPEVYT  438
Query_HUMAN  374   -KV--INAKDVDFMYV-ELIQRCK   QMFL    TQT-DTGDDRV    YQMPS-FLQSVASVLLYLDT    VPEVYT  432
Query_MOUSE  374   -KV--INAKDVDFMYV-ELIQRCK   QMFL    THA-DASEDHV    YQMPS-FLQSIASVLLYLDT    VPEVYT  432
Query_XENOP  394   -KA--VNAKNVDLMYI-ELIQRCK   QMYL    TEA-DTEEDNV    YQLPN-FLQSVASVILHMDS    IPEVYT  452
Query_PLATY  374   -KV--INPNDVDIMYI-ELIQRCK   QMFL    TEA-ETIDDHL    FQLPS-FLQSIASVLLHLDK    VPEVYT  432
Query_TASDE  375   -KS--INAKDVDFMYI-ELIQRCK   QMFL    TET-DTIDDHM    YQLPS-FLQSIASILLHLDT    VPEVYT  433
Query_CANAR  370   -KA--IHPEDVDNMYV-ELLQRCK   QMFL    TEA-ETIDDHL    YQLPS-FLQSIASVIFHIDK    VPEVYT  428
Query_OSTRI  341   -KA--IRSKDVDYMYV-ELLQRCK   QMYL    TEA-ETVDDHI    YQLPS-FLQSIASVIFHLDT    IPEVYT  399
Query_TURTL  379   -KA--INSKDVDFMYV-ELLQRCK   QMFL    TEA-ETADEHV    YQLPS-FLQSIASVIFHIDT    IPEVYT  437
Query_GECKO  378   -KA--INSKDVDFMYV-ELLQRCK   QMYL    TEA-ETMDDHV    YQLPS-FLQSIASVIFHLDT    IPEVYS  436
Query_DANIO  389   -KV--VCPQDVDLMYT-ELIQRCK   QMYL    TES-DRDDDNV    YQLPS-FLDSIASVLVHLDR    IPEVYT  447
Query_SNAKE  378   -KV--INPKDVDFMYV-ELLQRCK   QMYL    TEA-ETIDDHV    YQLPS-FLQSIASVIFHLDT    IPEIYT  436
Query_STERL  383   -KA--VCLGDVDFMYT-ELIQRCK   QMYL    TET-DTEDDNV    YQLPS-FLDSIASVIFHLDR    IPEVYT  441
Query_MILII  378   -KA--VNSNDVDFMYI-ELIQRCK   QMYL    TEA-DTEDDNV    YQLPS-FLDSIASVIFHLDK    IPGVYT  436
Query_BAMBO  379   -KA--VDSHHVDFMYT-ELIQRCK   QMYL    TEA-DTEDDNI    YQLPS-FLDSIASVIFHMDK    IPDIYT  437
Query_ALLIG  339   -KA--INSKDVDIMYI-ELLQRCK   QMYL    TEA-ETADDNV    YQLPS-CLQSIASVIFHLDT    LPEVYT  397
Query_GAR__  378   -KA--VCPQDVDFMYT-ELIQRCK   QMYL    TET-DTEDDNV    YQLPS-FLDSIASVLFHLDR    IPEVYT  436
Query_TETRA  364   aKHL--GQDQLQLFFK-QMINLSE[1]RLLK[10]VDFqPVEDETM[13]KQVIS-FITSFANIISQLDK    LYEDES  448
Query_PARAM  359   aNLR--GEKQLGDFFD-KLNDLGF[4]QVYE[ 8]-NMkRGAQFKV[ 5]KQLNS-FLISYANIIKNLSK    ISDKQI  435
Query_SLEMN  360   -KTK-SGDDALREYLS-KLTTLSQ[4]KEFE[ 1]-QEdPKNDYKN[ 8]KQLLS-FIESYSFIIQNLDS    EPSDEV  432
Query_STENT  359   -SKI-QGEAYLKGILE-TIMELGT[4]SE--    ------DEHTD[ 8]KKLSS-LLLAYSDIARNMNS    IPESAL  423
Query_NEMAT  375   -KLF-MTAEDVKFMFT-EMMQRSE   QIYL    RQS-EMQEDRV    QQLPN-FLEALSFIIEELDD    ISDGFL  434
Query_PISTI  378   -KLF-MQADDVRFMFS-EIMQRSE   QFYL    SQS-DLSEDRA    QHLPT-FLEALSSIIEELDK    ISDSVL  437
Query_CORAL  373   -KVF-LKDEDIKYMFT-SIIQMSE   QMYF    NDS-ELKEEKL    QNLPS-FLEALGFIVEEIDE    IPPAFL  432
Query_HYDRA  366   -SNV-MTENEKISVFD-EILKRSE   MILL    ------NNELF    HQLPT-YLESLSCILKIFTS[ 9]------  423
Query_LATIM  375   -KA--VCPEDVDLMYT-ELIQRCK   QMYL    TET-DTEDDNV    YQLPS-FLQSIASVLFHIDR    IPEVYT  433
Query_SHRIM  372   -RQF-LSPVEVHTMFS-QVLTASQ   HAFY    -QSgDMLENKL    TNLPS-YIESLAIIISNMDV    VNSGIL  431
Query_MCRAB  370   -RIL-LSPMEVHTMFL-QVLSASH   HAFY    -QSsEMLENKL    SNLPS-YIEALADIICNMST    INVGIV  429
Query_AMPHI  372   -KQL-LPSQQLRQLFS-TVLTASQ   HAFN    -QAsELFDTKL    ANLPN-FIQSLASIIKHLPD    ADSASV  431
Query_BARNA  375   gRLHp--PQELALLVT-WLADRAE   RLYL[ 2]-ESaPPPAELL[ 1]-QLPG-HVSSLGCCLAALPT    LPAAQL  436
Query_COPEP  374   -ARF-LSEEKYLNLCL-EVLGQTE   QFHI    -ESsQEEQEKW[ 2]-QFPS-HLKSIAVMISSLTT    -TSINL  433
Query_DAPHN  365   ------TGDNIWHIFM-YLFNKTA[1]-ICI    ---qEDTSKSL[ 2]--LPV-CLESLGLVLTHCTT[ 1]LGTERT  419
Query_STARF  376   -KLF-MTDDDVKFMFT-EMIQRSE   QLYL    QQG-EVSDDKL    NQLPN-FLEALASIMRQLDV    LSDSFQ  435
Query_STRON  381   -KAF-MKEADVRFMFN-EIILRSE   QQFL    VQG----EERL    SQLPT-YLEALASIVGQLDT    VSDTFL  437
Query_SEACU  376   -KAF-LGENDVKFMFS-EMIQRSE   QLFL    QDS-NIDEDRF    MQLPH-FLEALANIISELGE    VPESFL  435
Query_FUNGI  417   -CQF-MGQNELKKMLS-KLVTMCQ   RFYS    GGNe-SMEEQS    AHLAS-FISAFASII-----    LQETYL  471
Query_DPURP  413   -KSF-MGEWELKSLLS-SLFKFSE   KLMV    VKIe-NIDEIT    LHLSS-FINAFASILYELSE    LEFWYL  472
Query_HALBU  407   -KAF-IGEWELKSLLG-SLFKFSE   KLMV    VKLe-NIDEIV    MHLSS-FINAFSNILFEIKD    IEYWYM  466
Query_GLOBO  342   -KVL-LAEKELKELQG-ILVRK--   -L-S    SLSsGSNDAKY    AHISA-FMDAFKMIVHESVK    VDTELL  398
Query_PALUS  369   -KQL-LDPGQLNLMLE-HLVAKNA   FL-N    SVGtGNREDIL    SHIPS-FLEAYTFIADELDV    IDTELL  428
Query_CALIF  370   -KKY-FNESEFNIIIN-ELLKKAD   SIFS    NDNl--NEEMF    LHFSS-FLQAYNSILKVSDI    VDNSLL  428
Query_GLOMU  384   -KRL-DKEERYKPIFQ-NLLNPSN   WAIS    NLNeIEITHAL    SQIPL-YIEAYCFIGREYDE    IFQQLM  444
Query_LOBOS  374   -KHI-APDQLKNLLVQ-LLRKSSF   LVSS    NLNe-GVDGST    SHLAA-FISAYTNVARVYTE    VPELLM  433
Query_RACEM  369   -KKF-MQQSEVDGLVQ-ELVKKSA   WIFT    E-GtLEHMRLL    NYIPS-FIASYNQIIQHVDN    VSSQFV  428
Query_PVIOL  406   -KSF-MGEWELKSLLS-SLFKFSE   KLMI    VKId-NIDEIT    LHLSS-FINAFASILYELNE    LEFWYL  465
Query_BOMBU  363   -KIY-LTNNDVKNMFS-LIANCAM   PLCS    --RdDIEKIHL    ESICN-YQEALSEIILHTSD    LSIDQI  421
Query_AEDES  371   -SCL-MTEEYMNELLL-LVMQRTE   FVYL    --VeE--KSNE[ 2]EHLPD-FVQALSDIMSHVRE    LTGVQI  429
Query_ANT__  374   -KVH-MSDFKIRSMYS-ITTSYAL   PLCS    --SeYSHSTHI    ESICC-YQEALSEILCHMTD    VTMEQV  432
Query_WHITE  368   -KK--YSGDDLSNLFS-DIVKRAE   CLCA    --EkNESDTAA[ 4]ARIGE-FLNCLATLVCNDKD    ITHQQL  429
Query_BEETL  366   -QLQ-LEQEEIKKLFA-QIIQNFE   QTYI    --LnLDPSTDV[ 1]EFLPF-YLQTVANFL-QFKT    VTNNEL  424
Query_THRIP  350   -KL--LSPEEIPVIFN-IIAQKSH   EESF    --KsNH-----    HRLAD-YLDSLSMFSVEAGV    VHSLQL  402
Query_TERMI  366   -RVY-MTPQDVLYMFT-VVMQRAE   QIYF    --KeGDISNAD[ 1]QNLPD-YIESLSSIMRQLSA    VTNEQL  425
Query_FLEA_  358   -SAL-LSQDDIIPLYK-IIIQQTE   VMYT    --SeTQHKS--[ 1]QNLPE-YLESLSIIISHMPS    INNVHL  415
Query_MIDGE  378   -QVL-LESKYLSERFD-LVMQRTQ   YSYY    --TnDRLKRRD[ 2]EHLPH-YVESLSKIMTHLNE    ISGLQL  438
Query_LAODE  378   ------DHSEIEKLFE-DLVQKID   -LSL    ----SDVDSNI    FYLSE-YLGALARVGAVLHE[ 1]LKQQQL  430
Query_LOTTI  367   -RRF-LNVTDVVFMLS-EMLNKCE   QQFL    LKVgEN-QDKM    YGLHS-YLDAIANIIKEIDE    LSETYA  426
Query_APLYS  373   -RMF-LSEADTQFMLN-EMISKCE   HQFF    SQV-EELDEKL    ISIPS-YLTALSHIIRQVDS    VSAPYA  432
Query_SCALL  374   -KTF-LKQADVQFMFS-EMMSKSE   QQFL    SQD-ENLEERL    FNLPS-YLEALGNIIKQLDE    ISETYA  433
Query_CRASS  373   -REF-LQQKDVQFMFS-EMITKSE   QLYL    SSDkEDIEDKL    YSLPS-YLEALGNIIRQLDH    VPETYA  433
Query_OCTOP  364   -KKF-HTQDDVNFMFR-EMLEHSK   HLFE    KES-ENLHDQL    YSLRG-YLEALSSILKEIDC    ISDESL  423
Query_POMAC  370   -SIF-LEPADVTFMFS-EMITKCE   QEFL    RTN-EPVHEKL    SSLPD-YLHALASIIKVVTK    VSEVYA  429
Query_WHIPW  383   -HVL-FENGEVNRILR-QILFRFH   QFLC    HDE-GPVEQVI    HYVTS-PLEAVARALPLLSN    VPFDIW  442
Query_TRICH  384   -KAL-YGEEEVCKVLR-KLLQRFE   YLST    TSS-DAQQQNV    YHITN-ILEAASRGLLCLSK    IADDVW  443
Query_AASTA  389   -PMF-MGADALPKIHA-RLIKYGE[2]-VAI    -------REG-[ 4]WMLLCrYTTCYGRFVQKMQC[ 5]VQN---  449
Query_AEUTE  383   -PVF-MGDGALSKIHT-RLVKYGE[2]-MAM    -------REK-[ 4]WMLLCrFTTCYGHFVRNMQ-[ 5]VEN---  442
Query_PLASM  381   -FTY-LGAKGYIKIED-KLRVYGE[2]-LAL    -------DAKT[ 4]WSLICqYVQCIGQFVQHRS-[ 3]LDEGYI  442
Query_BREMI  380   -YLY-LGAKGYTRIEN-KLKLYGE[2]-LAL    -------DTKA[ 4]WPLFSqYLQCIGQFVQQRH-[ 3]LGESFV  441
Query_OLIGA  381   -RMY-LGEKAFLKVED-RLRKFGE[2]-LVL    -------DDKA[ 4]WSVFSsYVECVGCFVKQER-[ 3]IDDTNI  442
Query_LIGNO  364   -KKL-LAPPDMRQVYA-DVVQRAE   DLFT    AES-DSSDERL    SNLPS-FVESLSNLALHLDE    LTGVFL  423
Query_CESTO  277   ------DSTQILALSVqELTRLSG   ----    ----EISNNPV[ 4]YSLID-LLGALLNLLTKLPE[ 5]MRESLL  335
Query_OVI2_  289   vRIT--SANEFLNIID-KLLDRIE[3]---L[ 2]DSSsEVEPFDL[ 7]SSLSTvTWSKISTMQSSLDQ[ 5]LYNRLN  364
Query_PRIAP  377   -KLF-LSEADVQYMFT-DMVVRSR   QIFH    SGE-EMIDSRM    VELPP-LLEALASIVTELDQ    VSESFL  436
Query_ROTIF  370   -REL-MDTKDMKLMLN-IIIDICD   RNFL[ 4]AQHnEIYDEKI    FQLSS-FIESLSCVCSQIDS    FSEGSL  434
Query_SPONG  370   -RLY-GSMEDLHFMFG-EMVQRSE   QLFL[ 6]TSSsSVAFEER    FLLPT-FIEALANISKEIQN    LPDVCL  436
Query_TARDI  384   -RIL-STAAQNEQMLS-EMLQKSE   QTFF[ 3]TEV--IGESRL    EHLPT-FLDTLSAMVQDCGS[ 1]VPDAFL  446
Query_TRICH  371   -KTL-LKKDDVRFMTM-DLLQRSE   QLFF    --I-DVNVDKI    QYLPS-FLEAFASILNILDE    VNDTFT  428
Query_OSTRE  388   -LELgVRGADMNAYLG-----RLE[2]SMYT[ 9]QGRfESIERQV    ----S-LLSAYSDLLRKREI[ 1]VDEALI  453
Query_NITEN  383   -LKF-SGEEELRKALN-RLLPFGA   RSTV    STGy-GDDN-L    GHSVT-LVGAFADMMVYLPS    VEDSLI  441
Query_CLUBM  381   -VKY-SGKSDLEKAFQ-RLVPFVG   -SFS    RGGd-SHEE-L    GYAVT-VVGAFANIVVYIDR    VEDFMV  438
Query_MOSS_  375   -VKF-SGKESLKDAFH-RLEPFG-   -SLS    RLGd-GREGSL    GYAVT-LVGAFADMVIHLDD    VEDLMV  432
Query_LIVER  379   -VKY-SGKETLEKVFH-RLSSLGG   TTTT    RGGe-VYENYI    GHSVI-LIGAFADIVNHFDR    VEDFMV  438
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT   435    F-SLERL-VVLQMQVLPRLHVK-LHNQCYIA---LIRLFL    GLALKGS-IF-KGFLERIIYQ--SLIRSCSHPIVLE  500
Query_LINGU   433    V-SLEKL-VVCLTEHFPSLYKK-RKLDAVQS---ILGIFH    ALISKGP-TL-RIFLGRVVYQ--GLIRTCSHPVATD  498
Query_LIMUL   437    E-RVVKL-VVMQIEHYPQLADV-HHFLSYRS---IIRTLL    AFAPKKA-VF-KRFLGNIVYQ--GLIRTCSHPVVAE  502
Query_SPIDE   441    I-SLEKL-VVFQIDKFPKVPTA-YNFIVSKA---IIRMFL    AVLPKRD-MF-IGFLSQIVYQ--GIVKTCSHPVVVE  506
Query_SCORP   433    N-YLEKL-AVFQIQNFPYLPSP-YHFLCYNS---LIKLFL    AVHSKGP-QF-KGFLNKIIYQ--GIIRTCSYPVRFE  498
Query_TICK    425    A-LLEKL-TVLQIENFPLVAQN-LQMRYCKS---LLFVLL    AVY---P-IS-RSSLSQIVYQ--GLVRTCSHPIAAQ  487
Query_MONOS   407    P-----L-------AMPRVSLT-----ASRGt--------    -----GApSF---------Q--GVVLACVYGVQAS  440
Query_ROSET   359    P-ILERL-IVLLLEAYPSMYRT-QRASVHLS---LTRCFV    SLYHKGS-TL-QTFFAKVANHv--FLSACSQSIRDT  424
Query_BELCH   440    T-SLERL-VVVLFENFPRLGKK-MHLICQKA---ILKVLL    ALVPKGA-TL-RTFLSRIVYQ--ALIRTCSHPNVIE  505
Query_CIONA   441    T-SLQRL-SVLLIERYPKLHRK-YSFLCHNA---LISVAA    AFLEKQE-IL-KTLLESIVYQ--GLVRCCAYNVVVG  506
Query_CAECI   439    P-VLERL-LVVQIDSFSQYSVK-MQSVCCRS---IIKVFL    ALAGKGP-VL-WNFISTVVHQ--GLIRVCSKPVTLG  504
Query_HUMAN   433    P-VLEHL-VVMQIDSFPQYSPK-MQLVCCRA---IVKVFL    ALAAKGP-VL-RNCISTVVHQ--GLIRICSKPVVLP  498
Query_MOUSE   433    P-VLEHL-MVVQIDSFPQYSPK-MQLVCCKA---IIKLFL    ALSEKGP-VH-WNCISAVVHQ--GLIRICSKPVVLQ  498
Query_XENOP   453    P-ILERL-LVVQIDSFPQYSLK-MQSSCCKA---VLKVFL    SLAGKGP-VL-WSLISTVVHQ--GLIRVCSKPVVLA  518
Query_PLATY   433    P-VLERL-LVVQIDSFPQYSIK-MQSACCKA---IVKIFL    ALAGKGP-VL-WNFISTVVHQ--GLIRVCSKPVPFL  498
Query_TASDE   434    P-VLERL-IVVQIDNFPQYSIK-MQTICCRA---IIKVFL    ALAEKGP-IL-WNFISTVVHQ--GLIRVCSKPVILQ  499
Query_CANAR   429    P-VLEHL-VVLQINSFPQYTEK-MQPVCCRS---IIKVFL    SLSGKGP-VL-WSFISTVVHQ--GLIRIFSKPITFS  494
Query_OSTRI   400    P-VLERL-VVVQIDTFPQYSEK-MQSVCCRS---IVKVFL    ALTGKGP-IL-WSFISTVVHQ--GLIRIFSKPITFS  465
Query_TURTL   438    P-VLERL-IVVQIDSFPQYSVK-MQSACCKS---IVKVFL    ALAGKGP-VL-WSFISTVVHQ--GLVRVCSKPITFS  503
Query_GECKO   437    P-VLERL-IVVQIDSFPQYSIK-MQSTCHKS---IVKVFL    ALASKGP-VL-WSIMSTVVHQ--GLIRVCSKPLAFL  502
Query_DANIO   448    P-VLERL-LVVQMDSFPQYSQR-MQHATCRS---IVKVFV    AMAVRGP-VL-WSFTSSVVHQ--GLIRVCSKPVLQS  513
Query_SNAKE   437    P-VLERL-MIVQIDSFSQYSAR-MQTACCRS---IVKVFL    ALAAKGP-VL-WSFMSTVVHQ--GLIRICSKPLVFS  502
Query_STERL   442    P-VLERL-LVVQIDSFPQYSQR-MQPVCCRS---ILKVFI    ALAGKGP-VL-WSFISTVVHQ--GLIRVCSKPVVLT  507
Query_MILII   437    P-VLERL-LVVQIDSFPLYSAR-MQPICCRS---IVKVFV    ALAGKGP-VL-WNFISTVVHQ--GLIRVCSKPVLLP  502
Query_BAMBO   438    S-VLERL-LVVQIDSFPLYSVR-MQPVCCHS---IIKLFI    ALAGKGP-VL-WSFISTVVHQ--GLIRVCSKPVLLP  503
Query_ALLIG   398    P-VLEHL-IVMQIDSFPRYSEK-MQPVCCRS---IVKVFL    ALAGKGP-VL-WNFISTVVHQ--GLIRVCSKPIVFS  463
Query_GAR__   437    P-VLERL-LVVQIDSFPQYSQR-MQPACCRA---IVKVFV    AMAGKGP-AL-WSFISTVVHQ--GVIRVCSKPVLLT  502
Query_TETRA   449    Q-SLLRL-CMVGFKQFSSFYEK-YKEPLNEG---IVQLMR    EMKKHQSiYI--IWLENLVQN--GLVQSIKLPDNSF  514
Query_PARAM   436    Q-HLFAL-CSLCIKQHYNYFEN-YRPYLYDG---MANIII    SIFGHQTiGI--QFVGSLLSK--GVIEGLKLGQNQL  501
Query_SLEMN   433    TqHFLKV-CEIGVRHHRKLYQK-YRDRFYEA---LASFTK    ALSKHQKaFA--NWIQKFVDL--SLIQTLKIPDAVL  499
Query_STENT   424    L-HFSDV-AFEIFKRNIKIHKR-YRQRMYNS---IAQLFA    SLHLHHEkFI--PWFRVFIAK--SMKQLLEIPEEFI  489
Query_NEMAT   435    V-SLERL-IIVLFESFPRLSKS-AKYLCHKA---LVKMLI    NISAKGS-VL-KNFLSEVVYQ--GLIRTCSHPVLLD  500
Query_PISTI   438    A-CMERI-VRVLFDKFAILNKG-GKFLCCKA---LIRVLF    HLSSKGA-IL-RNFLSEIVYQ--GLIRTCSHPVILK  503
Query_CORAL   433    S-FLERL-VVVLIEKFPELGRG-PQFICLKA---MLKLFL    ALTTRGA-SL-KSFLSEIVYQ--GLIRTCSHPVLLD  498
Query_HYDRA   424    ----EKI-IIFIFENICNLGKT-GLALYMKG---IATFID    SVSSTQS-MI-KTSLSRIVYQ--ILIRTCSHSTL--  484
Query_LATIM   434    P-VLERL-LVVQIDSFAQYSQK-MQPTCCKS---IVKVFL    ALAGKGP-VL-WSFISTVVHQ--GLIRVCSKPVPLS  499
Query_SHRIM   432    E-NVQQL-SSALVEHFPLVSSF-KRFLAYRA---LFRLFF    TVYNKPG-YF-HEFLHTFVYK--AIVLSCSHA-PEV  496
Query_MCRAB   430    E-NMLQL-ALALVENFPSVSVY-KTFLAYRA---LFKLFY    AIHGKAE-YF-QDFIHSFVYK--SVIHTCSHA-PEI  494
Query_AMPHI   432    S-AVLEL-SVALTRHYPQLAAP-ERGVAHAA---LLTLLA    TSTAAPG--L-SCIRSSWVHQ--SILLTCSHSVPGD  496
Query_BARNA   437    A-TGERL-LLLLVEHYPRVPPT-HRFLADGG---VARCLL    ALLGRPE--LaAPCLRRLVYQm--LLRCCSHALAVE  502
Query_COPEP   434    SfSLEVL-SFNLVKKFPVLPHQ-YHGFAIQA---LLQVFQ[ 1]-------tAVsDTMLDNVIYQt--VIESCSHPAVGE  496
Query_DAPHN   420    Q-QLQRM-FVHLFKRYLDIDEG---NRCGRS----AAVFL[ 5]AIENQID--F-STFVRDIVIQ--SLIRSTRPELDSD  486
Query_STARF   436    A-ALERL-VVVQVSSYPRVPNK-YQYLSHQA---ILLVLI    ALSSKGA-TL-KNFLSRVVYQ--ALVQTCSHPVVLD  501
Query_STRON   438    T-SLEKL-VVLVFGSYPHQARS-SQWTSHSA---ILQILI    ALSAKGT-SL-QGFLSRVVYQ--WILRTCSFPVLLE  503
Query_SEACU   436    N-SLEGL-TIVAFRSYPSIPVR-YRFLSHRA---LFRIFL    VLHPKGT-CL-QQFLSKVVFH--ALAETCSHPVLLE  501
Query_FUNGI   472    T-HFEAI-LGTVFRIYPDLYF---RDGYCLA---ISRLII    ALCSKGT-SL-QSLLDRVIYQ--ALLITCS-----V  530
Query_DPURP   473    D-HIEQI-LETYFIIFPHLFVK-SRDRYFKA---VNRLIS    SLFYHGE-YL-KIILSRFVKK--GLLITFSKPNPSI  538
Query_HALBU   467    E-HLEQA-VATYFIIFPYLFTV-QRERYYSA---VNRLLS    SLYHRPE-FL-KTLLGRIVRQ--GLLITFSKPHEAL  532
Query_GLOBO   399    N-AIEEL-LTSMLLNFPKVREY-YRKPLVSA---FVELMW    ELNIHGL-LW-S-FWGGIAKR--LLIISCVDKEE--  461
Query_PALUS   429    A-AVEAI-AKTMLTKFSSMGAA-YRGLSVSA---FRNLLG    VLNKKGP-IG-TALWSKIAYE--TIVITCTSTSS--  492
Query_CALIF   429    R-GIELI-ITQLMISFQLLNDK-LKFINAKS---IQHILW    ELNNKSI--M-DTFFNKIASD--ILVLTCLNKEQKE  493
Query_GLOMU   445    E-TIEHE-VGLMMIYFPRMVYD-LRLRCVNA---FEHLLW    MLYLKGRgTM-EKFVSKSFYQ--TLILTCSDAIQ--  509
Query_LOBOS   434    S-ALNQM-VNAAIANFIRMSTY-ARIDCTIS---IVRLLV    MLFYKGEgAL-RGFFDKFSYK--LLVFTSADIAR--  498
Query_RACEM   429    A-TLTRA-IDVFVTDYRRLSSW-ARESGTVA---IQDALH    SFYSKSEgDL-RAFATTTFYN--ALIWTCTETRR--  493
Query_PVIOL   466    D-QIEQV-LETYFIIYPHLFVK-SHHRFHMA---VNRLIS    SMYFRGE-YL-KILLSRIVKK--GLLISFSKANKSV  531
Query_BOMBU   422    N-IITKL-SILLIKRFPEFPVS-NQNFAVSL---LVNTII    NIMTINRdLL-DEFLYNLIYN--GVVWTSLHTLFVD  488
Query_AEDES   430    I-ALQNI-IIGLVKDFHYLSSS-YHELIVSS---LMKTFD    NLGKLGGgVL-DNLLEKIVLR--GIIWSCSHMLVVD  496
Query_ANT__   433    N-ILVKL-SAYTIKRFPDLTKS-NNALAVLT---LIKTIS    NLATVNKsLL-QRYLDSIVYD--GVAWSCSHTLALD  499
Query_WHITE   430    V-SLETL-TIMVMKMYPDCGRN---WKASQT---IERTVL    CLSLKSEkTL-DTYLGVIVFQ--GIIRTCDTLILQD  494
Query_BEETL   425    F-CLQKG-TILLMKIFPQLPYY-CHPLVTEA---LVLTFY    YLSLNSEhIF-NTFLENVIFQ--GVVWTCSHQHITE  491
Query_THRIP   403    S-ILESL-SVQLICQYAKEEHPdWQRVMVDA---VICTLQ    RLRKCG-vAL-DSFMNVVVHQ--GIIRSVTHPVIED  469
Query_TERMI   426    V-ALQCL-TVQLIENFPKLPKP-FHWLAIRA---ITVTLS    SLAESGGtFL-EEFLSNVVYQ--GIIRSCKHQIAID  492
Query_FLEA_   416    K-ALQNL-IVILFQEFHQLSGS-HHNLCIES---LMKTFF    NITQINGcFY-EEVLEVTIYH--GVIGTCSHPIAIE  482
Query_MIDGE   439    Q-SLETI-MVILIKDFHFLSTA-HHPLVASS---LLETLL    NLQKLGGkML-HDVMESIVWQ--GILWTCSHQSVYD  505
Query_LAODE   431    E-NMERI-SVKLISACNTLSSN-KLDAAVNS---IVACCI[ 1]AHVARGKvGS-KDFINVIVYQ--GIIKACCQQPFVK  498
Query_LOTTI   427    S-IVEHL-LVMQIQNMPMLNEK-MHFMATKS---VIHVLL    ALMCKGS-VF-NQVLSGVVYQ--SVIRTCSHPIVLE  492
Query_APLYS   433    V-VLERL-IILLMENFHRVVKK-LQFVTYKS---ILLLLL    TLVDKAA-VF-QQILSTIVYQ--GLIRTCSHPPLAE  498
Query_SCALL   434    I-SLERL-LVLLIEKLPSLPQT-LSFICLKS---ILWLLL    AVMPKGT-TF-QQIMSGVVFQ--GLIRTCSHPPVVE  499
Query_CRASS   434    T-SIERL-MIILAENFPTIHQK-QHFFSVKS---LLWVLI    TLMPKGE-VF-SQVLSGFVFQ--MLIRTCSHPPLIE  499
Query_OCTOP   424    K-IIEYL-LVAMIENFPQVYFK-YHFKCVKA---ILKVLL    ALLPKET-HL-REILRETVYQ--GLLRM-SHLATPL  488
Query_POMAC   430    L-TLERL-LVLLLENIPTIYRG-FHFTCLKG---ILHLLI    VLQSKGS-TF-NKVLSGFVYQ--GLIRTCSHPVVMG  495
Query_WHIPW   443    Q-IVEDI-HVKMFINIPALGAK-WTARCKQVt---YAIFA    AALTKNR-SY----ITSVVDQ--ALLHVLTLPPTLD  505
Query_TRICH   444    Q-LVERF-HVTLFHNFPLMEKR-FYHRCKVIn---VATFA    AAYGHNY-QY----IQSIVFQ--SLIRVIASLPSV-  505
Query_AASTA   450    --FSVEL-VCRLLDAYPSSAIY-VKYQAELA---IVSMAD    AFSSTDV--M-K----RILQH--GMVLTVSNRIDTP  509
Query_AEUTE   443    --FAVTL-ACRILEAYPSSAIY-VKYQAEIS---LVSIFQ    ALPTGNA--I-D----RVIRH--GILLTVSTQLNPS  502
Query_PLASM   443    M-FLERI-LCHLMTAYPQCLLK-SKSVVHKS---IAALFV    SLSKWTV--I-DPVVDQFVHH--TLLFSIADITDPD  507
Query_BREMI   442    K-FFGDM-MCHLMTAYPQCHWK-SKVVVHKP---IAAVLV    SLSKWTV--L-DPLINRFVLH--TLMLSISSVTESD  506
Query_OLIGA   443    T-FLGDL-ICHLLDAYPRCLWK-SKLTIYKS---ILSILC    AVHSSGD--V-QLLLDRIVHR--AFLLSISNVTETE  507
Query_LIGNO   424    S-LLARL-SVTVCHMYPRVSER-MGRVGLRA---VAHALH    RLRSRRP-GLaKAFAAQVVYQ--SLVRTCSHEHPVY  490
Query_CESTO   336    Y--SLKLgIISCFQAYPRLSQT-STESIAVA---IVKIIE    KMENDKDvLL---ILDEAFYHi--VLLTCSHSPFNL  400
Query_OVI2_   365    Q-LLLRL-CVRIVELFPLSSVS-MSTTFSAA----IRLAI[10]ALNTMQP--L--NFGEQLVYR--AVLLSCRHPPIVF  437
Query_PRIAP   437    A-SLEQL-VVLLFGNVPLLRRE-SR-LDHRAcgaVVRVLL    ALRSKRA-TQ-RSLVKRVVFQ--GLLSTCSHPVAME  504
Query_ROTIF   435    N-TMEKL-VILAIDCYPKLVKR-YNNQISLA---IAALFL    AIQ-RTN-FS-DEFISRIIYH--SLIRIFSYKTLYS  499
Query_SPONG   437    S-SLQKL-SVLLFHCYPFLSSH-VRYQGSVS---LLKLIM    AVSSIPT-SF-KQFLSHIVYQ--GVIRSCSLPVVIE  502
Query_TARDI   447    A-AVQRL-CSLVIQMFPQLGPH-IQHTACQA---VVKTVM    VVARACP-TA-NDFVGEIVYK--GLRGSCLLEMPSD  512
Query_TRICH   429    R-SLEKL-TIVLFENFPYVINE-LHSLCYKS---LFELLL    ALHSKGA-AL-KNFLSRVVYQ--GLIRTCSHPVYVE  494
Query_OSTRE   454    V-TVGDL-VEWLWEHYAFEQSR-RKPLIQNA---LWNLFV    ALSSKGC-AL-QTLLTRIGAKllGLTLRVSPPDPVD  521
Query_NITEN   442    S-FISEV-AGAIFFRYPLLWEK-QKPAVHTA---FVNMLC    SLYPKGE-AF-TSFLNNIVRS--SLVRTMSADDPAA  507
Query_CLUBM   439    S-LITEV-TGRIFFQYPSLFPK-QRVAIHST---ISSLLK    ALYPKGN-LF-VMFLMNVVWS--SLDRATDT-----  499
Query_MOSS_   433    T-FIAQV-TGSIFFLYPQLFAK-QKQAVYTA---LRNLLK    ALYPKGP-VY-FLFLRNIVWS--SLQRAMEPVDEAL  498
Query_LIVER   439    S-FIAEV-TGNIFYQFPELYPK-QKQGVYNT---MSAFFK    AMYPKGP-IF-SIFLSNVVWS--TLERTMETSEGVS  504
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPITE 501   S---   --[ 1]EAKSEN   VDEVNLKV---DVSYQDYLP-MWKALLTAGK---LK   ALGQY-NITAE   QHRVL  552
Query_LINGUL 499   L---   --[ 1]EVEAEV   DMEMQMESE-KKITYKDHLD-FWNMLLGSSH---LK   --VHS-KLTPD   EKSAL  550
Query_LIMULU 503   VEAR   K-    -DEVET   YDP-------RQITYQNYLS-LWHAILDIVN---SK[2]EVGTL-QVPLA   ERQSL  554
Query_SPIDE  507   AEAM   KD[ 3]GKEDNQ   KNDNFISAS-KIITYKDYLT-LWRDMLNVVR---FK   ELNSL-NIPLE   RRQYL  567
Query_SCORPI 499   IEAR   EE[ 3]GDNLEL   KNP-------REITSKSYSS-LWENLLYIVR---NK   ELNHI-GISLE   DRHNL  553
Query_TICK   488   VEAS   --    -REDDG   -----TSQS-RTITYRRYSF-LWAYLLNAVA---VQ   ELNQM-QISLD   ARQML  537
Query_MONOSI 441   NNPD[4]ER[ 2]------   ---------------RNIMVeAFAHLVSQPPptsVT[2]------DGGAS   -GTDV  484
Query_ROSETT 425   EDMA[2]-A[ 2]------   ---------------QSYLA-LLRQLLKGSGl--VK[2]------DEEMT[2]DRNAM  465
Query_BELCHE 506   GVTA   ED[ 6]SGEEGG   EQQPQQPPV-KRFSYRDYLE-LWGNLLDSSR---LK   ELSQW-GVAVA   DRQAM  569
Query_CIONA  507   TGRD[2]VD[ 6]GEDDLN[2]-----LLDP-RTITYRDYIP-LWLALLGESE---KK   EVPYAvNDILM[2]HRKVC  572
Query_CAECIL 505   KEGS[1]KG[ 2]GSEEPP[1]AEEVRIGKW-KVPTYKDYLD-LFRSLLDCDK---MK   ESIFS-DEVFS[3]PLQSL  569
Query_HUMAN  499   K---[1]PE[ 1]ESEDHR[1]SGEVRTGKW-KVPTYKDYVD-LFRHLLSSDQ---MM   DSILA-DEAFF[4]SSESL  560
Query_MOUSE  499   K---[1]VE[ 1]RSDNRS[1]SEEVRTGRW-KVPTYKDYVD-LFQHLLGCDQ---ME   DFILG-DETFL[4]SLKSL  560
Query_XENOPU 519   QD--[1]KE[ 2]EAETAA[1]TGEVRAGKW-KVPTYKDYLD-LFRNLLRCDQ---LK   DSIFS-DEIFS[4]PLQSL  582
Query_PLATY  499   KEAY[1]RD[ 2]ESDDSS[1]FGEVRVGRW-KVPTYKDYLD-LFRSLLNCDE---MK   DSQL--DETFL[4]PLQSL  563
Query_TASDE  500   KEAF[1]KE[ 2]ESEEYP[1]SGEVKTGKW-KVPTCKDYLD-LFRNLLSCDQ---MK   DSLLA-DETFL[4]PLQSL  565
Query_CANAR  495   KDFS[1]KE[ 2]GTEEAF[1]AGEVGTQRS-KLPTYKDYLY-LFRSLLSCDT---MK   ESVFE-DANFL[4]QLQSL  560
Query_OSTRIC 466   KEYF[1]KK[ 2]GSEESP[1]AGEAGAGKW-KVPTYKDYLY-LFRSLLNCDT---MK   ESVLE-DENFL[4]ALKSL  531
Query_TURTLE 504   REIF[1]KE[ 2]KLEEST[1]AGEVRTGKW-KVPTYKDYLD-LFRSLLNCDT---MK   ESVLA-DETFL[4]PLLSL  569
Query_GECKO  503   RENS[1]TG[ 2]ASEEHP[1]SGEVRSGKW-KVPTYKDYLD-LFRSLLSCDT---MK   DSLLA-DETCL[4]PLQSL  568
Query_DANIO  514   DERG[4]VS    Q-----[3]STLVRSGKW-KVPSSKDYLE-LFKGLLDCEN---LK   DTGFV-DGAPA[4]NLRDL  577
Query_SNAKE  503   MDFG[1]KN[ 2]KSEEYD[1]SGEVRTGKW-KVPTYKDYLD-LFRTLLGCDT---MK   NILLA-DEVCL[4]PLQSL  568
Query_STERL  508   GEEG[1]QG[ 2]Q-----[3]KGQVHTGKW-KVPSYRDYLD-LFRNLLECDK---MK   DTGFV-DGAFD[4]SLHSL  570
Query_MILII  503   GEGT[1]RS[ 4]QQQSAS[2]---MWTGKW-RVPSYKDYLD-LFRSLLNCGK---MK   EFGLV-D-TLQ[4]PLQSL  567
Query_BAMBO  504   GEEM[1]KA[ 6]QQSLSN[2]---MGTGKW-RVPSYKDYLE-LFRSLLDCGR---MK   ESGLIyE-TFQ[4]PLKSL  571
Query_ALLIG  464   REFD[1]KG[ 2]KSEEPP[1]VGEVRTGKW-KVPTYRDYLD-LFRSLLNCDA---MK   ESVFA-DENFL[4]PLQSL  529
Query_GAR__  503   EEEG[4]AE[ 2]R-----[3]AGQVRTGKW-KVPSYRDYLD-LFRSLLDCDL---MK   DFGFL-EGVFE[4]PLQSL  568
Query_TETRA  515   LSMQ[2]--[ 2]------   -------------TLKSCSQ-FWKSLFNCQLwt---   ------DETLH   -----  546
Query_PARAM  502   DGDL[2]--[ 1]------   -------------TCQEMSK-FWIGIIEQKVwt---   -----------   --QHY  530
Query_SLEMN  500   FGGE[2]--[ 2]------   -------------SLQDAVQ-FWREWLNKDKiw---   -----sEYTCQ   -----  532
Query_STENT  490   DSSN[1]--    ------   --------------IELASK-LWTGIFQQECl----   -----nEIILD   -----  517
Query_NEMAT  501   E--G   PD[ 1]------[3]VGVVTGSRS-DQVTFKDYLP-LWEHMLDKTK---FK   MYLLF--QSLS[5]KRESL  558
Query_PISTI  504   EVNG   ES[ 1]------[3]EREVYEGST-PNTSLKDYLE-LWEYLLDKDK---LK   D------ESLL[5]ERKVM  559
Query_CORAL  499   E---   -G[ 1]------[3]EG-VTEFDV-RHISYKDYLE-LWECFLDGGQ---YK   D------QSIM[5]LTQDI  549
Query_HYDRA  485   ----[2]EI[ 6]KKDAL-   ------------ITIKTFLP-IWNFLLEEKN---YS   ISVYSrSELCV[1]----L  532
Query_LATIM  500   TEGI[1]KG[ 2]QSEEEA[4]PGEVRTGKW-KVPSYKDYLD-LFRNLLECDK---LK   QSALV-DETFG[4]PLQSL  568
Query_SHRIM  497   ESSP   GE[ 6]---GKA[2]GEEWKGLKN-EVVTFSSYIP-LWTALMVPEK---VA[2]KVSGMpATVMT[3]-----  560
Query_MCRAB  495   DSLK   KE[ 5]ESNDSE[2]GGKWEGFKK-EAVTYKNYLP-LWKAIMMPEK---VS[2]KIPGVsTTILT[3]-----  560
Query_AMPHI  497   DAVP   AS[ 1]-----P[2]GRRWFGVAD-EPVSCASFFP-LWRALLCPTA---AS[2]KCSVLsVTCVT[5]RRSAV  560
Query_BARNA  503   TELQ[4]VS[ 2]------   ---------dRPLSYKDYLP-LWDALLAAHG---LR[2]--------TRK[5]ALQAL  552
Query_COPEP  497   TQIL[2]--    ------   ----------DVLTFKKYVP-LWNGILQSGNnp-SR[2]-----------   -----  528
Query_DAPHN  487   SVKK   --    ------   -----------------YTS-LWITLINFSGqa---   NLPFSlEERAY   ----V  517
Query_STARF  502   G--S   PS[ 6]GNDDQD   TEADILADA-RHITYKDYVK-LWASLLDSTH---VK   EVQAM-GMNYA   DRRKI  563
Query_STRON  504   S--G   GD[ 1]GEEQRS   DEMEIEQDS-RRVSYKSFMD-LWTRLLDSKF---IK   TSEAH-GSSYA   ERLNI  560
Query_SEACU  502   G--A   GV[ 6]GGEEDD   LEVA----N-RMTTYKSYLD-LWASLLDSQY---LK   EIDAM-GMTYV   ERKKI  559
Query_FUNGI  531   RPRI[3]EL[19]------   -------------QTNEREA---SSASSAYY---QS[2]-----------[2]AKEEI  584
Query_DPURP  539   KALI[3]HT[22]KDLASS[1]-------------------K---FNQQQNNS---NN[3]INNDKhNIITV[5]IAQSI  611
Query_HALBU  533   LPHL[3]DT[20]ISRGIQ[4]QHVVQQVKQeEGEKEKEKEK---DSISI--------   -----------[1]TSEDI  602
Query_GLOBO  462   ----   --    -DRTAG[1]---DELVAEsREPTWKDYLV-FWETLFV--------   AKNEAwDESKV[2]----F  505
Query_PALUS  493   ----   --    -DDIVN[1]---GQPAAVrF--AFDDYLF-FWEQLFKWT------[1]AEKERlDSEDR[2]----L  537
Query_CALIF  494   TFQK[3]ES[ 7]DDQPAN[1]INVIDEMAKeNEHSYNDYYYsFWKVILDKNN---LP[1]RNEFKmEDDTL[5]KRNQF  571
Query_GLOMU  510   ----   --    -LQNRS[4]ANSKDMIDVvAEHTYQEMFF-FWENIFKDST---LS[4]QSGFLiDENRL[2]IQLEF  572
Query_LOBOS  499   ----   --    -PQPPS[2]IDQTITTDQiAWHSYSIYLF-LWRNIFKPTV---LI[4]KTRIDiSGEDL[1]---RF  555
Query_RACEM  494   ----   --    -DAPNE[1]------------QVYVEYLA-FWESIFKT-------   ---YEaQPSDL[2]---QF  527
Query_PVIOL  532   LGLI[3]DT[22]IPEAFR[1]-------------------K---SNIQQ--------   -----------[1]VNTSI  581
Query_BOMBU  489   AELQ   RG    ------   ----LNNLSeRPLCYKDYLP-LWLQLLDQNK---YS[1]------DFKYS   --RET  531
Query_AEDES  497   VNQG   RG[ 2]------   ----ENSNWkDYITYRNYLP-LWRGLLTPGS-----   ----SlSLDRG   ---PL  539
Query_ANT__  500   AELQ   QE    ------   ----LNDLKqMPVCYKNYLP-LWTGLLNAER---YA[1]------E----   ---QL  537
Query_WHITE  495   VEFH   KS    ------   ------E-NmEVTSYKDYLP-MWESLLRLS------[3]SIEPTiPECEA   --ARI  539
Query_BEETL  492   AELL   ET    ------   ------S-SqKVVTIKDYLN-MWKSLLKIVP---VR[3]KFRIFlQERKY   ----I  537
Query_THRIP  470   AEIL   RD    ------   ------T-QqKVITYENYLP-FWQILLTLDT---SN[3]LREASiEEKRD   ----I  515
Query_TERMI  493   VELQ   KG    ------   ------QFHeNVFTYKDYIP-LWKELLSVNN---FQ[2]--STSlTQTQL   ----I  536
Query_FLEA_  483   SFE-   --[ 1]------   ----NEDFYkNNISYKNFLP-LWIGLLSQK------   ------DGVQK   ---EL  518
Query_MIDGE  506   LEEN   L-[ 1]------   ----NVKDWkETITYKKYLP-LWSRLLSP-------   -----lDENHE   ---EI  543
Query_LAODE  499   QLTR[2]--[ 1]-EETAS[1]-----------KVVVDDYLP-FWEKLLLVHKne--K[2]----DtDSLYR   -KEVI  546
Query_LOTTI  493   T---   --[ 1]QTDDDT   LYNVDVEK--QSISYKSYLN-IWNLLMDSPE---IK   VVLFE-SFALF   ILTQL  545
Query_APLYS  499   T---   --[ 1]QSEDRL   GDMDVPGAN-KTVTYKDFVP-LWLGMLDSER---IK   DLAAE-DFTLA   DRREL  552
Query_SCALL  500   L---   --[ 1]QTEEST   NPQPEGVMV-RKITYKDYIS-LWTGLLNSAR---IK   DFSVA-DISLD   VRHRL  553
Query_CRASS  500   T---   --[ 1]DMTESN   QSEQTGT----KVTFKDFIE-LWDTLFNSPR---IK   EFSVA-DISVE   VRLAL  550
Query_OCTOP  489   KEME   QP[ 3]GTF---   ---RVWSVK-KEVSYKDYIS-LWSNLLKSSS---LK   ELSTV-GYSFS   DREKL  543
Query_POMAC  496   A---   --[ 1]ENQEIS   TAETDLTV--KQITYKDYLE-VWSGLLESGK---IK   DLIQD-GIGFA   ERDQL  548
Query_WHIPW  506   GNEQ[2]--    ADEGYN   --------G-RQFTAANHAS-LWRKLLAYDI---VD[2]DFGYD------[2]TQKLT  554
Query_TRICH  506   GKES[2]--    LAKDEN   -----------TFHLADYLP-FWKMLLDSEF---VD[2]EFGFD------[2]MQVNI  552
Query_AASTA  510   DGEM[3]PD[10]------   ----------------EYEG-LWRGCLKRMQ-----   -----------   -GEEL  546
Query_AEUTE  503   QEET[3]PD[10]------   ----------------EYEI-LWRNLLKRLK-----   -----------   -DQKN  539
Query_PLASM  508   QA-V[3]PG[10]------   ----------------EYEG-FWLNFLRLRServVG[4]PTEAGsGDVQR[5]FDRTL  569
Query_BREMI  507   EV-V[3]PN[10]------   ----------------EYEG-FWLAILRLRTeglVD[4]PVETSnIDVPM[5]TRRTL  568
Query_OLIGA  508   ELMV[3]PE[10]------   ----------------EYED-LWVALLRGKKhtkYK[4]VTPHTsSDT-R[5]TITAI  569
Query_LIGNO  491   QSDT[3]SS[ 3]ASDSAA[2]VQGDASNSV-RKISYKDYLR-LWLCLFSLEK---EK[1]---LC---DYV[5]EARSI  557
Query_CESTO  401   DVSE[4]ED[ 9]DDSLEK[4]NVPSQSEVSiPLPSFRQYLP-LWKYLLGLPTg--IR[2]-----kNHALI   -PKRC  474
Query_OVI2_  438   TPVD[2]-S[ 2]-AESPR   SKPLQMSDA--PVSYQDYVP-LWKALLGSGLdg--K[2]-----------[2]SNESG  491
Query_PRIAP  505   T---   --[ 5]QNADDV   DDDFQERDP-RNITYIDYMD-MWRHLLGSDK---MK   GLNEL-GVSMS   DRKAL  562
Query_ROTIF  500   VNQE   -D[ 3]ETNEPM   ------SSI-IVTTSLDYVM-LWSNFLGLQE---FS   DLKSD---NLD   DKKKI  551
Query_SPONG  503   AVQG   --    --DDSA[3]AAPLADSND-KVMSYLDYVQ-LWIQLIDYKK---LK   GMGMDfDLELR   --QSI  558
Query_TARDI  513   VDGL   --[ 1]SPPPAG   ----SMGTK-KSFSYKDYLP-LWNKLCGCTY---MK   ELLHElEASEV   ELQRT  566
Query_TRICH  495   T--D   SD[ 1]------[3]QDSVILETG-PRISYKDYIS-LWKNLIQATE---IK   GL------TLI[5]HIIYV  548
Query_OSTRE  522   KALY[2]--[ 2]------   -----------VPMWPRYVE-LWTTLLGRVDdt-VS[2]----------S[5]DIQIA  565
Query_NITEN  508   PP--[2]PK[17]-VLSAA[4]KEEASSEADsETPVASRI-----SVNAADY------   -----------   --AHM  567
Query_CLUBM  500   DP--[2]DF[14]---NIS[4]TENLQDD---DGIVGVDL-----AVLSVNV------   -----------   --SDL  551
Query_MOSS_  499   EE--[2]VQ[17]-DQGLQ[4]LDEAESF---EELVTSDV-----GNVGIIE------   -----------   --SSL  555
Query_LIVER  505   GD--[2]LP[16]-KVSVG[4]VPVAETD---DEHMSGELLL---GNGTTDV------   -----------   --TEI  562
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  553   ASAIYDELLSAMLKMISRLDLSS   NST--     --D-SeSQPSTSEPALS---[3]SDPLRGLHAVCPEDFIL    L  613
Query_LINGU  551   TQAIYDELVRSTLKILSKLDLTS   VRGSE     DtD-QgDSAGTADKGRP-VT[5]SDPVSGLQASKPKDFQV    F  619
Query_LIMUL  555   AENVYDEIIQAVFAVLKKLDLS-   -----     SsEnTvSDEVNADQLSVVEE[4]-----GRQAACPKDYQV    F  613
Query_SPIDE  568   MECIYDEIIRSSLNLLNKLDLSL   QQPEF[ 1]GkEdDnESEISCDPLHGVQP[4]-------------DYYI    F  625
Query_SCORP  554   LQLIYDEIISVVLTLLKKLDLA-   -----     TtHsMqDSEIKDSNLKQFET[7]SDPFYGLQLSKPKDFRV    F  620
Query_TICK   538   VRDVYDGLLSSALSVVSKLNLSV   RQQAS     ----------SDDGEARVDV   SDPTCGIEAQCPKDFNI    F  593
Query_MONOS  485   MDHMYAATLEAIMQVLEKLDLSA[8]RSQSA[ 2]---------APSSQAITMTS[2]TASLDDLRATRAKDFQV[10]L  563
Query_ROSET  466   LRTLYDETLKAFLHLINKLNLTT[2]-----     SsDaEdEDVAVATSGVKEEE[7]SDPTTSLIASRPKDIQV    L  535
Query_BELCH  570   HRMLYDELVTSVLKVIDKLDLTT   SKQVE     Tg---nEELTADVSGSQ---[4]SDPFAGLQAAKPRDFVV    F  633
Query_CIONA  573   HEMYTAAFVKAIVAIVDKLNLET   SAENK     -------------EVIENP-   -GLTDKLKAMNVTDFKI    F  623
Query_CAECI  570   NRLLYDELIKSVLKIIEKLDLTV   EKQSV     G-E-QqD-ASDEDSALVIPT   SDPAANLQPTKPKDFTA    F  632
Query_HUMAN  561   NHLLYDEFVKSVLKIVEKLDLTL   EIQTV     GeQ-E-N-GDEAPGVWMIPT   SDPAANLHPAKPKDFSA    F  623
Query_MOUSE  561   NHLLYDEFIRSVLKIVEKLDLTL   EKQTV     GeQ-E-D-GSTAD-VWVIPT   SDPAANLHPAKPSDFSA    L  622
Query_XENOP  583   NRLLYDELMKSILKIIEKLDLSL   QKQDT     G---QeD-DGGINNLLINAT   SDPAGNLYATKPKDFTA    F  644
Query_PLATY  564   NHLLYDELIKSVLKIVEKLDLNV   EKQNT     G-E-K-D-ESEVSNLWVIPT   SDPVANLHPSKPKDFTA    F  625
Query_TASDE  566   NRLLYDELIKSILKIVEKLDLKL   EKQNV     GnQ-E-D-ESEASGVWVIPT   SDPTANLHPAKPKDFTA    F  628
Query_CANAR  561   SRVLYDELIKSVLKIIEKLDLTV   QKLDI     S-E-Q-D-GNAGDSALVGPS   SDPASNLQPTKPIDFIA    F  622
Query_OSTRI  532   SRLLYDELLKSVLKIIEKLDLTV   QKLNV     N-E-Q-D-EDETDSALVVPT   SDPASNLQPAKPTDFIA    F  593
Query_TURTL  570   NRLLYDELIKSILKIIEKLDLTV   QKLNV     S-E-Q-D-ENEAGSDLVVPT   SDPASNLQPTKPKDFIA    F  631
Query_GECKO  569   SRLLYDELIKSILKIIEKLELTV   QKQNV     N-DaK-D-ANEADADLTIPT   SDPAANLHPTKPKDFTA    F  631
Query_DANIO  578   NRHLYDALVQSVMKIVEKLDLSV   QKVSA     A-D-E-V-QSDASAGIVL-S   SDPTANLMPNKPKDFIA    F  638
Query_SNAKE  569   NHMLYDELIKSVIKIIDKLDLTV   QKQNI     N-EeK-D-GSEGDAHLAIPT   SDPAANLQPSKPKDYSA    F  631
Query_STERL  571   NRLLYDELIKSILKIIEKLDLSV   QKQSA     D-E-Q-EGRSDASSSMLIPS   SDPAANLQPSKPKDFTA    F  633
Query_MILII  568   NRLLYDEFIKSILKIIEKLDLSV   EKQNQ     R-E-E-EGSGDVSNGNVIPT   SDPAADLQPVKPKDFTA    F  630
Query_BAMBO  572   NRLLYDEFIKSVLKIIEKLDLSV   DKQSQ     N-E-E-GGSGDVNSSIVIPT   TDPSANLQPTKPKDFTA    L  634
Query_ALLIG  530   NRLLYDELIKSILKIIEKLDLTV   QKLTF     S-E-Q-A-ESESGDVMIIPS   SDPASNLQPAKPTDFIA    F  591
Query_GAR__  569   NRLLYDELVKSILKIIEKLDLSV   QKQNP     D-E-L-E-ESDVSSGVVIPS   SDPTAHLLPSKPKDFTA    F  630
Query_TETRA  547   --HFSAQFFQYIIQLLENTDLSY[8]EEDEI[ 4]QmEiEqDKESKIQDIIREIQ   ---QVQYVANNEVDLQF    F  619
Query_PARAM  531   LTKFVKQLFDQIAQLIQNCNFES[1]QKEVN[ 1]--------------------   ------IDEQNQQQYTL    -  571
Query_SLEMN  533   --YVYGELIQRVIRIISDLNLTY[4]GRNNN[ 2]GnQdEsNNNNNGRDPQILNQ[6]QDYIVQYIANNNTDQQF    L  608
Query_STENT  518   --DMAQEIGSFFYHQISTLELSY[4]G----     --------------------   -----ELKAENSEDQQL    L  556
Query_NEMAT  559   HELVYDELVGAILRIIHKLDLTS   TKHG-     --QqDlQVRDQPQE------[6]SDPVAGLQPRVQKDFLV    F  621
Query_PISTI  560   HGLVYDEMIRGILRIIHKLDLSS   SKGSS     AqEgDiLGKSSASDVPASAT[7]SDPVPGLQPKVPKDFLI    F  632
Query_CORAL  550   HRAIYDELIQAILKVAHKLNLSS   FRDVE     TqDeDlQTEQPSSPDAPLLT[5]TDPISGIQPEAPKDFQI    F  620
Query_HYDRA  533   KELLYDCILSAFLNVLNKLNFVT   ENKEA     -------------EKLRDPV   SDPTLNLVPKMPQDFII    L  585
Query_LATIM  569   NRLLYDELIKSILKIIEKLDLNV   QKQKV     G-D-Q-EGETDPSSTIIIPT   SDPAANLQPSKPKDFTA    F  631
Query_SHRIM  561   -NFIYSEFISCILEISEKLDLTT   YKQNE[ 1]QqQeE---VSGSDEGDPATS   -DPVSGLKAKVPKDFQV    Y  622
Query_MCRAB  561   -SKLYSEYITSTIEIAEKLDLST   FKQEE[ 1]EtNdI---TS--DDGDVATS   -DPISGLSAKTPKDFQV    Y  620
Query_AMPHI  561   CSALFTEFSSAADLILGQLDLST   QQQQQ[ 1]QeDeLpPQALDATDENPVRT   -DVVLDLVAKTPKDFQV    Y  626
Query_BARNA  553   ANEVFAALVDASQKIISELDLST[2]-----[ 1]------------DGDAPPAS   GDPAVSLQARVPKDFQV    F  604
Query_COPEP  529   ---ILDSFLRSIFLIVDKLDLTV[8]QPEPE[ 3]MeIsLtQSDVDAVRSVPIPS   MAPVLGKRAAVVKDYTV    L  602
Query_DAPHN  518   SAILDEQLVSAVSTIISRLDLST[2]ENSSE[ 1]--------------------   -EFKAEYLPSNPADWQL    F  565
Query_STARF  564   HIILYDELIFSILKILSKLDLTS   NKDTT     E---LqDQVA-VDSDST---[4]SDPLHGLHATKPKDFQI    F  626
Query_STRON  561   HTMIYNELIGGILKVLKKLDLSS   NKHNE     Ee---aSTSGGVTSDLE---[4]SDPLEGLQPTKPKDFQV    F  624
Query_SEACU  560   HATLYNEMMSSMMKILTKLDLSA   TKSQA     EgE-IqGEVSNADTNTS---[4]SDPVLGLQAAKPRDFQI    F  625
Query_FUNGI  585   SKTVYNSLMGGMILFIQKVELAS[1]-----[ 5]GeTvQqDTTSQTA-------   ------------DYAL    F  631
Query_DPURP  612   TKLVYNEIISSILVVIGKLDLQY[4]--TSA[ 9]---eEeEKD-----ISIIE-[1]ESEVNQLKPNTSKDIDL    F  680
Query_HALBU  603   IQLLYNEIVQTMITFIRKLDLSY[8]RQEKE[14]GeEeKeEEEQQEEEDVTMED[7]NIGEDTLKPSTPKDIEL    F  697
Query_GLOBO  506   RSLLYDGMVTAILELPGNLNLSV[2]-TGPD[ 2]-aStVvPI------------[2]D--ASKLVADNAADFVI    F  561
Query_PALUS  538   DDMLYDETMKALIQLPQNLNLSL[2]-VNQP[ 1]-----iVE------------[2]D--DTNLQAVNGVDVAI    L  588
Query_CALIF  572   IQKLYDIIIQTILDFPKHLNLAL[8]RINDN[11]NvTiTsTMRSGENNNEIIEE[7]NV-QT-LKPYCIRDFDI    F  661
Query_GLOMU  573   FSILYDEFISSVFKMIRTLNLAV[8]GTSIT[10]LiAiAg--------------[2]DL-ET-LQPLVSKDFVL    F  642
Query_LOBOS  556   LCVVYGCTMESFKRLILLLNLTV[7]EEGVD[11]SvStVgPM------------[2]DI-AS-LQANCAKDFLI    F  627
Query_RACEM  528   LGILYDSFLTALLRLLKLFNLEL[8]ELHMP[11]DtRvAs--------------[2]TL-ETSLQPVNEKDFVL    F  599
Query_PVIOL  582   RKLIYDEMISSILSLIDKLDLQY[8]NNNKN[11]DqLeEeEEDENNATIKTIE-[1]ESEFDQLIPKTPKDIEL    F  666
Query_BOMBU  532   VQYVIDSTMHVGIALIDKLNLCT   K-TKE     ------DNIFSDVAFSQI--   --------AENEADFRT    F  580
Query_AEDES  540   IKVIYAQLMKTLFLILDKLDLST   RKRTV     ------RDDSGEDQDV----[2]CDPNIDLVPVKPKDFHI    F  597
Query_ANT__  538   AQLVANTMMDVCITLINRLNINV   KRTNE     ------DTVLSDVAFTQS--   --------AVNQADFRV    F  587
Query_WHITE  540   RKLLYRHLIDSALLVIDRLDLST   VRKED     ------DENQPVIN------   -DIKSKYDVRRPKDFTM    L  592
Query_BEETL  538   LVKVIDKLMRTLLILINKLNVSV   TLNDN     ------SEAVT---------   -DIQSAYKLNAVNDFAI    F  587
Query_THRIP  516   SRALVSKILDSVQDIINRLDLSL   QENNS     ------SESIDHC-------   -NVEKCVEATNPQDFTV    L  567
Query_TERMI  537   VVKVYDKLVNTVLVLMKKLDLST   RKKTN     ------IEALSTSGGASEAD[3]LDPESGLEANNRKDFQV    F  599
Query_FLEA_  519   SLRIYDNIMKAFFAIINKLDLST   -KRKV     ------KNKSNDDDIL----[1]CDIKSSLEPNKLKDYHI    F  574
Query_MIDGE  544   SSLVYMHFINNLFKIVESLDLST   KKRKF     ------YDESNNVEKE----[3]SDPSLDLEPLRAENFQV    L  602
Query_LAODE  547   VEPVYEAIIESIVKVINNLDVNF   EESNY[ 2]--------------------[1]TELTSGLKAHNDADMAI    L  595
Query_LOTTI  546   TDALYNELMNSVIKVLNKLDLTS   TTSTE     LtE-QmDEASTSGGDVT-ME[5]ADPTSSVHASKPKDFQV    F  614
Query_APLYS  553   TKAVYSQLMAALMKILRKLDLEA   TVA--     ------DSALEGDAEME--E[5]ADPTYGVIAKKPKDFYV    F  613
Query_SCALL  554   TEVLYDELIACVLKILCKLDLTS   TSTSE     LsS-QsDQNSKAGASGEGLE[5]GDPTQGIQPNRPKDHQV    F  623
Query_CRASS  551   TEAVYDELINSVLRMLGKLDLTS   QSSTE     VaQ-QqEGPTSSNESED---[2]SDPLHGMQPNKPKDFQI    F  614
Query_OCTOP  544   TQLLYDEILTSVIHIIESLDLTI   -----     -nTsIsQDVVIQPSASTISP[7]ADPVHNIKLNCVQDHQV    F  610
Query_POMAC  549   THAVYNELMQAILKVISKLDLTS   TTQRE     SeT-SgTEEEKSDITQS---   GDPVHEIEANRPMDFQV    L  610
Query_WHIPW  555   EEVFYEVMMRCLMEILQTVKLNA[3]DEDAS     ------------------PE[3]RSTTAGLLPANQADFNI    F  608
Query_TRICH  553   QRTLFDCFMKSFLEIVRLLSFET[3]QDVEE     ------------------YT[3]GNPVFNVIPTKQADYDL    F  606
Query_AASTA  547   QQAMVNAMADTMLTILQRLDLRY[4]DTAET     ------------------ST   ---SSQYTPTVARDHTI    V  595
Query_AEUTE  540   HGKIIDVFMKSILEVFERLDLRY[4]NKTE-     ------------------ED   ---ASEFEPLISRDQTI    A  587
Query_PLASM  570   QAIILDSTVKHVLAIIRKLNLSY[4]QTTKI[ 1]------------------RD   ---ATGYSPTVPRDHSI    M  619
Query_BREMI  569   QAIFFDSTVKHVLDIISQLDLSY[4]HATNN[ 1]------------------KD   ---TIRYTPTVPRDHSI    M  618
Query_OLIGA  570   RGMLLDSLFKSVQAVVGGLNLAY[4]Q--DG[ 1]------------------AQ   ---GGGYMPNVARDHSI    M  617
Query_LIGNO  558   SGGIFDAFLSSVLAIIERLDLSA[1]-RLSD     -lEpPvTAASEASSSSATTD[7]TDTLQGMVASTPKDIVI    F  629
Query_CESTO  475   SALVLKFLLDSTLVIVKRLDYSY[4]--ANT     --------------------   TDLETQVDVKIPHDVLI    L  522
Query_OVI2_  492   ASTVCRHFLKITMQLFQRMDLHY[8]WTNTD[ 2]-vKsEcLGEATGDFYGQLDT[7]FDKSQLVSFRTPKDVLI    F  573
Query_PRIAP  563   TCCVYDEIVTSLMKIVGKLDLST   SKSTD     PsQ-VhDGSGGASSESP---[6]SDFADGLVPCRPKDFLV    F  630
Query_ROTIF  552   VEIIFNEYIESIIKIMNKLDLAA   VKNDG     SeVnGgEQENVDKNQSDLGV[1]SDPIIGLKPQKPRDFEI    F  618
Query_SPONG  559   HQLVYSEFIQSLLNILRKLDFTT   ASTTE     DgImAnEGEGGGGSGGGDPV[2]SDPVLGRQPAVPKDFQI    F  626
Query_TARDI  567   SKLIFDEFIKATIKMIDAMNLQL   LETES     --------EDDRVTVDLMPV   -----DKRPANDLDFKI    F  619
Query_TRICH  549   QRSLYNELLHSLQRILLKLDLTS   SQQTT     -----iEEETQSSE------[2]SNPVAGLKPNVPKDFQV    F  605
Query_OSTRE  566   TAAVYGSFVSELLNFCITLNLGI[2]---ST[ 2]StSlQqDDEVDAVAVAVSLG   ----DEANASNPGDMQT    F  628
Query_NITEN  568   RELVYDSLVQNMLTIVGGLDLRY[8]AAAPS[ 5]KtSeDvEPELALE-------   -----SIIPLNPNDMRM    F  634
Query_CLUBM  552   KTTIFDSLIRSIMKLMMELNLKY[7]---AN[ 5]G----qEPEAEAP-------   ------IEPLDKNQMLK    Y  609
Query_MOSS_  556   RESIYDAIMHCVIKVMNELNLKY[7]GKYSE[ 5]P----iEPEAEMD-------   -----LIEPLDMIHMRK    Y  617
Query_LIVER  563   QGLVFDTIVRSIMKALTELNLKY[8]TSAQE[ 5]L----vEPETELE-------   -----NIEPLNLVHMRK    Y  625
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  614    TNLVELSRD-LLPVTQT   QKFSS--WVFP   FGHQLILQ   SSKLPLVSGLYKMLTVCMQV--ASQISYFK-EM     676
Query_LINGU  620    INLVDFSRD-LHTDQYC   ALFKR--WVFT   FGHQLIVM   STKYPLVSGFYKLLAICMHA--TRKTAYFK-DI     682
Query_LIMUL  614    INIVDFTRD-LLLKKQW   KYFEK--WVYT   FGMEIIEH   ATRSPRVSGFYKLAAVCLKI--CSRIKYFK-NI     676
Query_SPIDE  626    INLVDFLRD-LLSHSCT   ELFSK--WVFV   FGKEMIFY   AVKYPFHSGFYKLISMNLSI--CSKIGYFK-TL     688
Query_SCORP  621    MNIVDFCKD-LFISKHN   ELFEK--WILI   FGLEIIQF   INKNPYISGFYKLITVCLTI--CKKTRYFQ-HV     683
Query_TICK   594    INLVDFFRD-TFTKKHL   ELFAK--WAFV   VVKQMVEY   SARHPLVSGFYKVVAASIGI--CEKSSYFE---     654
Query_MONOS  564    QQLVEFTTA-VLRQSRR[2]--FRN--WYFL   MAQRLVRR   ATQHPYVSGLYALLTTANEL--AEELGLLN---     624
Query_ROSET  536    IHLVTFGQD-LFSTTRT   DLFPR--WLHI   YYSDIVTR   SSKYPLVSSLYKLLELANTI--ATEQRYFEaAS     599
Query_BELCH  634    INVVDFCRD-LLPHHHP   EYFQR--WVFT   VGHALILH   STRLPLVSGFYKLLSVCMTI--CTKVKYFK-DL     696
Query_CIONA  624    VNLVDICWE-LL-PKAV   NHRRS--YFLA   HLPDLICF[4]SEKYPLVSGFYKLISTLMIL--SIKLLYFE---     687
Query_CAECIL 633    VNLVEFCRE-ILPEKHI   EFFEI--WVYT   FGYELILQ   STRLPLISGFYKLLSVAMKN--AKRTKYFE-GV     695
Query_HUMAN  624    INLVEFCRE-ILPEKQA   EFFEP--WVYS   FSYELILQ   STRLPLISGFYKLLSITVRN--AKKIKYFE-GV     686
Query_MOUSE  623    INLVEFCRE-ILPRKHV   GFFEP--WVYS   FAYELILQ   STRLPLISGFYKLLSIAVKN--ARKIKYFE-GI     685
Query_XENOPU 645    VNLVEFCSE-ILPKEHI   EYFES--WVYV   FGYELVLQ   STRLPLISGFYKLLSVVMKN--AKKSRYFE-GF     707
Query_PLATY  626    INLVEFCRD-ILPEKHV   QFFEP--WVYS   FGYELILQ   STRLPLISGFYKLLSIAIRN--AKKMKYFE-GV     688
Query_TASDE  629    INLVEFCRE-ILPEKHI   QYFEP--WIYS   FGYELILQ   STRLPLISGFYKLLAIAIRN--AKKLRYFE-GL     691
Query_CANAR  623    INLVEFCRE-ILPEKHV   EYFHP--WVYS   FGYELIVH   STRLPLISGFYKLLSVTMKI--AKRIKYFE-DI     685
Query_OSTRIC 594    INLVEFCRE-ILPEKHV   EYFNP--WVYS   FGYELIIH   STRLPLISGFYKLLSVAMKI--AKKIKYFE-GV     656
Query_TURTLE 632    VNLVEFCRE-ILPEKHV   EFFEP--WMYS   FGYELIVQ   STRLPLISGFYKLLSVAMKI--AKKIKYFE-GV     694
Query_GECKO  632    INLVEFCRE-ILPEKHI   EYFEP--WVYS   FGYELIIQ   STRSPFISGFYKLLSIAMKI--AKKIKYFE-GV     694
Query_DANIO  639    INLVDFCSE-LLPSRNP   EYFAQ--WMHP   LCHELILQ   SIRFPLVSGFYKLLSLSMGI--AKKTQYFQ-DV     701
Query_SNAKE  632    INLVEFCRE-ILPDKHI   KYFEP--WMYS   FGYELIIQ   ATHLPFISGFYKLLSIAMKI--AKKTKYFA-GV     694
Query_STERL  634    VNLVEFCSE-LLMKEHV   DFFEQ--WVYG   LGHELILQ   STRFPLVSGFYKLLSVSMKI--AKRIKYFQ-GV     696
Query_MILII  631    VNLVQFSTE-LLLERHE   EFFDQ--WVYR   FGYELILR   STQFPLVSGFYKLLSATMKI--AKKINYFK-GV     693
Query_BAMBO  635    FNLVEFCGE-LL-EKHT   DFFEH--WVYR   FGHELILR   STQFPLVSGFYKLLSVTMKI--ARKIKYFE-GV     696
Query_ALLIG  592    INLVEFCRE-ILPKKHL   EYFKQ--WVYS   FGYELIIQ   STRLPLISGFYKLLSVVMKI--VKKIKYFE-DV     654
Query_GAR__  631    INLVEFCSE-LLLKKHV   EFFEQ--WVYP   LGHEIILQ   STRFPLVSGFYKLLSVAMKI--AKKIKYFK-GV     693
Query_TETRA  620    YRVTLLFDS-VLPSIKP   ALFSN--YIYQ   FYKSIILK[1]-RKLPRSIVLLKLLRCLNEYi--EKNNILK-GL     682
Query_PARAM  572    -KIQEFRNFaVNPTQLT[1]--FHR--FILM[1]--SALIPK[1]----PIESAVISILNLLNLVl--IKIRKYPkAL[ 9] 639
Query_SLEMN  609    VRLSNFAEM-FLLQCKD   QWFLK--WIPQ   FTDTLIKK[1]-SQTPRISKLYQLLKVCMQIc--SNHRYFE---[ 1] 670
Query_STENT  557    FQASEFAQL-FLPKIGK   RLTM---WVMP   YINILSKK[1]-TELPFLPSLYIFIKVSFQIy--EVSNYSE---     616
Query_NEMAT  622    INLVDFCKA-LLPGQST   DSFGR--WVYV   FGRDLIVS   SSRFPVVSGFYKLLEVCMKI--CKRLRFFQyEV     685
Query_PISTI  633    ISLIEFCRC-VLPNYCV   TLFER--WIYM   FCRDIIIM   SSRFPVVSGFYKLLEVCMKM--CEKLSFFK-NI     695
Query_CORAL  621    INLVEFSRK-ILLKTKT   KWFEN--WIYT   FGKELVVL   SARFPLVSGFYKLLDIVMKI--SRKLDYFKgMT     684
Query_HYDRA  586    VNLVDLFGE-LLCKKSL   DSFSR--WVYI   FVEETLKF   SMKLPFVSGFYKLLRHSLWI--SEQIGYFE---     646
Query_LATIM  632    INLVEFCRE-LLPEKHV   EFFEP--WMYG   FGHEVILQ   STRLPLVSGFYKLLSVSMKI--ARKIKYFE-GV     694
Query_SHRIM  623    MNVVSLLEK-VLPLQ--[2]DQLEP--HIPR   LVWFFVTQ   SSLQPLVSAHYNATTTVLTC--SRQTQFFS---     683
Query_MCRAB  621    MNVIALLEE-VLPLQ--[2]DQLQG--HVSR   LCWFFIRC   SSEQPLVSAHYTSLTTILVC--AQKMRYFE---     681
Query_AMPHI  627    DSLVALMSD-VLEARCP[5]DNLIS--HSSE[8]SAFCVVAR   AIREPLVSQHYALLTVLLRY--LHAHQLWQ---     700
Query_BARNA  605    INLVSVLEHaLLSDGAE[2]--LRP--WLPL   LSGQVVAT   SVQRPLVSGLYRLARIYLLT--AGRLGCWQ-AV     668
Query_COPEP  603    VNLVDICVP-VLTSCQD   -LLKP--WLPR   LWYKISVW   AYEHQDVSSFYKFANKLLELd---PVKYIE---     661
Query_DAPHN  566    HNLVDLLEG-LLKAVEP[1]ERLQNs--LIP   LGEAVIGL   SMKFPLIPGLYRILATLLSIi-GQVLPIVQhPR     631
Query_STARF  627    VNLVDFCRD-LLPANQP   ALFKQ--WVYR   FGHEVILL   SCRLPLVSGFYKLLTVCMNI--CTQIQFFK-CI     689
Query_STRON  625    INLVDFTKE-LLPTRQI   HLFDP--WVYR   FGHELIQL   SSQLPLVSGFYKLLTICMTM--SRKTAFFK-EV     687
Query_SEACU  626    INLVDFCRD-LLPRQQM   EYFKP--WVYQ   FGHQLILL   STRFPLVSGLYKLLTACMVI--CQKTGYFK-QP     688
Query_FUNGI  632    LNFVEFTKL-ILSTRMT   GYMVHsdWMYL   LAKEIILK   SHQCPRISGFYKLFTTLLNA--SREHAYFStT-     696
Query_DPURP  681    LGLVEFFKL-FISKNHT   ELFVQ--WIYI   FGKEMINY   SKKYPLISGFYKLIQIVMKI--CKSQKYFNtIN[15] 759
Query_HALBU  698    LNLVEFSKL-FLARQHT   NLLTR--WIYI   FGKEIIFG   SRKYPLISGFYKLLHVLLKV--AKKVNYFGfI-[ 7] 767
Query_GLOBO  562    VTYVSLVES-LLPKIEV   MFFDK--WIYV   VGQKFISY   STQNPLVSGFYKIFGVVLAL--SDSLKYFS---     622
Query_PALUS  589    VTFSRFCEH-FLTRVRT   ERFSN--WIYI   VGNKWIEL   SLKNPMISGFYNLFSVILRL--SKSTGFFEgAL     652
Query_CALIF  662    INYVEFTKLiLLASSTR   ELFDK--WLYS   FGKEIISF   SGKYPLVSGFYKLFECCLII--SEDLQYFNnIL     726
Query_GLOMU  643    QNLVDFWQL-FLPRIKS   ELFSR--WVYL   AGDTLITF   STRYPYVSGFYKMLGSCLTV--CESIKFFDgIK     706
Query_LOBOS  628    QNLTEFWQV-FLPEVRP   DLFKS--WSYV   IGSTLIEL   SSKHPLVSGFYKMFATCLQV--CQSISLFQqKQ     691
Query_RACEM  600    HNIVGFWCA-LVPQLRG   DRLIQ--WIQI   AGSELIQL   SVRQPLVSGFYRMLSTIIAI--TDRTGCFQeSQ     663
Query_PVIOL  667    LNLVEFFKL-FFTTNHT   ELFVQ--WIYI   IGKDMIFN   SKRYPLISGFYKMIQVVMEV--TKKESYFYsI-[ 5] 734
Query_BOMBU  581    VNIVDLYVD-I--INSL[2]SLFIN--TIHK   FLLKVISK   SHKNHLVSGFYKLVHAISKHi-C------E---     636
Query_AEDES  598    FNLVDLYQD-L--LKYN[4]DNFEE--WIPI   YFDYVVRK   SVKHPLVSGFVKLIDLGLF-t-ANQLQYFQiNF[ 2] 665
Query_ANT__  588    ANLVDLYVD-M--INAT[2]SLFVN--TVHR   FLYEIVRL   SYKYPLISGFYKLIRVGMKIf-AHTLEEEE---     649
Query_WHITE  593    VNLMDLLKD-L-CPKL-[2]YLFLE--HLPR   FGRSIIDF   SVRFPLVSSFYKIIATALQ-i-ADEKGYFK---     653
Query_BEETL  588    LNLVDLYEK-I-FEYI-[2]NLLKK--CVYK   LIHYLIAK   CEKHQLISGFYKLLSFCLK-i-CKDINYFN---     648
Query_THRIP  568    VNLVNFTID-L-LSSLP[2]SLLAN--WFTP   LILYCIKL   SSKHPLVSSFYKLLANIVR-i-AENLDYFE---     629
Query_TERMI  600    INLVDLTRD-V-LMSV-[2]SQFEK--WIQP   FFYKVISE   SSNKPLVSGFYKLLTVALQ-l-CDRLHYFS---     660
Query_FLEA_  575    INM--------------   -------WIPT   YLKEVINC   SYKFPIASGFYKLSSYGLD-f-CDICEYFD---     617
Query_MIDGE  603    YNIVHFYSD-LILLQSN[4]KNFLQ--WIEM   WLENSIKL   SFKHPLVSGFLQLIETALK-v-INRLN------     664
Query_LAODE  596    MNLEFLFSK-LLPKIEN   RLHEK--YISL   ----LLRN[4]STLNPLVPSYYKLLATVLSI--AEQLGFFEaDN[ 2] 661
Query_LOTTI  615    INLVDFCRE-LLPNKHI   NLFEN--WIFT   FSHNIILL   STQNPLVSGFYKLLALCMKI--SNKLAYFQ-VN     677
Query_APLYS  614    INLVDFCKD-VLVTSHW   NLFED--WVYT   FAHTLVVM   STRHPLVSGFYRLLSVTMTL--ADKLSYFK-AF     676
Query_SCALL  624    INLVDFCRD-LLPTVHF   ELFEK--WMFT   FCHKVILC   STDNPLVSGFYKLLAMSMKI--SNKLNFFK-VM     686
Query_CRASS  615    INLVDFCRE-MLPEKHY   KMFGK--WLFL   FSHTLILQ   STEKPLVSGYYKLLAVSMKI--ANKLNYFQ-GM     677
Query_OCTOP  611    YNIVNFCCE-FLNKNEK   KMFRK--WLFR   YCNFLVKI   STSTPLVSGFHRLLTVCMKI--ARDLHYFE-GI     673
Query_POMAC  611    INLVDFSRD-LLVPKHF   QYFEK--WVFT   FTHTVVVQ   STEFPLVSGFYKLLAVTMKI--ANKLEYFK-DI     673
Query_WHIPW  609    FNLVDFCRF-IIGPVTK   SLFEP--WVLP[1]--FSVLVK[1]TIENSNVSGFYKLLSYCIAA--AEKLSLFD--V     670
Query_TRICH  607    YNLVELCRS-VIMNGSH   VLFEP--WLQP[1]--LTMLLE[1]SSRYPMVSGFYKLLAVWVTM--AEKLNYFN--Y     668
Query_AASTA  596    LNATEFFER-WVPTV-T   SQFQS--WVAL   YVQWLLPS   ARQLPLVSAFYRLATVISHLl--A-----A---[ 1] 651
Query_AEUTE  588    LNLTEFCER-WLPSV-Q   SSIQP--WIPL   FLKWIFSR   SRESPLVSVLYRLATVVSRIl--A-----I---[ 1] 643
Query_PLASM  620    LNLTEFVER-ILSQVLS   SQLCP--WIPL   IMNQIFAS   ICDLPLVSCFYRIGTVMTKAm--HELNYF----[ 1] 680
Query_BREMI  619    LNLTAFVER-VLGKMPN   SLLQP--WIPL   IVEQIISL   AAKAPLVSCFYRIGTVVAAAm--NGLVYF----[ 1] 679
Query_OLIGA  618    LNMTEFIER-VLPRLSP   ESLRA--WIPI   LLMLVIEK   SDELPLVSGFYRIVRVLTTVs--EKTKYFD---[ 1] 679
Query_LIGNO  630    VNLVDFVEE-LLQHHRA[2]--IAK--WFYP   LCQALIEK   SAENPLISGFYKLLATFVTR--ARDAGYFNdQP     693
Query_CESTO  523    SNIAAFLEH-ILPDCTI   ENFKEs-LLFF[2]MCLELIVK   ---YPLLSGLYRLLKLSVLL--AKRSGFFR---     583
Query_OVI2_  574    VNLVDLYEA--LFSQET[1]EVFKHc---LP   ---ELIQT[4]IRRYPLLTGFYKLLNLLVTSg--VRSGFLL---     634
Query_PRIAP  631    FNLVEFTKD-LLSTNQT   ALFRR--WIYT   YGHEVILL   STRWPVVSGLYKLLTVCLSI--CSSLQYFK-AM     693
Query_ROTIF  619    INLVDFTRE-LLMNKNF   DLFEK--WLFK   LSKEIILL   STSHCLISGFYKLATLIMKM--AVKIGYFN---     679
Query_SPONG  627    INLVELAKK-VIPSVSL   SLFTP--WILT   FSQEVVSL   ATRYPLVSGFYKLLTVCLSI--CNKTHFFD--V     688
Query_TARDI  620    FSLEQIVVD-TMSRMNV   DLVAS---QEQ   FCLQIMKL   SEKNPLVSGFYRLLEVWFRRlrSSKPEHDKmAL     684
Query_TRICH  606    INIVDFTKA-LLPTTRT   EFFLP--WIFP   FGKNLVLL   STRLPLVSGFYKLLEICMTI--SSSLSYFE-DL     668
Query_OSTRE  629    LQLVEFSQR-VIDESPP   DSLLP--WVDH   LSDELMRM   ANQNSSLSGFYKLIGAIVSA--ANRGGYFS-NL     691
Query_NITEN  635    LTLVEFAQE-FLQKCCV   SFFPK--WLYT   FGRDVIAL   SDQYPLISGFYKLLAVAIRV--GDALGYFAaES[ 9] 707
Query_CLUBM  610    LNLVDFVLR-LFASDHA   SLMEK--WIFP   FCRCLIDL   SRKYPLISGFYKLLAAVISI--GDQKGYFE---     670
Query_MOSS_  618    LNLVDFVQE-LLSGPCS   SLILK--WTYP   FGRCIIEL   SNRYPLISGFYKLLTVNITI--ANREGFLS---[ 6] 684
Query_LIVER  626    LNLVEFTQD-FLSRPYA   HLFMK--WVHP   FSCCVMEL   SHQYPLISGNYKLLTITITG--ADSIGFFQsKD[11] 700
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  677   NRQLKIET[ 8]DEEF[ 4]AQEEACFMLFTKFAKEVMT   RLRQYKDDLLASCLGFILALPRE--IIALD    VTT-  750
Query_LINGU  683   SLSRDQDG[ 7]EEKI[ 2]-DRETAFLLFSKFTKEVLV   RIKQYKDDLLASCLSFILALPPE--LVIAD    LPT-  752
Query_LIMUL  677   N-TLLRPK[ 7]DDPF[ 2]-ERILSFQLFSKFAKDAVI   SLQQYKDDLLASCLQLVLSLPSE--IVLSE    ISH-  745
Query_SPIDE  689   S-SN---T[ 7]SNDP[ 2]-LVPQSYTLFSKFVKEIAV   QHTQFRDDLLASCLQVMLVLPME--IVAKE    IHS-  754
Query_SCORP  684   K-KL-KTK[ 6]----[ 2]-DCLIAFELFAKFSREVVI   YLQQFKDELLASCIEVILSLPTE--IVESE    INI-  746
Query_TICK_  655   -----QPS    ----[ 1]-EAVACLRLVSGLLRDVLG   RLQQLKDDLLASCLELVLGTPLA--LVQEH    ARA-  707
Query_MONOS  625   ---EHQKL    -ESF[ 5]-------TLMGAFCATVMR   QCKQFKDELLTTCLRFLVTCPTV--VVKSR[ 1]--Ev  680
Query_ROSET  600   -----KTA    -ETA[ 5]-EVAATLELMSKFGKEVMV   RSKQYKDELLSACLTFVLSLPLP--VVKRD[ 1]--Av  659
Query_BELCH  697   DTKL----[10]EGRD[ 2]SEGQACFTLFRKFVKEVLV   RLKQYKDDLLASCLLLVLSLPHQ--IVQSE    IST-  766
Query_CIONA  688   --DEHEVA    EESM[10]NQQNNVKCIVQKYATQVAA   RCKQYKDELLTACLYFLLSLPND--VCKDI    LPQ-  757
Query_CAECIL 696   SPKSHKKC    PKD-    PERFACFALFAKFGKEVSA   KMKQYKDELLASCLIFILSLPHD--IIMLD    IKA-  756
Query_HUMAN  687   SPKSLKHS    PED-    PEKYSCFALFVKFGKEVAV   KMKQYKDELLASCLTFLLSLPHN--IIELD    VRA-  747
Query_MOUSE  686   SPKSLKHS    PED-    TEKYSCFALFAKFGKEVSV   KMKQYKDELLASCLTFVLSLPHD--IIELD    VRA-  746
Query_XENOPU 708   TSKIYKKA    PED-    PERLSCFALFAKFGKEVSS   KIRQFKDELLASCLTFVLHLPHD--IIMMD    IKA-  768
Query_PLATY  689   SPRSLKKS    PED-    PEKYSCFALFAKFGKEVAV   KMKQYKDELLASCLTFVLCLPHD--IIMLD    IKA-  749
Query_TASDE  692   SPKSHKKS    PED-    PEKYSCFALFAKFGKEVSI   KMKQYKDELLASCLNFVLCLPHD--IIELD    IKA-  752
Query_CANAR  686   SPRSLRRC    PED-    PEKSSCFALFVKFGKEVAA   KMKQYKDELLASCLIFLLSLPHD--IIMLD    IKA-  746
Query_OSTRIC 657   SPKSLRKC    PED-    PEKSSCFALFLKFGKEVTA   KMKQYKDELLASCLSFLLSLPPD--IIMLD    IKA-  717
Query_TURTLE 695   SPKSYQKC    PDD-    PEKSSCFALFAKFGKEVTA   KMKQYKDELLASCLTFVLSLPHD--IIMLD    IKA-  755
Query_GECKO  695   APKSKKKC    PEN-    PEKSSCFALFAKFGKEVTA   KMKQYKDELLASCLIFVLSLPHD--IIMLD    IKA-  755
Query_DANIO  702   KQCPKQVG    GST-    MEN-ACFSLLAKFGKEVC-[1]RMKQYKDELLAACLMFILSLHPG--MVALD    IKA-  761
Query_SNAKE  695   SPESTKKY    PED-    PERSSCFALFSKFGKEVTT   KMKQYKDELLASCLTFVLSLPYN--IIMLD    IKA-  755
Query_STERL  697   SPRSCKTS    PED-    PDKYACFALFAKFGKEVST[8]RMKQYKDELLASCLTFVLSLPHD--IVALD    IKA-  765
Query_MILII  694   STNTWKQC    HTD-    PEKCACFALFGKFGKEVTI   RMKQYKDELLVSCLNFVLSLPHD--IIELD    IKA-  754
Query_BAMBO  697   NPKTRKQC    PKN-    PEKFACFALFGKFGKEVAI   RMKQYKDELLASCLTFVLSLPYD--IIELD    IKT-  757
Query_ALLIG  655   SPGTQKKY    PED-    PEKSSCFALFAKFGKEVAA   RMKQYKDELLASCLTFVLSLPHD--IIMLD    IKA-  715
Query_GAR__  694   SPKSFQPN    PQD-    PEKYACFALFAKFGKEVS-[1]RMKQYKDELLASCLTFVLSLHHD--IVALD    IKA-  754
Query_TETRA  683   -----KKY[ 1]KKDI[ 1]-------SQIKIFFSELYK[1]-LGKYQDELLIDCLKTMLSIPFE--VLLID[ 4]FRKn  739
Query_PARAM  640   -LKYIKET[ 3]NQQF[ 3]RQKQILINLILKLKGLIQN[1]-LKSYSSQLQIEALNVLLTIPVE-fIVKDN[ 3]LKSe  711
Query_SLEMN  671   SDDIDSVA    ---K[ 1]GEKQNTYNMLLTFFKELIG[1]-QEEYQDELLTSCLDLLLHVPVQ--ILYRQ[ 9]----  736
Query_STENT  617   -SDLQSKS    ---T[ 1]G------NLLKFIFK----[1]-MQSFQDALLSSCIEVCLAAPIC--IVY--[ 6]L-Sl  669
Query_NEMAT  686   TAD-----    ----[ 7]SDCTNCFILFRKFIKEVLI   KMKQFKEDLLSSCLSLVLSLPYE--IVAVD    VAA-  745
Query_PISTI  696   KAEKPDLK    NENL[ 7]SDCMKCFVLFHKFTKEIVV   RMQQYKDDLLFSCLLVVLSLPLE--LVTLD    LVS-  764
Query_CORAL  685   LKDVSTGC    VDDM[ 7]SDCRNCFILFHKFAMEVVV   RMKQYKDDLKSSCLQLVLSLPNE--LVLCD    IDI-  753
Query_HYDRA  647   --ANN---    ----[ 2]DEKARLFKMLSAFVKEVSA   RVKQFKEDLLVSAVMFLLSLPVS--VIKES    KKF-  701
Query_LATIM  695   SSKSCKQN    SLD-    PEKYACFALFAKFGKEVCA   RMKQYKDELLASCLGFILSLPHT--IVELD    VKA-  755
Query_SHRIM  684   -----RKQ    ----[ 1]PEVETSIHLLMSYVRELLQ   RCRQYRDQLLASCLTLILHLPLC--IVQEI    FPQ-  737
Query_MCRAB  682   -----KTK    ----[ 1]PEVETLIHLITSYVKDVLQ   RCRQLRDQLLASCLTLILKLPVC--IVIEI    FPR-  735
Query_AMPHI  701   -----RQQ    ----[ 1]VASRNLGELVVSFASDCCE   TSKQFSDHLLASCLAFLLALPKP--LARQI    LPQ-  754
Query_BARNA  669   -----QEN    -SDS[ 7]--------LWAAHVSETAP   RCSQYRDELQEACVSMVLAAPAA---LAAD[ 1]-AA-  722
Query_COPEP  662   --------    -PDT[ 1]-----LFTLILQFLQEHLQ[1]SLSFDQVELRISAFNLVTNIPLK--FVPGL[ 1]--S-  710
Query_DAPHN  632   --------    ----    ---------YQSYMTSLIN   RLSQLSDDVLVSAVQAVLSSPTW-lVISLS    -SS-  672
Query_STARF  690   KP-----S[11]DLNF[ 9]SDRCICFTLFAKFSKEVLV   RLKQYKDDLLASCLLLVLSLPHE--IVQAE    VAH-  766
Query_STRON  688   NPNVTFRT[10]DDDS[14]SDKQACFILFSKFSKEVLI   RLKQYKDDLLASCLFLILSLPHQ--VVENQ    VHD-  773
Query_SEACU  689   GK-----A[11]DPH-[ 6]GDRQACFTLFAKFAKEVTV   RLTQYKDDLLASCLFMILALPLE--IVQQE    VNN-  761
Query_FUNGI  697   --------[ 6]----[ 5]GKIRMCSVLFSRYIKEVIG   RIDQFKDELLSSCIQLVLSSPHQ--VV--P    ISQ-  755
Query_DPURP  760   ELENDLM-[ 5]----[ 9]DK-KNCFILFKKYIKEVSN   FISHYKDELLSSCIELLLSLPKE--LI--S    IPL-  827
Query_HALBU  768   SLDANGG-[ 7]DSGK[ 9]NK-RNCFKLFAKFLKEVEA   ASTQYKDELLASVIQLLLSVPKQ--LV--N    IPV-  841
Query_GLOBO  623   VPTRNE--[ 8]EESS[ 1]-VTKAARSLFSKYLDEVLL   KMRQFKDDLLVSCLKMVLYSPVE--LT--T    IES-  688
Query_PALUS  653   APDRKEIY[ 8]EQAT[ 1]-EQRRTYILFFNYIREVLV   RLQGYKDDLLASCLRVVLACPSE--LI--S    VQD-  720
Query_CALIF  727   KKEYIGES[ 8]NEPY[ 1]GIKQRTYILYTKYINEIIS   RMTQYKDDLLVSCLKLVLSSPNE--LI--S    LND-  795
Query_GLOMU  707   NVDNENIM[ 8]STAA[ 1]PTQRASYFLFIKFFDEVSA   RLPQYKDDLLASCLRLVLSIPRE--LI--N    IQN-  775
Query_LOBOS  692   SL-MERIK[ 8]HRAA[ 1]--------LFKKYIYEVLA   RLEQYKDDLLASCLHLVLSSPSA--LI--E    PSN-  751
Query_RACEM  664   KLYHEQ--[ 7]EPTS[ 1]PDLYVTYLSFRDFLSETRY   RLQQFKDELLAACLQLVLTAPLT--FF--D    IQQ-  729
Query_PVIOL  735   DLDENNN-[ 5]----[ 9]NK-KNCFILFGKFIKEVVN   NVSHFKDELLSSCLELLLSLPKQ--LI--N    VSQ-  802
Query_BOMBU  637   ----LSSD    -EIE[ 4]-E------MLNKYIMNVLN   LIPTFSNELLTTCVHLILDVPML--YVENI    LNS-  690
Query_AEDES  666   ----LQRQ    -QNF    -STMSTTNLLVYYLE-LQI[1]KSRNSTGELQLTCLRFVLGAPVT--LLDSF    LDD-  721
Query_ANT__  650   ----KGEE    -EEE[ 4]-ESQQMEEILSNYLWHSLD   LIPGFSNELLIACLYLILDAPFV--HVKNL    LPR-  709
Query_WHITE  654   -----TET    -TST[ 2]-EAENFRLLVQSYLELVLN   RALQFTDELQIACLHVIMAAPTT--FIDPI    LPL-  710
Query_BEETL  649   -----SSC    -EDA    ---TMCYLTLRKFLTSLLN   KTQQFTGDLLIACLQVLLQYPIV--IIKDM    LST-  701
Query_THRIP  630   -----QEK    -DEM[ 3]-EWRSTVNLLSAFLYDVIS   RQAQYRSDLQVSCLRLMLATPSC--IALPL    LRV-  687
Query_TERMI  661   -----DYG    -DML[ 4]-EQKASYSLLCRFLQELLI   RLQQYKNDLQVACLETVIAAPVA--LVVPL    LPH-  719
Query_FLEA_  618   -----HYE    -ESE    -EKSSCYVLVLNFLQECAV   NVRMFNGDLQESCLRVLLTAPVI--FIPEL    LVE-  672
Query_MIDGE  665   ----VTNA    -TDF    -DMTKIEESLSFYIKSMII[1]RCQLLSSELQIATLRLVFQVPID--ILIEF    ADE-  721
Query_LAODE  662   --------    -EDS[ 2]-----CQKEINYYIGDLLV[1]-LHHYHKDVQEACLRVITSSSVE--LIASH    LQQ-  711
Query_LOTTI  678   ESNEQE--    ----    KQKAVCYLLIKKFSKEVLV   RMKQYQDDLLAACLTLILSLPKE--IITEQ    MID-  733
Query_APLYS  677   DSQSQK--[10]GEGG[ 6]QEVEVTFHLIKKFSKEVLV   RMKQYRNELLASCLTLILSLPKE--IIVDQ    MSS-  752
Query_SCALL  687   LYLDVMGT[12]DQQV[ 4]QQRDSSYLLMKKFSKEVLV   RMKQYKDDLLASCLTFILALPKE--IVAQQ    MVE-  764
Query_CRASS  678   VSL-----[11]DSDQ[ 4]HQRATCFQLIQKFTKEVLV   RMKQYKDDLLAACLTFILTLPKE--IVSEQ    MSS-  749
Query_OCTOP  674   SIDNLKIT[ 6]----[ 2]-EKIACCILIKKYSKEILV   QLKQYKEELLAACLALVLSLPTE--LIVDQ    IDD-  738
Query_POMAC  674   AVENSF--[ 8]DETG[ 4]TEKHLTYILIKKFSEEVLV   RMKKFHDDLLASCLTLILTLPKE--IIAQQ    MSK-  745
Query_WHIPW  671   SISENQVG    ----    -DRATCRILLSSLVDEIFP   RSHSFSGELSVSFLRFVVCVPSA--TALEK    RDL-  727
Query_TRICH  669   TEQDKTEN    ----    -CNSHYKLLLSDFIDHVLV   ESCRFQGELLVSCLRFLLALPTC--IVVEK    LES-  725
Query_AASTA  652   SAASRSPL[ 1]SDAL[ 3]---------YAAFVVQVSK   AMPQYHDELLIATSVFVVATPVS--MV--D    LAI-  706
Query_AEUTE  644   VS------[ 1]PPAV[ 3]---------YIDFILGVVN   AMKAYSNELLVSTSIFVLATPLS--MV--D    VKM-  692
Query_PLASM  681   SASSVDTT[ 1]QTRF[ 3]---------FTSCVQQVSS   QARFYQDELRLTSSEFVLTAPLG--LV--A    IHA-  735
Query_BREMI  680   SPATLTTT[ 1]Q--F[ 3]---------FTSFIEQVTA   QAQFYQDELLLRCSEFILTAPGE--LV--T    IRS-  732
Query_OLIGA  680   MRDTTDGG[ 1]DVRV[ 3]---------YFIFAKRVIQ   GIRFYKDELLVASSGCVLAAPLG--LV--P    LVD-  734
Query_LIGNO  694   VPVVMEIA[12]QTDN[10]ENRIQCCAMLGSFAASLSA   RCGQLRGDLLAASLRFLLSLPAS--LLAVD    RPDd  778
Query_CESTO  584   ---SNELA[ 1]ETQL[ 3]-----AHFLINSFPPTGAL[1]-----ADDLSIARSSCLLSLLPS--VLFHS[ 4]IGH-  641
Query_OVI2_  635   -----RTG    DEGS[ 7]EARFETQRRLAEFSSDLCV[7]-DRTEADQRIASRLLCLLSLPGAtfIAPIG[16]FASh  723
Query_PRIAP  694   RKVN----[11]DSPD[ 4]AEILNCFTLFKKFSKEVLV   KLKMFRGDLLVACLQLVMALPYE--IVVDI    VPE-  766
Query_ROTIF  680   KIDSIQTN[11]IEDC[14]YERVNCFHLFKKYLLEVSI   RMKQFKDELLVSCLEFFLSLPKE--FVIIS    MDH-  766
Query_SPONG  689   HDLSGQGT[ 1]DDDV[ 7]TDRQSCYFLFSSFIKEVVV   QMQQYKDDLLASCLHLILVLPKE--MVVSD    FKE-  758
Query_TARDI  685   KDRQTDSL[12]DEAE[14]EPSSANHVTISFFMRNLLG   RLQHFSDEVLISALQLLLNFPIY--AVIDN    LSA-  772
Query_TRICH  669   NIETVTKC    TINA[ 6]IDKYLCYLLYYKFYKEVLV   RLKQYKDDLHVSCLSLILALPYQ--LIQDD    MAL-  736
Query_OSTRE  692   -----KSA    ----    -ESERVLSLFQSFLRDVLQ[1]-TRRFSGELLAAVLRLLLSLPSV--ILPMA    ---d  741
Query_NITEN  708   DVPMSEAS[12]CTEL[31]GDREQCTELFRKFLREVLV   ASRRYKLEHLAACLQMCLSAPTV--LL--E    IRQ-  810
Query_CLUBM  671   --------[ 6]V---    --------MIQEFLHEVLI   TSRRYKLELLGSCLHLCLSCPPT--LL--D    LAS-  717
Query_MOSS_  685   -LDADGGY[12]AGTT[ 4]EDRRLCKELFRKYLQEVLV   ASRRYKLELLAACLRLCLSAPPL--LV--G    IAS-  759
Query_LIVER  701   MVTEEEWK[12]QRGL[ 4]EHQRLCQQLFRHYLQEVLV   ASRRYKLELLASCLRLCLSAPPV--LI--D    IGS-  776
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  751    FVPALQRS---LEIGRSY--LPLARTSLNTLESWSKSLPP   DVMKPF   YEQLLPYLDAYLT-S-----SADEKN--   811
Query_LINGU  753    IVPALKIT---FKLGLSY--QPLAVEGMEAIEFWRDSIPN   HLLQPH   YQDILPCLDDFLK-LQSEQAGNETNS--   818
Query_LIMUL  746    LVPSLECT---FQLGASY--LPLAIAGLEALEVWSTKLPP   EVLEPY   YPQILPHLNGYLVgKSTNDLDESAGV--   812
Query_SPIDE  755    LIPALESA---LKLGLSY--LPLAEKAVNALEKWSTNLEL   NILKPY   FPRLLPLLNSYLLvTAAEAGSVLSNE--   821
Query_SCORP  747    LIPAVQIS---LQLGLSY--LPLAMSCLDALETWLYSLPS   ELLKEH   YCKLLPSLNDYFLaSDNISLQPLKNI--   813
Query_TICK_  708    LVPAVRMA---LQQGLSY--LPLAKASLDALRLWEKQLPA   DVADDV   FRQVLPGLCPYLA-----LGVRAGIT--   769
Query_MONOS  681    FIEALDTS---LRLGIGY--LPLAEAALASLQRWLDVLGL[2]--LQPH   LAPLLRRFDQYLT--AQASATDDLNAd-   746
Query_ROSET  660    FFPAVETA---LSLGLGFy--PLADAGLSALETWLPPALA[2]---RPF   -AQLMPLLDKYLS-AAVAASDIDVEE--   723
Query_BELCH  767    LVPALQTT---FRLGLSY--LPLAHAGLDALERWLLNLPR   SALEPH   FPHILPYLDSYLQtAADTGSSSSSDG--   833
Query_CIONA  758    LIPAVQRA---CQLGLSY--SPLSEALMSALLRWRSIFIS[1]---RKC   YLDIVGCLQPYFN------LSEETTN--   816
Query_CAECIL 757    YIPALQTA---FKLGLSY--SPLADAGLDALEDWSAYIPQ   YVIQPH   YSEILPYLDGYLK-TASSA-DEYQNS--   821
Query_HUMAN  748    YVPALQMA---FKLGLSY--TPLAEVGLNALEEWSIYIDR   HVMQPY   YKDILPCLDGYLK-TSALS-DETKNN--   812
Query_MOUSE  747    YVPALQMA---FKLGLSH--MPLAEIGLHALKEWSVHIDK   SILQPY   YKDILPCLDGYLN-TSTLS-DETKSH--   811
Query_XENOPU 769    YIPALQTA---FKLGLSC--PPLADVGLNALQYWSTNIPS   DILKPY   YKDIIPLLDGYLT-NLSST-NESLST--   833
Query_PLATY  750    YIPALQMA---FKLGLSY--PPLADVGLNALEEWSIYIPR   HVIQPY   YKDILPCLDSYLK-TTALS-DETKNN--   814
Query_TASDE  753    YVPALQMA---FKLGLSY--TPLAEVGLNALEEWSLYIPK   HTIQPY   YKDILPCLDSYLK-TTALS-DDNKDN--   817
Query_CANAR  747    YIPALQNA---FKLGLSC--TPMADLGLDALEDWSAHIPR   HMMQPY   YKDVLPLLDGYLK-SSTST-VESQNN--   811
Query_OSTRI  718    YIPALQNA---FKLGLSC--IPMADLGLDALEDWSAHIPK   HIMQPY   YKDVLPLLDGYLK-NSAST-VEPQDN--   782
Query_TURTL  756    YIPALQNA---FKLGLSY--TPMADEGLDALEDWSAYIPK   HIMQPY   YKDVLPLFDGYLK-TTGSA-DESQNN--   820
Query_GECKO  756    YIPALQNA---FKLGLSY--TPMAEVALDALEDWSAHIPR   YTMQPY   YKDILPCLDGYLK-TSTST-DDSQNN--   820
Query_DANIO  762    YIPALQAA---LRLGLSH--APLATAALDALESWSSFIPA   AILQPH   YTDILPHLDGYLK-TTSSS-EKDDSN--   826
Query_SNAKE  756    YIPALQNA---FKLGLSY--TPMADVALDALEDWSSYIPR   NIIQPY   YKDIVPCLDSYLK-TSAST-DESQNN--   820
Query_STERL  766    YIPALQNS---FKLGLSH--TPLANTGLYALENWSQFIPK   YVMQPH   YKDILPYLDGFLK-TTASN-DEEQNN--   830
Query_MILII  755    YIPALQAA---FKLGLSH--TPLADAGLDALEDWSAHIPK   HIIQPH   YKDILPHLDGYLK-TA-AF-NEAQNN--   818
Query_BAMBO  758    FIPALQTA---FKLGLSH--IPLADAGLDALEHWSTRISK   HVIEPY   YKDILPHLDGYLK-TA-AS-NEFESN--   821
Query_ALLIG  716    YVPALQNA---FKLGLSY--TPMADVALSALEDWSSHIPK   NIMQPY   YKDVLPLLDGYLK-TATSP-DEFQNN--   780
Query_GAR__  755    YVPALQTA---LKLGLSH--TPLASAALDALEEWSSRISQ   PLIQPL   YRDVLPCLDGYLK-TAASN-GSDQNT--   819
Query_TETRA  740    LIVALKKG---FELGVQN--VSMGFAAIECLQKIIKRSKN[1]-VVFKY   LKLVIPYLSEYLSiTKGELEEITSDMk-   807
Query_PARAM  712    YIQIIINA---FQMNSNS----LSWSAIDLLIQLIQSEYN[3]QKNLKY[4]LKNVIPKLSSYMQ-----LVPEQEKPfy   780
Query_SLEMN  737    WKVIMQKA---IDLGRNN--NQLAWNSINMLEQWF-----   -NLLPM[6]YHDILPKLSDFLQ------VDDNKKQ--   797
Query_STENT  670    FEPIILQA---LTLGLSH--LPLAHSVISVLNAWR-----   -ASIPI[6]LPKILPALSPYLT------LN-------   725
Query_NEMAT  746    IVPALELT---FKLGLSY--LPLAQAGLAALEEWNASLPG   DVLRRH   LRNILPYLDGYLK-SATDAVEEEQDQ--   811
Query_PISTI  765    SVAALKLT---FKLGLSY--TPLAEAGLDALESWTTALPA   DKLKPF   LREILQCLDGYLT-TLGDI--QLHNE--   828
Query_CORAL  754    IIPTLKTA---FKLGLGF--LPLAEAALDALDTWTAELPS   EILHPY   LREVLPQLDGYLK-TASEL--ETSNE--   817
Query_HYDRA  702    IIPIFQVA---LQLGLSY--YEIAELCLDTLDRLQKKL--[1]---ESD   MESIVRLVLPHFNdYIQPIFKEEDSY--   764
Query_LATIM  756    YIPALQMA---FKLGLSY--TPLADSGLDALEEWSNRIPK   WVLQPY   YKEILPCLDGYLK-TSISA-GDTPSS--   820
Query_SHRIM  738    LVSPLQMA---LQLGVSY--LPLAEVCILALQLWTKSLSQ   EILSEH   MPHVLPLLLPYLR-SQ-ETGGEAEVH--   802
Query_MCRAB  736    LIVPLQMA---LHLGVSY--LPLAEVCIAALCKWTHILPS   DILHQH   LPAVLPLLLPYLR-SK-ETG-EVEVQ--   799
Query_AMPHI  755    LSQPMHTA---LAMGLSF--LPVALSAVACLQRLVDDLAF   SQDEKY[7]VQSVLPDLLTHVQ-SP-QRE-RSELQ--   825
Query_BARNA  723    --PLLAVArlaLRLGRSH--LPLAARALDALESWLPALAG[6]DCVEPL   -RAMAPALSQYLApEATVAAGGDRAP--   795
Query_COPEP  711    LLPDAWNQi--FKLSSSIp--ELANSALKSIHGYLETFPE[2]--LELI   LEKVLPNLRVYLS------YQHQTSE--   772
Query_DAPHN  673    ITGALASA---LKLGLTYp--PLAEWPLDALEHWQTTLPF[2]-----F[3]MEEVLPCLEGYLS-----DVEELAKE--   734
Query_STARF  767    LVPALQMT---FQLGLSY--LPLAETGLEALERWMLHLLP   DSLQPH   LKVILPLLDGYLR-GTITQGSDAADV--   832
Query_STRON  774    VIPALKTA---FKLGLSY--LPLAESGLSALQAWIRYIPS   DVLKPY   YAQVLPQLDGYLG-AG--SGSEDEGS--   837
Query_SEACU  762    LVPALQAT---FKLGLSY--LPLAEAGISSLQMWTNNLPI   DTLTKP   LKKVLPLLDGYLR-A---RSSEDAAP--   824
Query_FUNGI  756    LIPAMRAA---FRFGLNY--IPLAKIGIDTMKRWISDDYE[2]-KV--H   LSAILPLMEAYLL-INVDTSTSSEQ---   819
Query_DPURP  828    LVPIANLA---FQYGKSY--LPLAHIALNAIEYWITVVPK[2]-HS--H   LDKLLPSLNDYLK-ISTNGL-DSVN---   890
Query_HALBU  842    LIPTLSTA---FRLGLSY--LKLGHVGLNTIEYWLSVLPE[2]-DA--H   LYAILPALSDYFL-ISTTNTSSSNQ---   905
Query_GLOBO  689    LCLPIIES---LKMGLSF--TPLASVALDAIAVWSSKLDR[2]-LASSY   AK-ILPSLNDYLL----LDLE-------   746
Query_PALUS  721    TVQPLQIA---LRLGLSF--PPLANIALETLERWIADFSA[2]-LRPAY   KG-VLPLLADYLV----LDVE-------   778
Query_CALIF  796    FIEPLKII---LKLGIGE--YSLAKLGLDAFEKCISYIPK[2]-LKKYY   KD-IFPLFNNYFI----NGEE-------   853
Query_GLOMU  776    LISPICIA---LKMGLSY--QPLATIGINAIEKLVDL-QG[2]-DISKW   LEQILPLFNEYLL----VKAD-SVDNsi   839
Query_LOBOS  752    IIPPIKLA---LKLGLGY--FPLASIAMEAIERWVDIIDH[2]-EI--W   LSQVLPYLSDYLM----VELTVNGDNdl   815
Query_RACEM  730    VVQPLQTA---LELGLTY--HPLATSALDALDQITAPDAP[2]-ANLCF   LGSVLPYLNEYLM----DRQE------i   789
Query_PVIOL  803    LVPIVCLA---FQYGLGY--LKLGHVGLNAIEYWLKVVPQ[2]-HN--N   LHLILPHLNDYLL-MSSAAVSNSSG---   866
Query_BOMBU  691    TIPTFKIA---FTVGISD--FELACKALNTLEKWTN----[4]EITAMF   LQEIVPFLEPYLH-SGESS---------   749
Query_AEDES  722    SKDLIEILcsaFQLGKGM--LSLANAALSCVRRLVAEGSP[3]ENRDRV   LVHVLPLLDSYLQ-TRDVALPPPISS--   793
Query_ANT__  710    TLPVFKMA---FTIGLSN--LDLAYTALTTLETWTTTMTT[6]EQTNEL   VREIIVYLEPYLR-STESS---------   774
Query_WHITE  711    ISSFLKVA---FRVGSSY--LPLADLALTTLERWIDAFPP   EKINPF   LKEILPCLDSFLR-SKDFEGDPEAAN--   776
Query_BEETL  702    CTPSFRTI---FHVGRTS--TKLTYLTLETLECWQNEIDP   DEFEPF   LKQILPLFDSYLQ-SKS-----------   758
Query_THRIP  688    CVPAFQSV---FQIGCSM--LSLAHEGLHTLSRWNQDLPG   EKMRIL   LKAVLPSFDSLLQ-TKDSVPVDEDDG--   753
Query_TERMI  720    AVPAFEVI---FRVGRSM--LPLARLGLSALDKWYCTLPS   HLVTPL   LMQVLPYLDPFLR-SQGVA--SGQDV--   783
Query_FLEA_  673    ASISLQII---FKLGYNH--LTLASLALETLERWTDHLSK   EQNTIL   LDEVLSCLEPYLQ-SQGLIGDPDVKV--   738
Query_MIDGE  722    ---FIPIFvigFNVGRSL--LSIANQALTCFDKMLDSLSD[2]KTREKL   LQEVLPCLESFLS-TRD-------SG--   782
Query_LAODE  712    -LPAAFEVl--FHIGRSV--LELSSMGVSALDRWMSKLSR[2]--LTPF   FKQVIPHLYVFLDnRRIKVAFPDEYT--   778
Query_LOTTI  734    VVPAIQVA---LTLGLSY--LPLAGIAIDALEYWCHNLPD   HIVKPH   YSCILPYLDNYIR-T-VDKGMVAFS---   797
Query_APLYS  753    VVPAIQVT---FTLGLSY--LPLAEVGLQALESWFRQLPA   SVLEAH   YTSILPYLDPYLR-T-DDLGAEDVVK--   817
Query_SCALL  765    VVPVIQIT---LKMGLSY--LPLAKVALDALEYWSDTLPA   SVIQPH   YSQILPYLDTYLK-T-ADKGAEEVSV--   829
Query_CRASS  750    VVPAIQLT---LSIGLGY--LPLAMIALDALDYWSDTLPA   SVIRPF   YPQILPCLDSYLK-T-MDQGADDVS---   813
Query_OCTOP  739    LLPAIQMS---LSIGSTY--LPLANAAMDTLEYWSETLSQ   DALKTV   YSQVLPHLDPFI----------------   791
Query_POMAC  746    IVPAIQLT---FSVGLSY--LPLAEVGLDSLVWWSQVLPP   EVLAPH   YANILPCLDSYLK-T-VDTGAEDVSV--   810
Query_WHIPW  728    IKKIIRNG---FALGLTY--LPIASVCLDALEKWSEQLCC[2]SERQQF   YNELAPVLDEYLR---CTKFEGRKVV--   793
Query_TRICH  726    IKEAIKSS---LKLGLNY--LPIACVCMDALEKWSAEFPE[2]PQLDSF   YTDLLPIFDDYLR---CTSVDETLLV--   791
Query_AASTA  707    LVPALKQA---LALGVFH--FDTAVAAVEALERW---RSD[1]--VVPF   YPEVLPLLAPYLD-------EA------   758
Query_AEUTE  693    LIPALKQA---LSLGTSH--NDTAVAAVSALENWHQVQPN[1]--LKQF   YQELLPLFAPYL--------ST------   746
Query_PLASM  736    MVGIVKSM---LDLGRSF--IPAAVVAIAALERWQKQRPN[1]--LNDV   ICDIIPLISAYLN-------QEKLGEr-   795
Query_BREMI  733    MVGVVKPL---LELGRSF--LPAAIVAVTTLERWQIQRPN[1]--LEDS   ISDLLPLLLEYLD-------QEGLSDk-   792
Query_OLIGA  735    IVQAMVPA---LILGKSF--LPAADMAVTCLERWQQERPW[1]--LDLV   LDQIVPCLTAYFD-------QSDAVDp-   794
Query_LIGNO  779    ALAAVRRG---LRLGRFW--PPLLEAALAALERWTAAASS[2]--LSDR   ---VAPVLAGLLA----VLLSDGGSG--   838
Query_CESTO  642    VITSISQL-----VNFQGrcASLAIEIASAIEACLKENEA[1]--LSFY[3]FSALVPALVNLLN-----ANPDKQIAik   707
Query_OVI2_  724    IYPAISMA---VRIGAVH-fAFLWAAVISALNRWLSEFQL[2]SVTNSF[7]LRQILVWSDDSAKaFSAPSSISSIRSll   802
Query_PRIAP  767    IVPAVQVA---LKIGLSY--PPMAHTTLDALEMWVANLPA   DVMQPY   YKQILPHLDGYLK----TTEGADSVP--   829
Query_ROTIF  767    SFHCLNQA---LDLGVNY--LKLAEDALNSLEYWLKNIPT   QLIKPY   FPSILTKFENYLQ---ASSNKSDLDS--   830
Query_SPONG  759    LIPAIKLA---LRIGMGY--LPLAKSVMEALEKWTLYIVE[4]DTLQCL   VNEILPLLADYLKaTSVQLKEEEDGG--   829
Query_TARDI  773    AMDAIVIA---LQAGKSY--LPLAEAAVTTLRRWCFDSDI[2]QLLQPY   FSAILPVLDGFLR-------STTLDV--   834
Query_TRICH  737    IIPAIKVS---MNIGLSY--LPLANAAIDALERWSSDLPQ   NIIGPY   FSDILPCFDDYLK-SSADDTTETLVE--   802
Query_OSTRE  742    LISPLLRA---LELGLQYpgMPITGVALDTLERWIDGKHV[2]--VSQY   LPQIAKRLCAYLD---SAAKSDSLFS--   807
Query_NITEN  811    LVGPMQSA---LTMGLSY--QPLADLALDTLEGWLKNHFR[2]-EP--H   LPEIVPLLNDYLAeVDTSGADQMSGAag   878
Query_CLUBM  718    IVRPLKLA---LKIGLRY--HPLANIALDSLQRWYNLHPK[2]-QK--Q   LQDIVPCLRDYLFeV-------------   772
Query_MOSS_  760    IIPSLRNA---LKIGLRY--PPLAEVAMDSLERWQAAFPE[2]-QH--W   LPQIIPCLNDYLVdVEKADVSEPGTSlv   827
Query_LIVER  777    LVAPMQNA---FKIGLRY--PPLASVAMDALERWHRLYPR[2]-EI--W   LPTIVRCLSDYMAeVETTELSDSSPSdl   844
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  812   DTLSLVQKDPMKK-qGKKKI[1]VKL-L   KQTMFDAE   K[1]KESELKEIQRKCVLLLGSLGGSINHALVFG-QTA  878
Query_LINGU  819   PSTLIMEPKGHSY-sRKKGL[1]IKL-L   KPAVAKEE   S   --SQLELLKQRIVKFLGNLGGATNAAMVQC-TSE  882
Query_LIMUL  813   PLLSRTSS--KNK--FATKS[4]IKK-V   NKDNQSK-   -[1]QETEITVLRLRILKFLGSIGGKNNKALLEK-SEE  877
Query_SPIDE  822   KRPSKKAY----K--MSVKS   MKR-S   NKNTSSDN   S[1]EDPALGRIQLQILKYLGQLG-ENNVALLEQ-SHQ  881
Query_SCORP  814   HTFTKQSLVVKKK--LPLK-   ISQ-Y   KKIKQTD-   -[1]HESELIKVQHRILRFLGSVGNQYNFAMISA-DKI  875
Query_TICK_  770   EPAESTKRARSQG---RHKV[1]LKL-L   YQKKRDQI   K[1]--PETESLRLDVLCFLGGLQGH-RRALLLADLER  832
Query_MONOS  747   ---AVITDGQAKR------T[4]ARS--   -KPGAVED[3]-   --APLDRVKGTIVTLLGQFGLDVSHHLALIGSPE  806
Query_ROSET  724   -EMKLLSRSITRK-------   -RE--   -QAAVVAV   -[1]SDNPFAALKMRILRLLGAMGGQHNHELAALEHQG  779
Query_BELCH  834   FQTVSLTAASSTR--GKRKI[1]VKL-L   KA---AAE   R[3]QESDLVKARLRIVRFLGSLGGHTNACLVTGLTAE  899
Query_CIONA  817   STQNTL-----KL----KPK[1]RSF-A   QKL-MSLK   E[1]DEPARKKARNQVIRFLGSMGCSDDHGDLNEDAG-  874
Query_CAECIL 822   WEVMRLSRAAQKG--FNRFL   VQR-L   KKTKTLSL   D   QDVSLASIRNRVVQILGSLGGQINQSLVTASCAE  886
Query_HUMAN  813   WEVSALSRAAQKG--FNKVV   LKH-L   KKTKNLSS   N   EAISLEEIRIRVVQMLGSLGGQINKNLLTVTSSD  877
Query_MOUSE  812   WGLSALSRAAQKG--FNRHV   VKH-L   KRTRNSSP   D   EALSLEEIRIKVVQILGSLGGQINKSLVTATSG-  875
Query_XENOPU 834   LDMVRISRSLHKG--FNKQL   IQQ-L   KRMKTLSV   K   EESSLTAVRNRVVRILGSLGGQINRSLVTAASTE  898
Query_PLATY  815   MEVTMLSRATQKG--FNKVV   VRH-L   KKAKATAP   V   EDLSLESVRTRVVRLLGSLGGQINRNIVTAASAD  879
Query_TASDE  818   EEVTVLSRAAQKG--FNKVV   IKH-L   KKSKTILS   D   AEMSLEAVRAQVIRVLGSLGGQINRNLVTAASSD  882
Query_CANAR  812   WEVRKLSQAAQKG--RNKVV   IQR-I   RRAKTLSV   D   NSPSLQAVRTRVARLLGSLGGQINHNLITATSAE  876
Query_OSTRIC 783   WEVRKISRAAQKG--FNKVV   IQH-L   RRAKT-SL   V   GDPSLEAVRTRVARLLGSLGGQINHNLITATSAE  846
Query_TURTLE 821   WEVRKLSRAAQKG--FNKVV   IQR-L   KKAKTLSL   E   DNLSLEAVRTRVVRLLGSLGGQINRNLLTAASTE  885
Query_GECKO  821   WEVKKLSQAVQKG--SNKVV   IQR-L   KKAKSLAL   E   EDHSLEAIRMRVACLLGSLGGQINRNLVTAASAE  885
Query_DANIO  827   MEVTFVSTGSSKG--YGQVL   LRL-L   KKSKRFSL   G   DESPIAAVRRRVVRLLGHLGGQLNRSLVTAVSAE  891
Query_SNAKE  821   LEVKKLSKAIQKG--MNKVV   IQH-L   KKSKNLEL   E   ADPSLEAVRARVVRLLGTLGGQINRHLITGTSAE  885
Query_STERL  831   WEVMSLSRSTEKG--YGRVL   MRL-L   KKAKRFSL   T   EDSPTALVRRRVVRLLGHLGGQINRSLVTATSAE  895
Query_MILII  819   WEGMNINNPSRKG--GSKVL   FHH-T   KKSKEFLV   S   EDSAITAVRRRIVRLLGSLGGQNNRSLVTAVSAE  883
Query_BAMBO  822   WEMMSMNTSS-KG--GNKVL   IRR-L   KNARDLSM   S   EETPAAAVKKRVVRLLGSLGGQINRSLVAATSSE  885
Query_ALLIG  781   WEVKKLSQAVQKG--FNKVV   IQR-L   RKAKALSL   N   EDLSLEAMRSRVVQLLGSLGGQVNRSLVTAASEE  845
Query_GAR__  820   WEVMCLSSGAEKG--YGRVL   LRL-L   KRTKHLSM   A   EDSPSAVVRRRVVRLLGHLGGQLNRSLVTAASAE  884
Query_TETRA  808   ---YKVQSNNQQRdiFNLTV[1]-EL--[1]QRKKDLIE   -   ----------RVTNLLGDIGGESHLIVLSEKGFQ  860
Query_PARAM  781   TPIQYIDSISPNI-----LL[4]IDQ--   --------   -   -----NHIVDKVRYFLGSIGGLCHLIVDETYNES  831
Query_SLEMN  798   ---KKLNDATGGD--TYQYL[1]-ML--   K----SGN[4]E[1]-----KDIALKVMDLLGKIGGNAHSII-------  847
Query_STENT  726   ----KITQEVQGQ--VEDQL[1]-RQ--   K-------   -[1]-----KNIAYRIVKFLGSLGGACHYLV-------  766
Query_NEMAT  812   VSRRTATLRSTYS-rLGKIL[4]IDF-V   SKRNILPQ   V   GESSLVTIRVQILRLLGQLGGHTNLCLLGSDGGV  881
Query_PISTI  829   CNTIVLITSSHHT-sRGRKL[4]LKM--   -KPQSQHQ   M   TLTPLAKVRHKILRMFGGLGGETNILLLDNTG-I  895
Query_CORAL  818   KETKLRSRKSGKT-kASK--   -----   SSKRNPNE   E   ENSPLNKVRHRILILLGTLGGKCNLGLLEGGEGS  877
Query_HYDRA  765   TEHVTLTKSERKV----KRV[1]IRY-L   KALAMSQK   D[1]SESPLFNLELRVIKFLGCLGGRLSIGILQNAAK-  828
Query_LATIM  821   WEVLRLSRASQKG--DHKVL   IRL-L   KKAKAVSL   N   EDSSLGAVRGRVVRVLGSLGGQVNRSLVTAASSE  885
Query_SHRIM  803   TRVITVKMAFARQ---RRKV[1]QKK-M   KESKRVEE   T   ---AMQKIQLMILRLIGSLDPLLRLHTIPQDPDT  864
Query_MCRAB  800   TRVITVKMAYANQ---RRKV[1]QKK-I   LASKQVDA   T   ---AMKRVQSLILQLIGGLDPSLGQYVIPQDADS  861
Query_AMPHI  826   SRVINVKLSLLNQ---RRRV[1]VRK-M   RDARCSKV   S   ---PSIDLERNVLKLLGSLGPDV-CSALSSETGS  886
Query_BARNA  796   ---------SSRR-------[1]-----   --------   -   --SELTPLRRRALQLLGLLAAA---GLVADSSEA  829
Query_COPEP  773   ------NAEIKMR---RRRA[2]VKR--[5]KESEGIAK[4]E[2]---------QLALGILAKVSSKLKYLSLPTKEEI  833
Query_DAPHN  735   -----VTFISGKR-------   FKR--   RKPNIQST   -[1]ERESWEKIRRRVVRLLGRMGPKCSRLVAQNEDHV  788
Query_STARF  833   VQTVTMASASSSGraGKRKL[1]AKL-I   K---GAST   T[3]QDSPLRSVQLRILHLLGSLGGQTNAYLLET-SAE  899
Query_STRON  838   RPEVLVIQTASSR-vGRRRI[1]SKL-L   KK---AKD   N[3]QDSPLRSVQLQILRLLGSLGGSVNVGLLET-SAQ  903
Query_SEACU  825   NQFVEMKGTLGSG-aGRRRI[1]VKL-L   KTSGGAGK   V[3]T-SPLRDIQHKILQLLGSLGGKTNISLLES-SSE  892
Query_FUNGI  820   ---DNNSTA-----------   -----   --------   -   ------HIQEQVLRLLGKLGGD-NIYMITDSD--  850
Query_DPURP  891   ---SSNTQQIDLH-lMKETW   -----[2]HRKMKYKS[4]D[5]FKTSIEDLQNRIIRLLGKLGGD-NIHFLGEVSLL  959
Query_HALBU  906   ---ESANSKTGML-mGR---   -----[3]PRMNTKSI[3]D[5]QRTVVEEMQTRVIRILGKLGGQ-NYHLMGSVNLM  971
Query_GLOBO  747   -ETASTVS-------G----   -KK--[5]RTKHLKTH[4]E[5]IVQDLKSIQIRILKLLGSLGQD-SRLMLSN----  808
Query_PALUS  779   -SDQSSSSAPRQS-lG----   -KKgP[6]SAANTRAK[4]A[5]TNTQLRAIRYRIIVLLGMIGGD-NRFLLEE----  849
Query_CALIF  854   ---------RKDA-eLDDYI   -----[5]KTMKKINK[3]-[2]VEETNNNIKIKVLRILGMIGDG-IPYLYSNFDIK  914
Query_GLOMU  840   KEFAQIRNTQRNR-rGNFKS[4]RKErV[6]KSYKTVV-[2]G[5]V--GLHDLQVRVLRLLGRLGGV-SKFLFEDIVNN  919
Query_LOBOS  816   PEATSLKSKQKG--------   -----[5]QSYKVVAR[4]P[5]V-ESMRDLQLRILRLLGRQAQY-NKLILAQRTAD  882
Query_RACEM  790   EEADDSVKRRKTF-nV----   ----A[6]QRWRTLNE[1]-[5]V--SLRDLQLRILRFLGRLAGA-SKLMIQTPLHQ  856
Query_PVIOL  867   ---EGSSTSISIA-fSEEKL   -----[5]NRKKKFKS[3]D[5]FKTSVEDIQNRIIRLLGKLGGD-NSFFLGDVNLL  937
Query_BOMBU  750   ---VEFLQDII-------KT[1]RKV--   IKRIILRD   -   DENTLERFQMRVLLFIASLDTNIIMDFIYKK-SM  804
Query_AEDES  794   -RLVKFQRKR--SvhLTSSK[1]EKI--   KLQLALEA   -   SESDLVKLQLRILVFLGDLDPNVCTKMILRSETD  856
Query_ANT__  775   ---VEISQDLT-------TT[1]KRV--   -KRIDVID   -   TECTLQNFQRRVLLFLGSLDHDLLTSFVHERASR  829
Query_WHITE  777   -GLDAEFKVL-----TSALR[1]SKR--   KIKKIHER   -   QDTELFRIQKRIILMLSKIDISTALTILSEDSTR  836
Query_BEETL  759   --LESMSSGI---ndNRRKS[1]QAL--   KKRKVLLD   -   TEPELLRLKTKMIKFVGQQNSLLCEA-LAVSEDN  818
Query_THRIP  754   -NLYGMTMV---KtvLQEQK[1]KKL--   WKYKVNSH   -   SDTELVKLQKEVVVFLSTLDSSVCLSLLDGSEDR  815
Query_TERMI  784   -IIEKELMVLGRRnrTNYSQ[1]KKV--   ARKKIIEH   -   PESELSQLQHEIISFLGQLDSTTCLALLKDPQQE  848
Query_FLEA_  739   -CLITDSRSQ--K-------   -KV--   KKQI-IEG   -   TPSELALLQRDIITFIGKLSSMDCSSLISK----  787
Query_MIDGE  783   -DSKDFQKS------LKNSF[1]IQT--   TI------   -   -ETDLTRFKKRILLTLGRFDPNDVKFILSKFDQK  834
Query_LAODE  779   ---EYAQRSSSRG---GSHV[1]-----   KRRKLVEQ[4]S[2]GETQLSRLQQAILQLLGRLDMQSLLSLLDQDYTD  842
Query_LOTTI  798   ---IFMWYISCGM-hIRFSL[1]FSS-F   SSFSTINC   L[5]YESKLGVIKDRVVQYLGSLGGSINYTLLTN-SQD  865
Query_APLYS  818   ETTVVVSAKKSSR-r--SKA[1]IRL-I   KNAQSADD   K[1]SKSQLSQVKQRILYFLGSLGGQVNHTLQDR-ADA  882
Query_SCALL  830   DGTVNMTKSKTGR-g-RRKL[1]VRM-V   RGPKKDQL   K[2]YESELSQVKLRVMRYLGGLGGSQNHAMLAD-NEA  896
Query_CRASS  814   ENTMEMNKSKSGK-g-RKKL[1]VRI-L   KGPKTDS-   K[2]FETQLAKVKQRVMTYLGRLGGEVNHALLSG-SD-  878
Query_OCTOP  792   ----KTSERDSDD--IDLTN   IDQ-K   KKRYFLD-   -   PDSELSKIKYRIAKYLGNIGGNFNKDLLAN-SDQ  849
Query_POMAC  811   ETLVTVSSSKSKY-r-GKKL[1]VRL-V   KERKPVDL   D[2]YESQLSVVKQKIVTYLGSLGGSMNFALLAR-REE  877
Query_WHIPW  794   D----FPEAAGIY--KAAQL[2]VRR--[6]GLQTKIDA   S   GETEAQRVLRRIAVFVGSHGHDVSKALLQVNSRM  861
Query_TRICH  792   NNGFNYNGEKSTF--FNKKL[2]VNS-D[6]KADDNDSE   E   EMPNPQSVLRRIASFVGSLGCYMNRFL--IDSAD  862
Query_AASTA  759   --------------------   -----   --------   -   -ANQSASFQISVLRLLGSLGGYSRFVVQEHHAAR  791
Query_AEUTE  747   --------------------   -----   --------   -   -SNEGLTVQLKLLKILGSMGGYSRYVVQE----E  775
Query_PLASM  796   -EHLLDKATLQGV-------   -----   ---KATGS   -[1]EASDLAQVQRRLLLLLGRCDGNILLLVSEPSSLI  847
Query_BREMI  793   -ELWLMKP--RGT-------   -----   ---DTVNT   -[1]EISNFTQLQRRLLLLLGNCGGNISLLMSEPPSVI  842
Query_OLIGA  795   -EIVVIKGDASDG-------   -----   ---GNSES   -[1]NATERSKLQRRIMVILGKMGGYGALLASRGANTA  846
Query_LIGNO  839   AGLTDAEAEAQRGgaDRRSL[4]VAR-A[6]RRARAKNY   A[4]AELGLAEMQKRVLLFLGSAQSSFKYQLAEIASDK  919
Query_CESTO  708   SALPPIGRRSSRS--YRRTA[4]LKT-I[1]KQPTTLSS   -   -DSLIREAQIKAIELIGQNVAHWNLLNTGETSKE  775
Query_OVI2_  803   HQHHERLAKSSGR--TRRDL[3]VRL-I[6]HQSIPSVT[4]G[5]PSQIRKNAQVELLHLLGRLGGLIIASLTSSSIQE  885
Query_PRIAP  830   TPTVAITSANATR--GRKKI[1]VKL-L   QG----SG   R[3]EDSALTTARMRIVRFLGGLGGANNISLISC-DAE  893
Query_ROTIF  831   DEKVILLNLSHKG-kGRKKL[1]----V   KLFDKSSL   N[1]--DVYEQIQLKILKILGQLSADMSHSVHLIDNED  893
Query_SPONG  830   RGKQYATKRYARM----KMV[1]-----   KWSEGLHQ   G   SRDDITSVRRRILLLLGSLGGVVNESLSKVPSQH  889
Query_TARDI  835   TERSKVVLGTGKG-------   --K-T   KRRSPVAV   G[1]DQVALTKVRELTLSFLGDLGGKYNHLIVDGDFLN  893
Query_TRICH  803   DSSKALKARTGKA----KKL[4]VNP--   -----TGG   S   KDSPLAQVRLRIVQLLGRLGGNTNSLVIRKGNDD  863
Query_OSTRE  808   ------SNETEDE---KRSI[1]----I[4]KKRKQLAG   K[5]DVVLKRATLIRITKIIGAMGGAGHGLLEDMKPSK  872
Query_NITEN  879   DEVAEKRNTA--D-mTRTQA   -----[3]REARAKDM   -   QASPFERLQLRVQRFLGRIGGH-SHLLVPDVNLT  939
Query_CLUBM  773   ---ADSRKKTMDS-kT----   -----[1]KLLEKEKT   -   KVVTFEDIQVRVQRFIGQIGAY-GHFLVPDANLL  826
Query_MOSS_  828   TEAVGTSAPTISS-mSRGQL   -----[3]KRARAKDS   -   QPSPFELIQVRVQRFLGCIGGL-GHFLVPEVDLL  890
Query_LIVER  845   TTLESSSSR--ES-lSRLQL   -----[3]KKTRAQAS   -   QAGPFELTQIRIQKFLGQIGGF-NHHLVPELNLL  905
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  879   -N-PPD    LVAWDS--DQHLKFDLPFL-D---IK-TSIYFD--AFLPKVVTLAT    VSSDRQTKVAACELLHSLVL  939
Query_LINGU  883   -DIAKI    AVAWDQ--EKHLKFDMPFM-D---VK-AALYLD--PFLPRVVELAT    LSSDRQTKVAACELFHSMVL  944
Query_LIMUL  878   -KISEK    AISWDNrvQDHLKFAVPFA-D---MK-PEICLD--NFLPQVVHLAC    NASDRQTKVAACELFHALVI  941
Query_SPIDE  882   -KVDEV    AIAWDPlcKNHLKYSVPFV-D---MK-PNVYFD--EFLPRILEIAT    KSSVRQVKVAACEALHSLVL  945
Query_SCORP  876   -KIHEV    AVAWDInkQNHLKFYLPFI-D---VK-LEIFLD--KFLPRIVTLST    TASQRQTKVAACELLHAVVT  939
Query_TICK_  833   -DTLKA    ALAWDPetKEHLVFALPFP-D---TK-VEIALD--AFLPRVVELAR    FAGDRQTKVAGCELLHAMVL  896
Query_MONOS  807   -DAEAV[ 2]WVAWSA--SQTLPLTVPFP-D---IK-PSIKLA--PLLPRIVDLAT    NAGDRQTRTAACELLHAIVL  870
Query_ROSET  780   ------[14]WVAWDMt--NHLKFAVPFQ-E---MK-PEVQLD--PMLPRIVELAL    HSGDRQTRTAACELLHAMVL  850
Query_BELCH  900   -DLAKR    AVAWDT--QTHLQFAVPFM-D---MK-PTVCLD--PFLPRICELAT    TSSDRQTKVAACELLHSLVL  961
Query_CIONA  875   ---LAV    -APWDE--DVNIQYPIPFR-D---LK-PCIRLD--KFLPRIMELAT    QGGERKAKISACELLHSITI  933
Query_CAECIL 887   -EMMKK    YVSWDT--EKRLSFAVPFV-D---MK-PIIYLD--SFLPRVTELAL    STSDRQTKVAACELLHSMVV  948
Query_HUMAN  878   -EMMKS    YVAWDR--EKRLSFAVPFR-E---MK-PVIFLD--VFLPRVTELAL    TASDRQTKVAACELLHSMVM  939
Query_MOUSE  876   -ERMKK    YVAWDA--ERRLSFAVPFR-E---MK-PVIYLD--VFLPRVTELAL    SASDRQTKVAACELLHSMVM  937
Query_XENOPU 899   -EMIKR    HVSWDT--EKRLRFDVPFK-D---LK-PVIYLD--MFLPHITELAL    STSDRQTKVAACELLHSIVA  960
Query_PLATY  880   -EMMKK    CVAWDT--EKRLSFAVPFA-E---MK-PVIYLD--VFLPRITELAL    SASDRQTKVAACEVLHSIVM  941
Query_TASDE  883   -EIMKK    CLAWDS--EKKLSFAVPFA-E---MK-PVIYLD--VFLPRVTELAL    SSSDRQTKIAACELLHSMVM  944
Query_CANAR  877   -EMMKK    CVSWDT--EKHLSFAVPFA-D---MK-PVIYLD--LFLPHVTELAL    SASDRQTKVAACELLHGIVT  938
Query_OSTRIC 847   -EMMKK    CVSWDT--EKRLSFAVPFA-D---MK-PVIYLD--LFLPRVTELAL    SASDRQAKVAACELLHSIVA  908
Query_TURTLE 886   -EMMKK    FVSWDT--EKHLSFAVPFA-D---MK-PVIYLD--LFLPRVTELAL    SASDRQTKVAACELLHSIVM  947
Query_GECKO  886   -ERMKK    FVSWDT--EKHLSFAVPFA-D---MK-PVIYLD--VFLPRVTELAL    SASDRQTKVAACELLHSMVM  947
Query_DANIO  892   -DMMKR    FVAWDC--EKRLSFAVPFK-D---MK-PVIYLD--SFLPRVTELAL    SSSDRQTKVAACELLHSLVI  953
Query_SNAKE  886   -EKMKK    FVSWDI--EKHLNFAVPFV-D---MK-PIIYLD--LFLPRVTELAL    CASDRQTKVAACELLHSIVM  947
Query_STERL  896   -EMMKK    FVAWDS--EKRLNFAVPFA-D---MK-PVIYLD--TFLPRITELAL    STNDRQTKVAACELLHSVVV  957
Query_MILII  884   -EMMKR    CIAWVT--EKPLSFAVPFM-D---MK-PIIYLD--SFIPRISELAL    SASDRQTKVAACELLHSVVA  945
Query_BAMBO  886   -EMMKK    CVAWVA--ERPLSFAVPFA-D---MK-PVIYLD--SFLPRITELAL    SASDRQTKVAACELLHSVVA  947
Query_ALLIG  846   -EMMKK    FVSWDS--EKHLNFAVPFV-D---MK-PVIYLD--IFLPRVTELAL    SASDRQTKVAACELLHGLVM  907
Query_GAR__  885   -EMMKK    FVAWDS--EKRLTFAVPFA-D---MK-PVIYLD--AFLPRITELAL    STSDRQTKVAACELLHSVVV  946
Query_TETRA  861   gNVERQ[16]LISWDY--EAKVLFPLPLL-----AKkIDINLT--KLLPRMVELTL    -NPDDKIKVYACEAIHAITI  938
Query_PARAM  832   ------    LLALDY--NSPITINIPVG-Q-rkMK---LRFG--QFLERILSLCL    SAPLEKSKIYACELLQAFAI  888
Query_SLEMN  848   ------[12]FIKWDP--EKRLKFTVPLY-----TKkVDIYFD--ACLPKVVDLAK    NSSDRETRVAACEFLHALII  916
Query_STENT  767   ------[ 7]HLAWDS--QPRLQMKIRFN-----LKyLEIDFD--KILPRVIHLAE    KSSDKKTKIISCELLHALIL  830
Query_NEMAT  882   -DS-GS    AVAWDT--VKHLPFAVPFQ-D---MK-PTIYLD--PFLPRVVELAT    TSSNRQTKVAACELLHTLVL  942
Query_PISTI  896   -NS-SN    AVVWDT--ENHLTFAVPFQ-D---MK-PSIYLD--PFLPRVVELAA    KSSDRQTKVAACECLHSLVH  956
Query_CORAL  878   -VA-SS    GIAWDT--KQHLSFAVPFQ-D---MK-PTIFLD--PFLPRVVELAT    SSSDRQTKVAASEVFHSLVL  938
Query_HYDRA  829   ---FSK    HIVWDT--NQHLPFALPFR-D---MK-PTIYLD--MFIPRVVELAL    YSSNRQIKVVSSELLHGLVI  888
Query_LATIM  886   -EMMKK    FVSWDT--EKRLSFAVPFA-D---MK-PVIYLD--SFLPRIAELAL    SSSDRQTKVAACELLHSVVV  947
Query_SHRIM  865   -IASAA    -VRWNV--QKLVKFATPFE-H---IK-IDIYLD--DLLPHVVDLAV    NSSDRQTKVAASELLHSIII  925
Query_MCRAB  862   -LAAAA    -VCWDS--QKHIRFAMPFG-Q---AK-IDIYLD--DLLPRVVDLAL    NSSDRQTKVAASELLHSIII  922
Query_AMPHI  887   -VANTS[10]GIRWDE--CDHVALKIPYE-H---ST-LDLHLD--SALPRIAHLAL    HATNRQTKVAAAELLHSVVI  958
Query_BARNA  830   ------[ 1]LVAWSR--RPVLGLDVPFS-D---VK-PKLWLE--TFLPRTVQLAL    GSGDAQARTSACELLHAVTV  887
Query_COPEP  834   ------[ 3]AVVWDQ--QKHLLVEVSLR-D---VN-LPLYLDl--LLPITLDLVE    KSVNRNTRVSAAEVLHATLL  893
Query_DAPHN  789   --IARS    ATGWQI--KGHLEFGLPLS-D---MK-PIIQLD--QFLPRILHLVL    HTSEIKTRMAACEFLHASVI  849
Query_STARF  900   -EIAKV    AVAWDT--QPRLQFAVPFV-D---MK-PNIDLD--RYLPRVVELAL    TSGDRQTKVAACESLHTLVL  961
Query_STRON  904   -ELSNV    ALAWDS--QPRLNFPVPFV-D---MK-PEIQLD--AFLPRIVELAL    TSSDRQTKVAACEALHTLVL  965
Query_SEACU  893   -ELAEM    AVAWDT--QKKLSFAVPFM-D---MK-PTIFLD--PFLPRVVELAL    SSGDRPTKVAACELLHTLVL  954
Query_FUNGI  851   --H---[ 4]GIAWDD--TKRVRFTLPFP-D---VK-VDLYVD--STLARIVELST    NANDPQTKTAAAELLHGILL  912
Query_DPURP  960   -AP---[ 4]GVAWDT--EQKVLIKIPFT-D-qtASaIEIHLD--TILPSVVSLAE    NSPNRQIKVAACELLHSVLL  1025
Query_HALBU  972   -AP---[ 4]GVAWDT--ISRIQLQIPLG-DlngVS---IYLD--SVLPVIVTLAT    RSTHRQTKVVACELLHSIVL  1035
Query_GLOBO  809   ------[ 5]LLAWDT--EMNLKVQIPFK-E----AiFDVYFD--NMLPRIVELAE    NSPDRKTKVAANELLHALVI  870
Query_PALUS  850   ------[ 6]LIAWDT--EKRVWFDMPFK-E----InCRIYMD--DILPRVMDLAE    SSPDRKTKIAATELLHACVI  912
Query_CALIF  915   -DY---    -LKWDP--ENGITFILHFV-D----FqTEINLD--NLLPRILELSQ    LSSNRKVKIWACELLHSITI  972
Query_GLOMU  920   -DLVYR[16]LLAWDP--EKRVSFNIPFP-D----TkVEILFD--EFLPRIVELAE    SAPDRKTKIVACESLHSIVI  997
Query_LOBOS  883   -EV---[ 7]LLAWDP--ETRLKLKLPFL-E----MkAELQFD--EMLPRIVDLAE    NSLSRQIKVVSCELLHSLMI  948
Query_RACEM  857   -NKDER[10]MLAWDP--ARPLKVRMAFQ-G----AtIDLSFD--ELLPRICDLAE    SSPDRKTKSAACEVLHALVL  928
Query_PVIOL  938   -AS---[ 4]GVAWDT--EQKLSFEFPFI-D-qkAN---IFLD--SILPGVLSLAE    KCTIRQTKVAACEFLHSILL  1000
Query_BOMBU  805   -NTG--    -ATWDK--KDLLKYSLTFS-D---TQ-ADIHFD--KILPRIILLAQ    NSGDRRTKVIACEVLHSIVM  863
Query_AEDES  857   -DAAQQ[ 7]LVLWDLdsNRNISLQLLSVtG---IR-PIIKLD--TIIARVCALAV    GSSDRKTKVAACELLHALIL  928
Query_ANT__  830   -STG--    -ASWNH--KDLLKYKLPLP-D---AR-LNIHFD--RMLPRIIALAC    DSSDRRTKIAACEVLHSITA  888
Query_WHITE  837   ------    -SKWKSt--KQLKFTIPFP-D---CR-LNLYLD--EMLPRIVESAL    SSSNRKTKVAASELLHAVVL  892
Query_BEETL  819   ------[ 2]LTVWAN--EQHLKIDLPYE-D---VK-LQVYLD--KLLPRIVNVAL    NSSDRKTRITACELLQATVM  877
Query_THRIP  816   ------[ 2]VTAWQK--ERLLKFPMPFQ-D---MK-VDISLD--DLLPRLVHLAV    NCSDRQTKTAACESLHAVIM  874
Query_TERMI  849   ------[ 2]VIPWATclQQPLVFPLPFK-D---MK-LNINLE--KLIPRVVELAM    TSSDRYTRFTACELLHGIVM  909
Query_FLEA_  788   -----N[ 7]LIPWDL--KQHLRFRLPCQ-G---VK-PLIYLD--KLLPRINKLVL    DSNDRQTRIAACELLHGIVL  852
Query_MIDGE  835   -LVRDH[ 6]-------------IRLLCEdP---LI-PNIFLD--NLIERVCQLAL    SSSERTTKIAACELLHGLVL  892
Query_LAODE  843   ------[ 3]--SWSR--KPVVTWLMQFQ-E---LH-VPLQLDi--FAPRAMEIAL    QSSDPANKDASFELLHAIFV  900
Query_LOTTI  866   -DISKQ    AIAWDT--EQHLRFETPFI-D---MK-PIIYFD--PFLPHVVELAS    TSSDRQTKVAACELLHSLIL  927
Query_APLYS  883   -DISNE    AVAWDS--QKHLRFDVPFF-D---MK-PVIFFD--PFLPHVVKLAQ    QSSDRQTKVAACELLHSLVL  944
Query_SCALL  897   -EVSRQ    AIAWDT--HRHLKFDMPFI-D---MK-PTIYLD--PFLPHVVMLAT    KSSDRQTKVAACELLHSLVL  958
Query_CRASS  879   --VCEE    AIAWDA--EQHLRFDVPFM-D---MK-PQIYFD--PFLPRIVELAT    KSGDRQTKVAACELLHSLIL  939
Query_OCTOP  850   -DVLKH    AVAWDT--KKHLKFDLPFA-D---MK-PTIYLD--PFLPPILELAT    KSSDRQTKVAACEFLHSAVL  911
Query_POMAC  878   -EISRR    AIAWDT--EQHLRFDIPFF-D---MK-PSIYFD--PFLPHVVELAT    QSSNRQTKVSACELLHSLIL  939
Query_WHIPW  862   -VKQVL    -IRWTT--DRLLNFPCPFP-D---VK-IDLCLD--DFLPHLCDLAL    SAIERPLKVAACEALHSCVV  922
Query_TRICH  863   -VGS-L    -LRWGQ--GSHLEFICPFP-D---MK-FEIALD--TLLPRLCDLAL    SAGDRPLKVAACEALHACVL  922
Query_AASTA  792   -DDVDK[ 1]-----A---AALTFPVALD-R---VH-ADMSIE--LLLRQMRDLAV    TSIDRPVKAAAGEAYHAMVL  848
Query_AEUTE  776   -DPQLQ[ 1]-----S---APLDISLAME-N---VH-VEIPIS--NILHQMKDLAL    TSIDRGIKSAAAEGYHAMVL  832
Query_PLASM  848   -TASGI[ 5]-----S---PHFRLELRLS-E---TP-LSLALD--SVLSHLGNLAS    QSSVRRVKVSACESYHAMIC  908
Query_BREMI  843   -NSNSM[ 5]-----S---PYFRLELQLG-E---VS-ASLAMN--EILSHLGTLAC    KSSNRRLKVSACEGYHALVC  903
Query_OLIGA  847   -DAK--[ 4]-----T---STLQLPITLN-D---IT-IDLPLN--KIFDYIGSLAC    DSTDRRVKVGACEVYHAMAS  904
Query_LIGNO  920   -AATAD    LAAWSPg-GRLLYFDLPFA-D---SK-PRLYLD--PLLSRCVALAT    ESSSRQSKVAAAECLHSIVL  982
Query_CESTO  776   --SDDR[ 9]--VFND-----LCLCLPFP-D---LR-PSIRLSdlRLFTRILHILL[14]KGFSRQTRVCAAEYIHEYIV  856
Query_OVI2_  886   ---FRP[ 4]-LEYSSv---GLRLCLPYA-D---IR-PSFSLTdhSLLRSVLEVSS    AAQTTPSTSAAVEYLYSLVV  949
Query_PRIAP  894   -DIAKV    AVAWDA--EKNLKFGVPFA-D---MK-PVIYFD--VFLPRVTDLAL    SSSDRQTKVVACELLHALLL  955
Query_ROTIF  894   ------    ITSWDT--VQHLKFYVPFT-D---IK-PTIYFD--RFLPRVIHLAQ    NSPSRQTKINACELLHALVI  950
Query_SPONG  890   -SL---    -VSWDS--DRRLGFAVPFS-D---LK-PEIYLD--PYLPTVCELAL    SSSDRQTKVAACEFLHAAML  947
Query_TARDI  894   -DVKNL    AIAWDMv--QHLQFSVPFD-D---MMlKSFYLD--VYLPRVTELCL    HSPDRPTKVAACELLHAITI  956
Query_TRICH  864   -GS-SQ    AVAWDN--HKRLPFAIPFQ-D---MK-PIIHLD--PFLPRVVELAT    SSSDRQTKIAACELLHALVL  924
Query_OSTRE  873   ------[ 2]-LPWDR----NCKLPTPIG---lvRSsTNLWLD--SLLPRASHLAL    FSQDRPTKVHACEALYGFSV  930
Query_NITEN  940   -KG---[ 4]ALAWDV--IKRIKFALPFR-D----EkPDIYLD--PLLPRVVELAQ    SSTDRRTKVAASELLHAVSL  1002
Query_CLUBM  827   -QS---[ 4]GLSWDV--VDRIKFTVPLR-D----EkLDIWLD--CLLPRIVELAK    SSSDRKIKVLSSELLHAMCL  889
Query_MOSS_  891   -QL---[ 4]GLAWDP--IDRVKFTLPFR-Q----DkPDIWLD--SLLSRMVEIAL    SSSDRRMKVTASELLQAVCL  953
Query_LIVER  906   -QL---[ 4]GLAWDS--VDRLQFTIPFR-D----EkYDIWLD--PLLSRVVELAV    SSSDRRIKVTASELLQALSL  968
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  940   LMLGRSASQ    P-------TGN   PMAA    IYRKLFPQMLKLAT-DVEQVSRQLFEPLMLQIIHWFTK   SN     995
Query_LINGU  945   YMLGRSVQQ    P-EGQQ--GRH   PLDG    LYRNVFPAMLQLAC-DVELVARQLFEPLTMQLIHWFTR   SK     1004
Query_LIMUL  942   YMIGISVQQ    P-KEWQ--KKY   PMTK    LYQRVFPALLKLAC-DVEQVARQLFEPLCYQIIHWFTN   NR     1001
Query_SPIDE  946   FMIGQGSML    P-DNLQ--EKY   SMEA    LYEKLFPGLLELSC-DVEQVTKQLFSPLVLQMIHWFTC   KK     1005
Query_SCORP  940   FMVGCGAQQ    Q-PELQ--AIY   SMKK    LYKNVLPALLQLAC-DVDEVARKLFEPFTFQLIHWFTT   KR     999
Query_TICK_  897   YAVGTGVQT    S-AERR--AKF   PMVH    LFKHLFPELLHLAC-DVDQVTQRLFRPLVLQLMHWFSS   GS     956
Query_MONOS  871   YILGREAVL[ 1]--EDNK-----[2]-MTD    LWRHIFPVMIQLSC-DLDSVPQQLFELLTLQVVRWFTD   SR     928
Query_ROSET  851   YMLGQGAGS    ---GTH-aAGE[2]-LSP    LYRHVLPAVLRLAC-DVDLVPKQLFSPLVMQIIHWFSS   SR[1]  911
Query_BELCH  962   YMVGQGAQQ    P-GEMQ--AKN   PMTR    LYKKVFPVLMQLAC-DVEQVARQLFDPLMMQLIHWLTN   NR     1021
Query_CIONA  934   YIIGRNAQI    S--NVE--RST[1]SMAS    LYKKLVPVLLQLAC-DVDQVCKQLFDPLIKQMIHWFTN   NK     993
Query_CAECIL 949   YMLGKASQM    P-EKDQ--SLP   PMYQ    LYKRTFPVLLRLAC-DVDQVTRQLYEPLVMGLIHWFTN   NK     1008
Query_HUMAN  940   FMLGKATQM    P-EGGQ--GAP   PMYQ    LYKRTFPVLLRLAC-DVDQVTRQLYEPLVMQLIHWFTN   NK     999
Query_MOUSE  938   FMLGRATQM    P-EG-Q--GLP   PMYQ    LYKHTFPVLLQLAC-DVDQVTRQLYEPLVMQLIHWLTN   NK     996
Query_XENOPU 961   FMLGKATQM    PdDKKT--GSS   PMYK    IYKRTFPVLLRLAC-DVDKVTEQLYKPLVMSLIHWFTN   NK     1021
Query_PLATY  942   FMLGKASQI    P-EGNQ--GSP   PMYQ    LYKRTFPVLLRLAC-DVDQVTRQLYEPLVMQLIHWFTN   NK     1001
Query_TASDE  945   FMLGKASQI    P-EGKL--GPP   PMYQ    LYKKTFPVLFRLAC-DVDQVTRQLYEPLVMQLIHWFTN   NK     1004
Query_CANAR  939   YMLGKASQM    P-EGCQ--GPP   PMYH    LHKRVFPVLLRLAC-DIDQVARQLYEPLVMQLIHWFTN   NK     998
Query_OSTRIC 909   YMLGKASQM    P-EGCQ--GPP   PMYQ    LYKRIFPVLLRLAC-DVDQVTRQLYEPLVMQLIHWFTN   NK     968
Query_TURTLE 948   YMLGTASQM    P-DGPP--GSP   PMYQ    LYKRMFPVLLRLAC-DVDQVTRQLYKLLVMELIHWFTN   NK     1007
Query_GECKO  948   YVLGKAAQI    P-EGQE--VSQ   PMSH    LYKRTFPILLRLAC-DVDQVTRQLYEPLVMQLIHWFTT   NK     1007
Query_DANIO  954   YMVGKGAQM    T-EDDK--SAP   PMYN    LHRKVFPVLLRLAC-DVDQVTRQLFEPLVMQLIHWFTN   NR     1013
Query_SNAKE  948   FMLGKAAQI    P-EGHK--NSP   PLFH    LYKRMFPVLLRLAC-DVDQVTRQLYGTLVMELIHWFTN   NK     1007
Query_STERL  958   FMLGKGSQI    P-EGEK--STP   PMYR    LYKRIFPVLLRLAC-DVDQVTRQLYEPLVMQLIHWFTN   NT     1017
Query_MILII  946   YTLGKGSQI    S-E--G--QQE   TLYK    LNRQLFPVLLRLGC-DVDTVTRQLFEPLVLQLIHWFTS   NR     1003
Query_BAMBO  948   YMLGKGSQI    A-E--G--QQE   TMYK    LNRHLYPVLLRLGC-DVDKVTNQLFEPLVMQLIHWFTN   NR     1005
Query_ALLIG  908   FMLGKASQM    P-EGRQ--GSS   PMYQ    LYKRTFPVLLCLAC-DVDQVTRQLYEPLVMQLIHWFTN   NK     967
Query_GAR__  947   CMLGKGSQI    P-EGEK--SSP   PMYK    LHKRIFPVLLQLAC-DVDQVTRQLYEPLVMQLIHWFTN   NR     1006
Query_TETRA  939   YMIGTQASS[ 1]-qGGQKkrGRD[9]EFSK    IYSKLFPTIFSIGS-NSSHPSCQLMNQLILQISRWFAN   SK     1010
Query_PARAM  889   YMIGDQAVN    ---GTKavAQN[5]EYAN    IYKKMFPVMIEISN-TMDHPCRTYFKEHLTRIVHWFAN   SK     953
Query_SLEMN  917   YMIGKNASA    PkGRGNqqQQE[2]PFAK    IYKKLFPVMIHLAT-EVENISRQLFEPLIFQIVRWFSN   ST     981
Query_STENT  831   YMLGKKTQV    DfN--------   -FAK    IYNNVFPVVYKLAT-DVELIIRQMFEPLSKQLVRWFAR   SF     884
Query_NEMAT  943   YMLGKGVTQ    P-DVKK--VAL   PMQN    LWKKLIPVLLQLAC-DVERVSAQLFHPLVMQLIHWFTN   NR     1002
Query_PISTI  957   YMLGKSVTQ    P-QGFK--KS-   PMEH    LWKHVFPTLLQLAC-DVEQVSEQLFKPLVFQLIHWFTN   NR     1015
Query_CORAL  939   FMLGKSVQI    P-DASS--KS-   PLEK    LWKRVFPVVLQLAC-DVEQVSQQLFRPLLFQLIHWFTN   NK     997
Query_HYDRA  889   FMLGIESQQ    Q--GKR--TSQ[2]SMER    LYKEIFPALIKLSC-SQEQVSEQLFSPLLFQIIHWFTG   RS     949
Query_LATIM  948   YMLGKGSQI    P-ERCG--GVP   PMYR    LYKRIFPVLLRLAC-DVDQVTRQLYEPLVMQLIHWFTN   NK     1007
Query_SHRIM  926   LMIGTGAQK    S-QETQ--LKE   PMAP    LYRHIFSAMLRLAC-DPDHVTRQLFLTLSYQTIHWFTN   NR     985
Query_MCRAB  923   LMIGTGSQQ    Q-EGIQ--AKY   PMAA    LYHHIFQAMLQLAC-DPDQVTRQLFFTLTHQTIHWFTN   NK     982
Query_AMPHI  959   VLIGSGRTA    T-R--------   PSAP    VYRHLFAACLRLAC-DVDHVSRHLFVTLMMQCTHYFSS   NR     1012
Query_BARNA  888   FTLGRGAQL[ 1]-mDGAG-----   SLSA    VYRRLLPALLRLAA-SPSLVARQLFAPLLTQITHWFTT[4]GR     949
Query_COPEP  894   FMIGKTSQQ    -tDDYE----K[2]PMTE    LFRRVFPVILKLCS-DPDSVIQGLFQPLFSQMIHWFTK   ER     953
Query_DAPHN  850   FLTGTHSRR    --------SED[5]PLAP    LYAKIFPTILNLAA-DSDEIIRQLFRALLIQLVHWFSR[4]RK[6]  919
Query_STARF  962   FMLGKGTQQ    P-EQRQ--AKS   PMQH    LYRRVFPALFRLSC-DVEQVARQLFEPLVMQLIHWFTG   NK     1021
Query_STRON  966   FMLGKGSQQ    P-DQRE--AKS   PMNQ    LYKKVFPAVLQLAC-DVEQVARQLFEPLVMQLIHWFSG   NK     1025
Query_SEACU  955   FMLGKGTQQ    P-DQQG--SQV   SMVN    LYKHIFQALLHLAC-DPEQVARQLFQPLVFQLIHWFTG   NK     1014
Query_FUNGI  913   YMLGTNTKS    -------eDHA[2]SYRK    LYDHLFPALLRLAS-DPDPVCRQLFEPLTMQLIHYHTR   PN     970
Query_DPURP  1026  YMIGTS--A    -------nQSI   SFSR    LYRKIFPSLLCLST-DVEQVTKQLFQPLVFQIIHWFTR   NK     1079
Query_HALBU  1036  YMVGTNSGG    -----------   SFHR    LYRRVFPALLRLAV-DVEVVARQLFSPLIMQLVHWFTK   NV     1087
Query_GLOBO  871   VMIGGSQSR    SeGDGR-----[1]PYHK    IYVKLFPVLLRLSV-DLDKVTRELFRPLVIQLVHWLTK   NS     929
Query_PALUS  913   VMIGRSAKQ    LtNVGRmsEAK[2]PFHN    LYRKVFPGLLRLAV-DLDKVTRDLFRPLVFQIIHWLTK   NS     977
Query_CALIF  973   YMVGLNAHR    IrDNSTm-KEN[2]IFNK    FYLKIFPVILRLSV-DIDKISRDLFSSLTFQLIHWFTN   NS     1036
Query_GLOMU  998   VMIGKSAFQ    LsRQSQ-----[1]PFYR    IYRKIFPALLRLAI-DTDPVPRRIFRRLVAQLIHWYTN   NA     1056
Query_LOBOS  949   LMIGSSAFR    -aRDAQ--DSK[2]PFHK    LFLKLFPALLRLAV-DVDQVPRSLFRPLVSQLIHWLTN   NA     1010
Query_RACEM  929   VMIGKSAHR    -pRQEK--SAT[2]PYHR    TYVRVFPVLLRLAV-DIDSIPRSMFRSLVFQLIHWLTT   NA     990
Query_PVIOL  1001  FMIGKSSTA    -------sEEV   SFSR    LYKKIFPGLLKLST-DVEQVTKQLFQPLVFQIIHWFTR   NK     1056
Query_BOMBU  864   YILGKTSQY    L--VS---NPD   QFVS    LYTMLCPALLNLGC-DYEEAARQIFYPLMLQLTHWLSS   KF     921
Query_AEDES  929   YVIG--IQY    Q--DK------   MT-K    LWTKLCDHLLQLAT-DTDIAVCQMFEPLLFQIIHYLTQ   PS     980
Query_ANT__  889   YAIGFTARY    L--TDP-eNRD   KPIT    LYSTLCKAVVTLGC-DSDEVVCGLFHPLALQLMHWLSS   RR     948
Query_WHITE  893   VLLGNDKQM    A--NL---ESN   QTAD    VLKELMPALLQLGS-DSDEVVRTLFQHLVRPMMFYFSS   KF     950
Query_BEETL  878   IFLGTTKQM    S--GS---GIS   DLDE    LLKHVSVSLLRLGC-DMDQVVEQLFRPLCFQLIHWYTH   PF     935
Query_THRIP  875   FMLGKEKQL    E--TA---HQG   HLTS    LWVHLFPKIFSLAC-DVDIIVSQLFKPLGEQLAHWYSS   KL     932
Query_TERMI  910   LILGSGQEA    Q--TS---HPD   KYSE    LFRKLSLPLLKLGC-DPDPAIHRLYNLLVLQLVHWYSS   KF     967
Query_FLEA_  853   YILGTGMQV    N--TA------   QLVH    LYTNIFPTLLQLAA-DSDIIAMQLFDPLMIQITHWFSS   QW     907
Query_MIDGE  893   YIMG--KNL    D--GA------   ETLP    LWKEICKNLIVLAT-DQDQTVRQLFEPLLMQMMHYFAQ   PS     945
Query_LAODE  901   IFLENAPDS    --------GET[3]--HK[ 1]FLRKMLQGMFKLGCkSSDSVGRNIFNNLLHQTCHWFAR[4]NK[2]  964
Query_LOTTI  928   YCLGRGAQQ    P-GEVQ--KRY   PMQK    LYKKLFPAMLKLAC-DVEQVAKQLFEPLTMQLIHWFTN   NK     987
Query_APLYS  945   YALGRGAQL    P-GDRQ--NKY   PMER    LYRKLFPPLFALAC-DVEQVCKDLFEPLVMQLIHWFTS   NK     1004
Query_SCALL  959   FSLGRGATQ    P-GEAA--KRC   SMAP    LYKKLFPVLLQLAC-DVELVAKQLFEPLTMQIIHWFTG   SK     1018
Query_CRASS  940   LCVGRSATQ    P---AR--ERY   SMAP    LYKKMFPSILQLAC-DVDLVTKQLFEPLLMQLIHWFTK   NN     997
Query_OCTOP  912   YSLGKSVQK    G-SSPE--GNY[2]SFEA    IYKNVFPVLLQLAC-DLDQVPKQLFETLVFQLIHWFTS   NK     973
Query_POMAC  940   YTLGRSAQQ    P-GEIQ--KRN   PMDS    LYRKLFPALLSLSC-DVEQVCRQLFEPLVKQLIHWFTG   NK     999
Query_WHIPW  923   FLLGMESYR    S---LS--ASE[2]SIAA    MYGKLLPALLQLTS-DTDEVPRSLFSSLFGQIVHLFAG   TP     982
Query_TRICH  923   FLLGNESFR    T---VK--MKD[2]SISK    AYSHLFPVLLKLAC-DVDEVPKQMFSTLFVQILHLFSG   TD     982
Query_AASTA  849   YLCGMTATL    P--TAKthASE[2]VYFD    HWHGIFPSLVQLAA-DADTITRSLFAPLLAQLLRWFSG   IA[5]  916
Query_AEUTE  833   YLCGKTATL    P--RSKssANE[2]IYYD    HWANVFPSLVMLAS-DADTITRSLFAPLLSQLLRWFSG   IS[5]  900
Query_PLASM  909   YLCGKTATH    P--HT---AGT[2]DFYD    LWREVFARVVRLAS-DSEKICRSLFEPLLFQLLRWLAI   NS[5]  973
Query_BREMI  904   FLCGKTTTH    P--HP---AGK[2]EFYE    MWREVLKRVMQLAT-DPEKICRSLFEPLQFQLLRWLTA   NS[5]  968
Query_OLIGA  905   YLCGKTATN    P--HT---SGS[2]AYFE    VWASVFSRLIILAT-DAEKVCRALFEPLLFQVTTWFSK   GS[5]  969
Query_LIGNO  983   LMTGRCAFA    StSGND--REQ   -LLT    LWRRLLPALLPLAA-DEDPVVKQLFQPLLMQLVHWFSR   SR     1042
Query_CESTO  857   FALSTQLSA    ---------SN[2]-VEN[ 4]YWQLLFQLVFILGA-DDDPLIRRLFRSLAFQLVHWYAG   YL[4]  919
Query_OVI2_  950   FMVGKAANT[22]PfDQSL--QGH[9]PVRS[14]LWRQLFLASFYVAA-NKMPPISRLFSNLLFQLTRWFGE   YA[2]  1057
Query_PRIAP  956   YILGTGAQQ    P-DSIK--SKR   PMSA    LYSRLFPVLLRLGC-DVEPVAKQLFEPLVMQTIHWFTL   NR     1015
Query_ROTIF  951   YTIGKSVSE    P-EGYS--QSY   QMTK    IYHRLYPCLFGLSC-DTDHFVRNLFQPLVMQMIHWFTG   NR     1010
Query_SPONG  948   FVIGKGTQQ    S---VQ--KKN   PMTK    LYQKLIPVALQLAC-DVDRVAEQLFRPLVFQMIHWFTR   NT     1005
Query_TARDI  957   ILVGTHSEV    SaDKRE--KAI   -LMD    LINKLMPKILELAC-DVEPVTKQIFEPLSVQMVHWFSR   TA     1016
Query_TRICH  925   YMIGRGAQQ    P-GGTG--SVT   NYGN    LYKKIFPAILTLAS-DVEQVSQQLFQPLVYQLIHWFTG   NS     984
Query_OSTRE  931   FLIGKSAHA    PsHESGdfTRT[2]QYRP    IFEKLFPIIFELAV-DTEPVARQLFNSLSFQIVRWYAK   NQ[2]  997
Query_NITEN  1003  YMIGTSAKA    PlQRRGeeLRE[2]QFQK    IYRKLFPALLKLAT-DVEPVTQRLFSEFVKQIIRWFTS   SK     1067
Query_CLUBM  890   LMLGNAAKG    PlRTRNdeTNL[2]-YVQ    LYRKLFPSVLWLAT-DVEPVTQQLFSEFAMQLIHWMTS   NT     953
Query_MOSS_  954   LMLGNAAKG    PmRRRGedSQS[2]-FDK    IYKRLFPAILRLAT-DVELVTRQLFLDFVKQLIHWFTS   NQ     1017
Query_LIVER  969   VMLGNSAKG    PlLRKNekDQS[2]-LDK    IYKRLLPAVLRLAV-DVEQVTRQLFNEFVKQLIHWFTS   NL     1032
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  996   NSHRPDTEALLNSLMDGIV--S   PSDS--AV-RDF    SAKCIREFLAWSIKQ--T--SP   KEMEAG---PA      1050
Query_LINGU  1005  VFENPETAALLNAVMEGLI--D   PSNA--SL-RDF    CARCIEEFLKWALKQ--T--SP   AQLRKN---PH      1059
Query_LIMUL  1002  GLMSEETMTLLQALWDAIT--E   REDT--WV-RDF    SGNCLKEFVLWSIKQ--T--AP   KEQEKS---PV      1056
Query_SPIDE  1006  VSSTPESAVILLSLWDGVT--N   PKSA--AL-RDF    SAICLKEYFLWSIKQ--S--SA   ---ESK---PV      1057
Query_SCORP  1000  MINSDETKQFVQVIWDGLV--Q   SKDA--VF-RDF    SARCLKEFVQWSIKQ--T--SV   REQEEN---PI      1054
Query_TICK_  957   TRGSRESVAIVECLWDTVT--Q   PQET--VL-KDF    AARCLREFVQWGIKQ--S--TP   KELEKG---PT      1011
Query_MONOS  929   MYEYPNTVALITAILDALC--D   ASNT--NL-REF    AARCVAEFFTFALRRKgAe-NP   ---ERKt---A      983
Query_ROSET  912   --EREEVVLLLDAILTGLA--D   AKNS--AL-RDF    CATCVKEFYVWAYKHS-Td---   KLLETGg---L      964
Query_BELCH  1022  KFESEETITLLDAIMDGLV--H   PTDT--AL-RDF    SSRCVREFLQWSIKQ--T--TP   KQLEKS---PI      1076
Query_CIONA  994   KFESKDTAILLNSIMEGLV--D   EENA--TL-REL    CADSVREFLVWSIKQNpS--AE   KQVRNN---A-      1049
Query_CAECIL 1009  KFESQDTVALLEAILDGIV--D   PVDS--TL-RDF    CGHCIREFLKWSIKQ--K--TP   QQMEKS---PV      1063
Query_HUMAN  1000  KFESQDTVALLEAILDGIV--D   PVDS--TL-RDF    CGRCIREFLKWSIKQ--I--TP   QQQEKS---PV      1054
Query_MOUSE  997   KFESQDTVALLEAILDGIV--D   PVDS--TL-RDF    CGRCVQEFLKWSIKQ--T--TP   QQQEKS---PV      1051
Query_XENOPU 1022  KFESQDTVALLEAILTGIV--D   PVDS--TL-RDF    CGQCIQEFLRWSIKQ--T--TP   DQQAKS---PV      1076
Query_PLATY  1002  KFESQDTVALLETILDGIV--D   PVDS--TL-RDF    CGQCVREFLKWSIKQ--T--TP   EQQEKS---PV      1056
Query_TASDE  1005  KFESQDTMTLLETILDGIV--D   PVDS--TL-RDF    CGQCVREFLKWSIKQ--T--TP   QQQEKS---PV      1059
Query_CANAR  999   KFESQDTVAFLEAMLSGIV--D   PVDS--TL-RDF    CGQCIREFLKWSIKQ--T--TP   RQQEES---PA      1053
Query_OSTRIC 969   KFESQDTVAFLEAILGGIT--D   PVDS--TL-RDF    CGQCVREFLKWSIKQ--T--TP   KQQEKS---PV      1023
Query_TURTLE 1008  KFESQDTVALLEAILDGIV--D   PVDS--SL-RDF    CGQCVREFMKWSIKQ--T--TP   QQQEKS---PV      1062
Query_GECKO  1008  KFESQDTVTLLEAILDGIV--D   PVDS--PL-RDF    CGRCVREFLKWSIKQ--T--TP   QQQERS---PV      1062
Query_DANIO  1014  KFESQDTVAVLEAILDGIV--D   PLDS--TL-RDF    SGTCIQEFVKWSIKQ--T--TP   KQQEKS---PA      1068
Query_SNAKE  1008  KFESQDTVTLLEAILDGIV--D   PVDS--TL-RDF    CGQCVHEFLKWSIKQ--T--TP   QQQEKS---PV      1062
Query_STERL  1018  KFESQDTVAMLEAILDGVV--D   PVDS--TL-RDF    CGQCVQEFLMWSIKQ--T--TP   KQQEKS---PA      1072
Query_MILII  1004  KFESRDTVAMLEAILDGIV--D   PVDS--TL-RDF    CGQCVREFLKWSIKQ--T--TP   EQQAKS---PA      1058
Query_BAMBO  1006  KFESRDTVALLEAILDGIV--D   PVDS--TL-RDF    CGNCVREFLKWSIKQ--T--TP   EQQQKS---PV      1060
Query_ALLIG  968   KFESQDTVALLEAILDGIV--D   PVDS--TL-RDF    CGQCIREFMKWSIKQ--T--TP   QQQEKS---PV      1022
Query_GAR__  1007  KFESQDTVAILEAILDGIV--D   PVDS--TL-RDF    CGQCIQEFLKWSIKQ--T--TP   KQQEKS---PA      1061
Query_TETRA  1011  QLESPEVAELLDNIMNFISskD   -NNN---L-REL    AANCLGQFIKWTLKQN----TE   QEIKVNh----[1]   1065
Query_PARAM  954   QYEAPDVAQLLDCLFEQPI--N   KEDP--QL-EQF    CSQLIGEFIQWTLKQT----TP   DQLKLNy----[1]   1008
Query_SLEMN  982   VYDHPEVESLLEALIDGAS--S   RNNS--AS-RDL    CSNAVAEFAKWSLKQM----TD   KNIKENpn---      1036
Query_STENT  885   SYENPETMALLDALVDGVC--N   QSNN--MQ-RVL    ASECIGEFCR------------   DVIMYNsnsIA      934
Query_NEMAT  1003  KYESPETIALLDALLEGVV--H   PTDT--SL-RDF    SAQCIREFLQWSIKQ--T--SA   KQLEKS---PI      1057
Query_PISTI  1016  KYETPETMALLDAIVDGIV--H   PVDT--SL-RDF    SAQCIKEFLQWSIKQ--S--SE   KQLEKS---PI      1070
Query_CORAL  998   QYESPDTIILLETLLDGIV--H   PTDT--AL-RDL    SALALEEFLRWSIKQ--A--TK   NQLEKS---PI      1052
Query_HYDRA  950   KYESPETIVLLDAILDGVS--D   PDNS--AV-RDF    SAKCVKEFLTWSIKQ--T--TE   KQMESN---PG      1004
Query_LATIM  1008  KFESQDTVAMLEAIMDGIV--D   PIDS--TL-RDF    CGLCIQEFLKWSIKQ--T--TP   KQQEST---PA      1062
Query_SHRIM  986   NFENVDTMVLLEALMEGIV--T   ANDP--AL-RDV    SAQCLAEFVKWSRKQC-R--QQ   NSVSS------      1038
Query_MCRAB  983   KNENPDTAVLLESIMDGIV--T   ANNP--AL-RDV    SAQCLAEFVIWSRKQS-R--EK   NQASL------      1035
Query_AMPHI  1013  GYEHPDTVALLDAIMEGLE--S   VNNP--AL-RDF    CGRCLCEFLVWSRKQN-V--QA   RTSSRAtgsAA[7]   1078
Query_BARNA  950   ----PESALLLDVLLEGLV--D[2]--DP--AR-RDV    SADALREYLQWSLKQRrAd---   ---------PV[1]   996
Query_COPEP  954   KCGSLETNCLLKVLLDTVCe-D   -EDV--SRkQDA    -AGYIQEFMVWSIKQSgSs---   ---SKNi----      1005
Query_DAPHN  920   TVLGAEAALLVEVLMESACr--[1]-LDA--SLvRDL    ACECLLEFLKWTIKQSsDt-TL[1]-------edPA      975
Query_STARF  1022  KFESEETTALLNAMLDGVI--H   PTDT--SL-RDF    SACCLKEFLKWSIKQ--T--SK   KQQEKS---PI      1076
Query_STRON  1026  KYESDETSTLLNAILAGIV--H   PTDT--AL-RDF    CARCISEFLRWSIKQ--T--SK   KQQEKS---PI      1080
Query_SEACU  1015  KFESEETMTLLNAILEGII--H   PTDT--AL-RDF    SGQCVKEFLKWSIKQ--T--PK   KQLEKS---PI      1069
Query_FUNGI  971   LAESEDTSCLLEAVSESLS--S   R-DA--AL-RDT    AGKCVSEFIKWSVKHS----SK   KQMESNv---G      1024
Query_DPURP  1080  KQENDDTMNLLNSIVDAVG--N   PYDG--AR-RDF    AGKCLCEFAKWSLKNT----SV   VQQEKNa---F      1134
Query_HALBU  1088  RAENDETMALLNSIIDAVN--G   E-DG--AL-REF    GARALGEFASWSLKHApTtkST   ANEARSa---F      1145
Query_GLOBO  930   KYENPETMALLQSCFDAVS--S   S-NG--PL-REF    GAECIAEYLKWSIKHAsE----   KELENNp---I      983
Query_PALUS  978   KAENAETMAMLNACVDAVS--G   D-DG--PQ-REF    AGECIAEFLAWSIKHTlP----   KELDQNp---M      1031
Query_CALIF  1037  LNDDSESLILLDCCFSAAI--C   Q-DG--PL-RDY    GTQCINEFLKWTMKQNfE----   ENKL--i---E      1088
Query_GLOMU  1057  QYENQETMALLQCCLDAIC--D   T-WG--SL-RDY    GAECLNEFLIWSIKQSsA----   THAEESs---M      1110
Query_LOBOS  1011  QYENPETIVLLNTCIDATC--D   T-MG--PL-RDY    GAECLGEFVKWSIKQSsS----   SSS--Gs---A      1062
Query_RACEM  991   QAENPETIALLEACLEAVC--D   S-NA--SL-QSY    GAECLQEFVEWSIKQSgS----   RQSG-Nt---M      1043
Query_PVIOL  1057  RQESDDTMTLLNAIIDAVG--H   PLDG--AL-REF    AGKCLSEFTKWSLKNT----SQ   KQQEKNp---F      1111
Query_BOMBU  922   MLESLTTTYFLDLLFKGLC--D   EANS--SL-REF    SGICLSEFIKWSIKQSdNa---   AISQS------      974
Query_AEDES  981   KIGLKGTEVLVNCLMESIS--H   QTDT--GV-RDL    AARSIREFLLWTIRQTpAd---   QRSALSsssSA      1039
Query_ANT__  949   MPMSPV----IDLLFDGLT--D   DSNP--AL-REF    SGMCLAEFTRWSIRQNtFs---   ----E------      993
Query_WHITE  951   QAHTNSVKVVFEALMDGIT--N   PTDS--AV-RDF    TAECIKGFTQGAIKQSkSs---   N-EKV----PL      1004
Query_BEETL  936   QSRSGHEAIVIEAILESMT--H   PTDT--AL-RDF    AGKCISEFVKWTIKQSnEk---   NLSKD----PI      990
Query_THRIP  933   NEKSEQADAFLRAILDGIT--H   QTDA--AL-RDY    SATCLHEYVKWSRKQGnMm---   A-K--------      982
Query_TERMI  968   KLKSKEVEILIQTLMDGIT--H   PTDT--AL-RDF    SARCLHEFVVWSVKQGsHv---   --N--------      1016
Query_FLEA_  908   KIKSIETSLLIDQLMIGTS--N   ETSA--SV-RDI    SAKYLNEFVKWSLRQAdPe---   DLRS----sSV      962
Query_MIDGE  946   QILSSLTTVIIEVLMEMIS--Y   KKSS--GV-QDL    SARLLREFILWVNKTTnRn---   QRQA----sPI      1000
Query_LAODE  965   ------VFIFENSLLEALTc--   PTDI--EV-REF    AAECLHEYVSWTFKPSdEe---   RSLSKAvkqDI[3]   1020
Query_LOTTI  988   MGESSETMALLDTIFDSII--Q   ADDT--SL-RDF    SGKCLREFLHWSLKQ--S--SV   KTVEKS---PV      1042
Query_APLYS  1005  MSESPETMALLDCIYDGLV--Q   PSDP--SL-RDF    SARCLREFLKWSLKQ--I--SK   KAAEKN---PI      1059
Query_SCALL  1019  KAESVETMSLLDAIFDGII--Q   PTDT--SL-RDF    CATCLREFMKWSVKH--T--S-   KKMETS---PI      1072
Query_CRASS  998   KAESEESMALLDAIYDGII--Q   PNDT--AL-RDF    CAVCLREFLKWSIKQ--T----   KNVEKS---PI      1050
Query_OCTOP  974   KFENAESMALLDAIYDGII--Q   SEDT--GL-RDF    CAKCIAEFFKWSLKR--K--SS   ---EKN---PV      1025
Query_POMAC  1000  MAESPESMALLDSIFDGLV--Q   QDNT--ML-RDF    AAQCLREFLIWSIKQ--T--SK   KAMETN---PR      1054
Query_WHIPW  983   FASEATALVLLDVIIDNLC--N   TTSS--SV-RAQ    SAAYLTEFLVWSMNGT-A--SR   KEATHRl-dKA      1040
Query_TRICH  983   FANKMVTSCLLDTLMDGLC--C   FWNS--SL-RDQ    SAYYISEYCRWSMRKV-D--SG   GQNMRSe-sKV      1040
Query_AASTA  917   -----EAHAFLDAIMHGLS--C   VDDGsgGV-RDV    AAQSVATFLKYASKQPpTs---   ------ranTV[1]   967
Query_AEUTE  901   -----EAQIFLDSLIDGIS--S   ADD--gGA-REV    CALSLSTFLKYASKQP------   -------gsTS[1]   945
Query_PLASM  974   -----EYSSMLDELTRSMS--D   -SEA--AV-RTL    SARCIATQLSLAVDTR------   --------lPQ[1]   1016
Query_BREMI  969   -----EYAGMVDILVSNFS--D   -PET--AV-RTL    SARSIAVLLSLAIKN-------   -----------[1]   1007
Query_OLIGA  970   -----EYSRMLDELVIGLS--D   RTNT--AV-REV    SGKCLAACLSSSIEAS------   --------gSV[1]   1013
Query_LIGNO  1043  EFEGPYTALLLDSLFDRLTa-S   SGDA--SV-RDL    AAQSLREFLTWSVRQ-------   -QMPSSgadST[6]   1103
Query_CESTO  920   ----SEAKLLLKTLLRLLKliE[7]IIEStwPLpESI[ 7]ASVCLRDFLYWMNFEApGt---   KRRRSSl----[2]   991
Query_OVI2_  1058  --NHQHAVCFFKVLLDILClgD[8]QSPSmaTL-SDA[13]AARCLYKFLIFSVRRArVssSP[7]TNQTGQdrqPS[1]   1150
Query_PRIAP  1016  MFESPETMCLLDAIMEGIV--D   ATDA--AL-RDF    CGSCIREFLIWSIRQ--T--TV   RQQERS---PG      1070
Query_ROTIF  1011  KYESSDTIQLLGCIMESLV--D   EKNA--AL-RDF    SVNCLKEFLKWSIKH--T--PL   SKQENSaptPV      1068
Query_SPONG  1006  QYESPDTVTLLEAILEGIE--H   PTNA--SL-RDF    SAQCIKEFLVWSIKQ--T--SK   KQQERS---PM      1060
Query_TARDI  1017  QTASAEATAVLDCVMDGLV--S   EGNP--TL-RSF    CARCLNEFLGWSGKNL------   PANSKAv--PP      1070
Query_TRICH  985   NYESVETIALLNALLDGVI--H   STDS--AL-RDF    CAKSLREYLHWSIKQ--T--TK   KQQEKT---SN      1039
Query_OSTRE  998   --ELDVTLCLFDAVLTGLS--N   SSKA-gAV-RDL    CASLTAELLDWSLRYVrD----   ---DKTfegSV      1051
Query_NITEN  1068  TREHPETMALLDAILEGLV--D   EDNG--AL-REF    SAGCFQEFMEWSIRHApP----   --EADNf---V      1120
Query_CLUBM  954   QKENPETMALLDAILEGLV--D   AENG--SM-REL    CSRCFSEFLRWSIRHApP----   --EASNf---V      1006
Query_MOSS_  1018  QKENPDTMALLDAILDGLV--D   PDNG--SL-REF    CADMFSEFLQWSVRHApP----   --EDDNf---V      1070
Query_LIVER  1033  QKENPETMALLDAILDGLV--D   AENG--SL-REF    CASCFGEFMSWSVRHTlP----   --ENDDf---V      1085
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1051 NPKRL---LKRMYSLLEHPSSMMRLGGALMFNS-IYSVFREE-DALVDQFSLQILSVLVNSLALAHR   DDKSLG   T  1119
Query_LINGU  1060 IPKSI---FKRIYSFALHPSPFRRLGAALAFNH-IYRVFRED-EVLLDIFTFEILETFIESLAMAHH   DDKSLG   T  1128
Query_LIMUL  1057 NIKAL---MKRLYSYCLHPNSFKRLGAALAFNH-IYAILREE-EPLVNMFILETLVHFVNSLAIAHK   DEKSLG   T  1125
Query_SPIDE  1058 --ERL---LSKLHSFSLHPNSFKRLGAALTFNS-LYNVFRNE-DFLVRKYTLQLLVLFVDSLAIAHT   DSSSLR   T  1124
Query_SCORP  1055 GMNVV---LNCLYNYSLHPNPYKRLGAALAFNS-IYTIFREE-ESLVDKFTLEIAVHFIESLAIAHY   DDKILG   T  1123
Query_TICK_  1012 NVQSV---LRRLCSYCRHPSALKRLGAALVFNN-LYVLLREE-ESLVDVFILDILAHFVDSLKLAHS   DERTIG   T  1080
Query_MONOS  984  -LTSL---LLRIRAMAAHPSIFQRLGAALAMRT-IYRKLRES-APLVSVFGLDLTVAFLTGLQHCEN   DPEGAG   T  1051
Query_ROSET  965  -VKVL---LKRIMSMAAHPNAFQRMGAALAMNN-IYTRLRES-DTLVDVYVCELTDVFMRSLQLAHR   DDPAMG   T  1032
Query_BELCH  1077 NIKSL---LKRLYSLALHPNPFNRLGAALGFNS-IYRDFREE-SSQVDMFTFELLVTFVESLALAHH   DEKSLG   T  1145
Query_CIONA  1050 --KSL---FKRLHSLALHPTPMKRLGAALAFNH-IYRVFREQ-NTLVDEFIFETLVTFINSLSLSHV   DEMFIG   T  1116
Query_CAECIL 1064 NTKSL---FKRLYSLALHPNAFKRLGAALAFNN-IYREFREE-NALVDQFVFEILVVYLESLALSHQ   DEKSLG   T  1132
Query_HUMAN  1055 NTKSL---FKRLYSLALHPNAFKRLGASLAFNN-IYREFREE-ESLVEQFVFEALVIYMESLALAHA   DEKSLG   T  1123
Query_MOUSE  1052 NSKSL---FKRLYSLALHPNAFKRLGAALAFNH-IYKEFREE-GSLVEQFVFEALVTYMESLALAHE   DEKSLG   T  1120
Query_XENOPU 1077 NTTSL---FKRLYSLALHPNAFKRLGAALAFNN-IYRDFREE-TALVENFVFEVLVIYMESLALSHA   DEKSLG   T  1145
Query_PLATY  1057 NTKSL---FKRLYSFALHPNPFKRLGAALAFNN-IYREFREE-DSLVDQFTFEALVIFLESLALTHA   DEKSLG   T  1125
Query_TASDE  1060 NTKSL---FKRLYSFALHPNAFKRLGAALAFNN-IYREFREE-EALVEQFVFEALVIYMESLALTHA   DEKSLG   T  1128
Query_CANAR  1054 NTKSL---FKRLYSLALHPSAFKRLGAALAFNS-IYREFREE-NSLVEQFVFEALVVFLESLALAHT   DEKSLG   T  1122
Query_OSTRIC 1024 NTKSL---FKRLYSLALHPSAFKRLGAALAFNS-IYREFREE-NSLVEQFVFEALVVFLESLALTHA   DEKSLG   T  1092
Query_TURTLE 1063 NTKSL---FKRLYSLALHPNALKRLGAALAFNN-IYREFREE-NSLVEQFVFEVLVVYLESLALTHT   DEKSLG   T  1131
Query_GECKO  1063 NTKSL---FKRLYSFALHPNAFKRLGAALAFNS-IYREFREE-NALVDQFVFEALVLYLESLALTHG   DEKSLG   T  1131
Query_DANIO  1069 NMKSL---FKRIYSLALHPSVFKRLGAALAFNS-MYRQFREE-SSLVEQFVFEVLVVFVESLALAHF   DEKSVG   T  1137
Query_SNAKE  1063 NTKSL---FKRLYSLALHPNAFKRLGASLAFNN-IYQEFREE-NALVDQFVFEALVVYLESLALTHG   DEKSLG   T  1131
Query_STERL  1073 NTKSL---FKRLYSLALHPNVFKRLGAALSFNS-IYRQFREE-STLVDQFIFELLVIFVESLALAHT   DEKSLG   T  1141
Query_MILII  1059 NTKSL---FKRLYSLALHPNTFKRLGAALAFNS-LYREFRQE-SALVDQFIFEVLVVYVESLALAHA   DDNSLG   T  1127
Query_BAMBO  1061 NTKSL---FKRLYSLALHPNTFKRLGAALAFNS-LYREFREE-SALVEQFVFEVLVVYVESLALAHA   DDKSLG   T  1129
Query_ALLIG  1023 NTKSL---FKRLYSLALHPNAFKRLGAALAFNN-IYREFREE-NSLVEQFAFEALIVFIESLALTHT   DEKSLG   T  1091
Query_GAR__  1062 NTKSL---FKRLYSLALHPNVFKRLGAALAFNN-IYKQFREE-SSLVDMFVFEALVIFVESLALAHS   DEKSLG   T  1130
Query_TETRA  1066 NIKSA---LRRIISFSGQPDLSKRFGALLALKE-VVNNIKNE-SFLVDKFIMETIQT---SMRIVQK[1]-----G   V  1127
Query_PARAM  1009 NIKTV---MRRIITYANHPDFRYKSMSLECILQ-FVKHTQQE-KFIVSKYFIEIV----EQLKFIFI[1]EDKDIS[1]I  1075
Query_SLEMN  1037 NIKSL---IRRIESNSNHPDPFKRLSSVLCFSK-IFSVIREH-DPLIDRFCMEIAHNVLGSLKMCHN   SIEFSQ[1]V  1106
Query_STENT  935  NFKSA---FRRVQALSNHPQSEKRRGAIMCLKK-ILAPLSLN-EVLIDRFLLEFAHVVFVTTRLSHY   D-KSSE[1]T  1003
Query_NEMAT  1058 NIKSL---LKRLYSLALHPNASKRLGASLAFNS-IYTVFREE-RSLVDVFVLEILVTYLESLALAHT   DDPAQG   T  1126
Query_PISTI  1071 NIKSL---LKRLYSLALHPSASKRLGAALAFNN-IYTVFREE-ASLVDIFVLEILVNYLECLCLSHG   DEEARG   T  1139
Query_CORAL  1053 NIKSL---FKRLYSLATHPSANKRLGAALAFNN-IYRVFREE-SSLVENFTIEILVTYLESLEMAHE   DEISLG   T  1121
Query_HYDRA  1005 NPRRL---FKQMISLACHPNPYKRISAAIVFNN-CYTVFREE-EILVNTFILELLLVYINGLSLSHK   DEKSFG   S  1073
Query_LATIM  1063 NTKSL---FKRLYSYALHPGAFKRLGAALAFNN-LYKEFREE-SALVDKFVFETLVVYMESLALAHM   DEKSLG   T  1131
Query_SHRIM  1039 NIKSI---IKRIFSFCKHPSAFKRLGGALAWNS-IYREIRED-GQAVDIWTLDLLAHLMTSLDLAQV   DDPALG   T  1107
Query_MCRAB  1036 NLKSI---LKRIFSFCKHPSAFKRLGGSLAWNS-IYREVRED-SLAVETWTLEILSHLMSALDLAHN   DDLVLG   T  1104
Query_AMPHI  1079 NLKSV---MKRIIHLMRHPSAFKRIGGCVAFNA-VYTEVRES-ETAVNVWLLELFAAVAACLTLADK   DHAALP   C  1147
Query_BARNA  997  KVAPL---LEALYGHLRHPDPYRRLGATLAFNN-LYRVFRES-DRLLDVFLLEVTAALLAALARAHA   DPSALG   T  1065
Query_COPEP  1006 NIKTV---LSRLEYLWKHPDIGKRLGGSLVFNS-IYRIFREE-KKLVSEHLLVILSSIFTAIKISHN   DPLFCG   T  1074
Query_DAPHN  976  TPEFL---FQKLRASATHPSHHQRHGAAIVFNK-LYRVMRES-DPLTSIYTLDLLTHFVTSLSLGER   DPFSSE   I  1044
Query_STARF  1077 NTKSL---LKRLYSMALHPSSFKRLGAALTFNH-IYTVFREE-DSLIDQFVCEILVTFVNSLALAHH   DDKSLG   T  1145
Query_STRON  1081 NTKSL---LKRLYSMASHPSAFKRLGAALAFNS-IYTIFREE-EVLVDQFTFELLVVFVESLALAHH   DDKSLG   T  1149
Query_SEACU  1070 NTKSL---LKRLYSMALHPSALKRLGAALAFNN-IYTVFREE-DSLVEQFTFEILATFVQSLAMAHH   DAKSLG   T  1138
Query_FUNGI  1025 AVKGL---FKRLFGLSRHPNPYKRLGSILAIES-SMEMLKSE-SALVDQFAMEILHHMIVCIRLCR-[1]DDWSLG   V  1093
Query_DPURP  1135 NFKSI---LKRVYSLAHHPDPFKRLGAAISFNE-LYRVFREE-DSLVDQFIFEILHNIFFSLRI---[2]DNSSDD[4]I  1206
Query_HALBU  1146 NFKSL---LKRVYSLAAHPSAQTRMGAAASLLS-LIRVMRED-DNLISVFLFEIAHQVLLSMRLCGA[6]DTTTLE[2]V  1222
Query_GLOBO  984  NAKSL---FKRLYLLCQHASSAKRLGASIIINR-IYTIFRES-NALVDVFTMEILYHLLFSLKISDR   DNPALG   T  1052
Query_PALUS  1032 NAKSL---LKRMYRLLSHSSSHKRLGGAVTFNR-IYRILREE-ESLVDQFSFEILYHLLMALRLAEN   DHTGLG   T  1100
Query_CALIF  1089 NIEVL---LEKMYQYCRNPNPYKRLGASLIFNK-IYVELREH-QKIVDKFLFEILYYILYNLKLSYY   DKDEIE   T  1157
Query_GLOMU  1111 NMKSL---LKRLYNLCAHPDYTKRLGASLAINN-IYRVFREE-EYLVDQFTFELLYWMLVNLRLSDN   DQSATG   T  1179
Query_LOBOS  1063 NVKSL---LKRIYSLASHSNPTKRLGASLIVNR-IYRVFREE-APLVNQFTMELLYWLLFSLRLAET   DHSGLG   T  1131
Query_RACEM  1044 NIKSL---LKRLYNMASNPRSTQRLGAAIVINH-IYRIYREE-ESVVDEFTLELLYWMLFSLRLADS   DHPSIG   T  1112
Query_PVIOL  1112 NFKSI---LKRIYSLAHHPNSHKRLGAAISLNE-LYRVFREE-DILVNHFIFEIMHNILFSLRI---[2]DNADDS[2]I  1181
Query_BOMBU  975  NIYKV---INKITNFALHPSTSKRIAGATAFNH-IYSILRED-EKTVSIYWLEILYSFIKSLDGCSD[1]------   -  1037
Query_AEDES  1040 NLSVM---LEKLRTFSLDSNPNRRLGAALAFNN-IYRILREE-DAQIERCWFDLFYVFCMNFVMTED[1]DGSTTN   -  1108
Query_ANT__  994  HIKEI---IERINNLALHPLTRKRIAAAIAFNH-LYMILRED-DDTISIYWLEIFYSFVCSLDGCND[1]------   -  1056
Query_WHITE  1005 YIKGI---ISRINLYCVHPCKFKRLGGAMIFNS-IHSCLREE-KPIINEHWLELMYYLVKNLSLSEE[1]NDCS--   -  1071
Query_BEETL  991  NIKIL---VKKMRFYSMHPDSNKKLGAALIFNS-IYRDIREE-LSLLSMFWIEILHIFVKSFSSLEN[1]GIDENL   -  1059
Query_THRIP  983  -VAPV---IKFVLSMFRHPCPQKRMGASLAFNS-MYAVLRED-LKIVDEFWIEILYGLVTGLSSCSI[1]GEDSSA   -  1050
Query_TERMI  1017 -IKYV---MKQIYSFCLHPCPSKRLGAALVLNS-VYTVLREE-ESVVNQFWLELLYHMVMNLSLSEE[1]DSILGG   -  1084
Query_FLEA_  963  NLEAL---ISKICDYSINSSVHKQLGAALTFNN-LYMIFREN-DNIVDQFWIQLSTYFVQSLMVNKD[1]----TN   -  1027
Query_MIDGE  1001 KLVDL---FFELRKMSIETNVSRRIGATLAFNN-IYRIVRED-EALIDVYWIYLLDTFAINFKNSED[5]DDNSEQ   -  1073
Query_LAODE  1021 EIDYI---LSKITSYCDHPCPFKRQGAAMGFNR-IYKELREQlTSIVDIFWLELFCAFVKNLLLEHP[1]------   -  1084
Query_LOTTI  1043 NASSI---LRRIFSFALHPSVFKRLGAALAFNN-IYQIFREE-KSLVDRYTFQILVHFVESLSLAHK   DDKSLG   T  1111
Query_APLYS  1060 NAKSV---LKRMFSFSLHPSAFKRLGAALAFNN-IYMVFREE-NALVDRFTFQILVNFVESLAMAHT   DEKSLG   T  1128
Query_SCALL  1073 NAKSV---LKRLFTYALHPSAFKRLGAALAFNN-IYTVFRED-NTLVDTFTFQTLVHFVESLAIAHK   DEKSLG   T  1141
Query_CRASS  1051 NAKSL---LKRLYSYALHPGAFKRLGAALAFNN-IYTVFRED-NAMVNLYTFQILVHFVESLAIAHE   DESSMG   T  1119
Query_OCTOP  1026 NVISV---LTRLYSYTLHPDPFKRLGAALAFNN-FYTIFREH-DILIDQFVFNLLVHFVKSLALCHK   DDQSLG   V  1094
Query_POMAC  1055 NVTSI---LKRMQSFALHPSAFKRLGAALAFNN-IYMVFREE-QSLVDRFTFQILVLFVESLGIAHQ   DEASLG   T  1123
Query_WHIPW  1041 --DLL---VDRLEDMALHSNWTKRFAAAFALNNgIVKTLRDN-KHCTDKYTLELFFVLMRSLIVATE   KEDYYG   V  1108
Query_TRICH  1041 --ELL---VQTIITFAVHSSAVKRAGAAVAFNN-LYKIIREK-RWLVDIYILEMFFTFMKSLYLAEN   DDELYG   T  1107
Query_AASTA  968  SAHAL---MERLFALCGHPLLSCRVGAALAINH-MYRDFREDd-DLVRTYALSMAKNVLVALKPPDD   ------   -  1029
Query_AEUTE  946  SAHTL---FERLFALSSHPVRSCRLGIALVVNQ-FYRDFREDq-EMVDIYAISIQKYLLLAIK-SEG   ------   -  1006
Query_PLASM  1017 KAHYI---FNRVFSLCRHPGSVQRSGAAVSISY-FLQSLREEdGEALTRFAMPCLKNLLYALRLCDS[2]-QYQIG[4]A  1091
Query_BREMI  1008 KTDGI---FKRIFSLCRHPGAMQRSSAALSISC-FLRSLNEEdGKTFSQCALSCLMNLLYALRLCDR[2]-KNRIG[4]S  1082
Query_OLIGA  1014 KAEVL---FHRLFALCRHPGPAERIGAATTMSY-FLRSLNEEnSAVISAFALRCMKNLLYSLRLCDD[2]-GGGNE[4]A  1088
Query_LIGNO  1104 NVKSI---VKRVCLLALHPASSKRLGAALAFNA-VYAQFREE-DSLVNAFAFELLHCFMRSIQLADA   DDASLG   T  1172
Query_CESTO  992  --ESLddlLRQILD-----GLCRQGSASVVFKDvVAPLLTGR-PELTRKYIYVLTDAFTRPGQCFSP[3]---KLA[1]-  1057
Query_OVI2_  1151 SDETFllvLSRVHSSLTGPQPEGLLGALHALEKvILPILCKD-QELVEVHVFQIFAHLLDLIVFVQA[4]TEDINN[4]V  1231
Query_PRIAP  1071 NTKSL---LKRMYSLALHPSAAKRRGATLVFNN-IYTVLREE-EMLVDTFIIEILVTFVDSLALAHN   DEKSLG   T  1139
Query_ROTIF  1069 NAKSI---LKRIFNFLTHPNANKRLGACLAWNS-VYTIFREE-ESIVNRHIFEILFYLIESLAMSEN   DDKMFG   T  1137
Query_SPONG  1061 NTKSL---LKRLYSMSGHPSTAKRLGAALAFNS-IYTVLREE-ESLVDVFIIEMLVVMVNSLKLAHK   DAQSIG   T  1129
Query_TARDI  1071 AATAL---LERIVSYAGHSSCAKRFGAALAFSH-CYKTFREN-SAVVDVFTIDLIVTFTKSLALCHR   DPPGFG   S  1139
Query_TRICH  1040 NAKSL---LKRLYSLASHPSAIKRLGAALAFNN-IYVVLREE-STLVDTFIFEILVNYVNSLALSHT   DEKSLG   T  1108
Query_OSTRE  1052 NHRFI---LRALFGLIGHSSAEHRLGAVTAIRRcLKHIAKNQ--FLIDEYVLKMLDVTLKSMSRSDE[3]GSSMRG   A  1123
Query_NITEN  1121 NVTSM---LRRLYNRLDHPNPYQRLGATMALYR-LYPLLVRQ-RGIVDRHIFELFFYCAKSLRLAEA   DSETIG   T  1189
Query_CLUBM  1007 NVTSM---LRRLYNRMDHPSPYHRLGASMAAYQ-IYPILRQN-SSIVDRYIFELAFFLLKSLRVGDT   DAEQIG   T  1075
Query_MOSS_  1071 NVRSM---LRRVYNRMAHPNPYHRLGAAMAAHR-MYPVLREN-GAIVDRYIFEILFFCIKSLKLGEL   DDEQIG   T  1139
Query_LIVER  1086 NIASM---LRRLYNRMDHPNPYKRLGAAMAAHR-IYPVLRGN-PNIVDRYIFELIFFCVKSLRLAES   DDVQTG   T  1154
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1120  QEQCKSALLHV    ERILI--KKQVTFN--KDS--KQ-RIKPRGWR-----D-aNMTTAVRWLLRQ--CGRPQTECRHV  1180
Query_LINGU  1129  QEQGVQVLNHL    ERIIV--NKAALLC--KDS--KI-RRKPKTWN-----K-aTLDICVRWLLRQ--CGRPQTECRHS  1189
Query_LIMUL  1126  QEQCKQALDHV    ERILR--VRAELLN--KPS--KQ-RRQPQCLD--G--T--TLSAAVLWLVQQ--CGKPQTECRHK  1186
Query_SPIDE  1125  QEQCVNALNHL    EKIIV--VKHSLFI--KSE--KS-RKKPECLE--E--A--TLQCIVLWLLKQ--CGCVQTECRHK  1185
Query_SCORP  1124  QDKCKQALNHL    ERIME--VKAKLLS--KTS--KK-RRPLRNIN--D--S--ILPEVVEWILKN--CGCLQTECRHK  1184
Query_TICK_  1081  RELSGCALDSV    LRIVS--TKSELLN--RPNv--K-RKWPSDTGGEP--V--VLSTAVRWLLLQ--TSRPQADCRHK  1143
Query_MONOS  1052  SERLTEVVMDC    KRMLI--HHAGILN--KPE---RhRRGPASWPQAV------LSNLVAHCFGGf--GWTSAPARHA  1112
Query_ROSET  1033  ATHTERVLTNL    RRIIL--KKADKFR--RTS---RvRRIPPGWETCT------LAEMTRWSLTH--ACDLHAAVRHQ  1093
Query_BELCH  1146  QAQCVAVLTHL    DRIIT--TKADLLN--KHS--KV-RRVPRGFDRES--P-pTLPDLVTWLLQQ--CGRPQTECRHR  1209
Query_CIONA  1117  QKVCKQALEHI    ERIIS--VKSSMLC--KPS--DI-RRTPRCFTNTT--S-iTLSWIVEWLFDQ--CGRPQTECRHM  1180
Query_CAECIL 1133  VQQCCDAIDHL    KRIII--HKAPSLN--KTL---K-RRLPRGFPSTK--SsfCLSDLVPWLLGQ--CGRPQTECRHK  1196
Query_HUMAN  1124  IQQCCDAIDHL    CRIIE--KKHVSLN--KAK---K-RRLPRGFPPSA--S-lCLLDLVKWLLAH--CGRPQTECRHK  1186
Query_MOUSE  1121  VQQCCDAIDHL    RRIIE--KKHVSLN--KAK---K-RRLPQGFPPLT--S-lCLLDLVEWLLAH--CGRPQTECRHK  1183
Query_XENOPU 1146  TQQCSDAVDHL    KRIII--RKAASLN--KAT---K-RRIPRGFPQ-G--NtvCLFDIVLWLLEQ--CGRPQTECRHK  1208
Query_PLATY  1126  IQQCCDAIDHL    KRIIE--KKHISFN--EEK---N-RRLPRGFPNAR--S-vRLLDLVYWLLGQ--CGRPQTECRHK  1188
Query_TASDE  1129  IQQCCDAIDHL    TRIIE--KKHITLN--KPK---K-RRLPRGFPPVQ--T-lCLLDLVKWLLAQ--CGRPQTECRHK  1191
Query_CANAR  1123  AQQCCDAINHL    KRIIK--HKAPSLN--KEG---K-RRVPRGFPPTK--S-vCLEDIVMWLLVQ--CGRPQTECRHK  1185
Query_OSTRIC 1093  TQQCCDTINHL    KRIIK--HKAPSLN--RAG---E-RRVPRGFPATK--T-vCLLDIVMWLLVQ--CGRPETECRHK  1155
Query_TURTLE 1132  VQQCCDAIDHL    KRIIK--HKASSLN--KAE---K-RRIPRGFPNAK--S-vCLLDLVMWLLVQ--CGRPQTECRHK  1194
Query_GECKO  1132  TQQCCDAVDHL    KRIIK--HKASTLN--KES---Q-RRLPRGFPPAK--S-iCLEDIVMWLFLQ--CGRPQTECRHK  1194
Query_DANIO  1138  VQQCCSSLDHL    KRIIK--HKADSLN--INS---K-RRIPRGFPADQ--S-vCLSNVVLWLLTQ--CGRPQTECRHK  1200
Query_SNAKE  1132  IQQCCDAIDHL    KRIIK--HKASTLN--QKS---T-RRLPRGF-PSR--S-iCLEDVVMWLLRQ--CGQPQTECRHK  1193
Query_STERL  1142  LQQCSDAIDHL    KRIIK--QKAASLNaeKTS---R-RRVPRGFPPDK--P-vCLLDVIMWLLVQ--CGRPQTECRHK  1206
Query_MILII  1128  VQQCCDAIDHL    NRIIK--QKASALN--KPI---K-GRVPRGFPPDQ--P-vSLKEVVMWLLTQ--CGRPQTECRHK  1190
Query_BAMBO  1130  IQQCCDCINHL    KRIIK--QKAHELN--KPV---K-RRVPRGFPADQ--M-vCLTDVVMWLLTQ--CGRPQTECRHK  1192
Query_ALLIG  1092  TQQCCDAIDHL    KRIIK--HKAPSLN--KAE---K-RRLPRGFQPAE--S-lCLLDIVMWLMMQ--CGRPQTECRHK  1154
Query_GAR__  1131  LQQCSGAIDHL    KRIIK--HKAATLN--KSS---R-RRVPRGFPPDK--P-vSLTDVVMWLLVQ--CGRPQTECRHK  1193
Query_TETRA  1128  SNEVEDNFNIL[ 4]TNIIQ--KNYKTLK-aKNEn-----------RSGQf---kLLDSLLDWLYSY-gFFNSSYLVRKS  1189
Query_PARAM  1076  VSNTNNIIDII[ 4]ERFTQ-----DLLQ----Em--------KGRQGKNk----TLYSLHEFLLNE-aMIGKDGGWQQA  1133
Query_SLEMN  1107  IDNCIMMLEKV    SKVVQ--RKISILM--NVN------------PKRS--IhqSVFTFILYLFEK--FVAIETVCRRE  1162
Query_STENT  1004  LDYCKTTLNDL    KALLE--NKFSVFC--VSN------------DKRT--YhkDLFSLLEWLWEQ--CKRPERVARKF  1059
Query_NEMAT  1127  QDQCSQVISHL    ERIIK--TKATLLD-kENK---Q-RRIPRGFQMLD--S-iTLSHLIPLLIKL--CGKPQTECRHQ  1190
Query_PISTI  1140  QEQCVQVIIHL    ERIIH--SRAQLLV-kENH---R-RRIPRGFASQS--S-iTLIHMIPWLVKQ--CGRPSSECRHQ  1203
Query_CORAL  1122  QSQCIEVISHL    ERIIV--SKSSVLL-nVNK---K-RRVPRGFSA-D--T-tTLRHLIPWLVRK--CGCSQTECRHQ  1184
Query_HYDRA  1074  IDQYELSLQHM    EKIIK--KKCDILL--KPF--KN-RRHPIQFKDTT--D-fNLRHVSMWLLSE--CGSPQTACRHQ  1137
Query_LATIM  1132  VQQCCDAIDHL    KRIIE--QKAPSLN--KAA---T-RRVPRGFSPDK--V-fCLSDVVMWLLVQ--CGRPQTACGHK  1194
Query_SHRIM  1108  HEQTKSAINHI    ERIIR--VKSDLFQ-kKSG---Q-RRLPPGF--GDg----TLVDVLKWLFEQ--CGRPKALVRHT  1168
Query_MCRAB  1105  HEQTRAAILHV    RRILI--VRKELFW-kMSS---N-RRIPIDF--EGg----TFEDMLLWLLEQ--CGSPKTLVRHM  1165
Query_AMPHI  1148  AALASTSLRHL    CRIAT--HYKNLFN-aRSE---M-RRIPDCLTGSDg----TLAALTEWMLRL--CGHRASALRHA  1210
Query_BARNA  1066  DTAAADALGHC    QRILAa-RASHLLE--KSEq----RRVPAGWQGAT------LPDLLSWLVGQ--CGAPQTECRHA  1126
Query_COPEP  1075  SSASLQNLEHL    KRILF--HYLPLFQ---TNp-GE-RRIPVGFSGGT------LAHILDWLIDL--TFLQLTEARHA  1135
Query_DAPHN  1045  LHQSKQALYHL    TQTVNl--KRNIFE-hPDSr----RRVAPFLTQGT------LPELLDFLLNH--ALGSQRHCRQI  1105
Query_STARF  1146  QLHGVEVLNHL    ERIIR--VKADMLN--KES--KI-RRIPREFPAGV--C-pTLSMLVIWLLRQ--CGRAQTECRHQ  1209
Query_STRON  1150  QEQGSEVLKHL    ERIMR--VKAALLN--KES--KN-RRVPRGLPEDC--S-pTLSHMVMWLMTQ--CGRPQTECRHQ  1213
Query_SEACU  1139  QEQASEVLKHL    ERIMK--SKANILN--KEN--KH-RRIPSNLP-GN--C-kTLSHLLVWLINQ--CGSPATECRHQ  1201
Query_FUNGI  1094  AEKAERVIDQL    GEVIA--SSAKVLV--KVN--RK-RREH-----ED------LRGLVEWLLMN--IGRQEMESRRV  1149
Query_DPURP  1207  SSKYSTVLVAL    TKVIV--RKATMLN--KTQ--QS-RREH-----KD------LSHFVQWLFQV-gCSRAESVVRYE  1263
Query_HALBU  1223  GEKLQAVIDAV    VRVIE--RRAELLN--RTN--NN-RREH-----KD------LADFVGWVFGE--VGRMERRVRFT  1278
Query_GLOBO  1053  HALVKIAIDNM    AKIVV--VKHELFL--ATT--SS-RRSFPDLPNCN------LGTVTRWLFEQ--SKHKELEYSYK  1113
Query_PALUS  1101  QQMTIKGVQHL    LKIIK--SKSDLFR--SSK--ST-RRSFPGLKEAD------LGGLVDWIFAE--IGRLELTYAQQ  1161
Query_CALIF  1158  TEQCKVAINHI    IRILK--KKYSIFL--MSK--ES-RNHFPGLENTD------LKSLVEWLLHE--TGALEEIYSKT  1218
Query_GLOMU  1180  RQQAISAIRHL    KKIII--VKSNLFL--NDT--PR-RRKFPNLNRSD------LSSLTEWLFIE--TSRKEQDYSSM  1240
Query_LOBOS  1132  RHQACLAIARL    QRIVQ--VKAALFL--KDS--KD-RRRIPGLEEAT------IDGLVHWLLRE--TSRPEMEYTKI  1192
Query_RACEM  1113  REQSRKAISHI    KRIIR--EKAPAFL--KPA--SG-RRTFPGIEEPT------LQETVAWSFQE--VGQLERDYARS  1173
Query_PVIOL  1182  ASKFSNVLVAL    TKVIQ--RKADLLN--NPD--TS-RREH-----KD------LSDFVSWVMKD-cCSRPENIARYE  1238
Query_BOMBU  1038  --SIINALNHV    ERVLI--AKKNLLN--AEH--HN-RRKPYEFEGST------LVDAVKWLLTQ--CGCLDQHCRIK  1096
Query_AEDES  1109  LEQVSATIDHL    VRVLV--VRKVVFN--RES--SV-RIVPRVFGGSL------LKDLVLWLFGQ--CSSRESSYRHK  1169
Query_ANT__  1057  --SITNALSHI    KRVMK--VKANLLN--TLN--SSkRRKPHEFDNAT------LTYALYWLLSQ--CGTYDEHCRAK  1116
Query_WHITE  1072  -DQLRSALRHI    LRVFEkeEFRKMLN--SRN--PN-RKKPIEFEDET------LQGATLWLLKQ--TSSPSIVCRTE  1133
Query_BEETL  1060  VTQINNALSRL    ERGFV--EKSDLFN--NLD--TQ-RRIPNEFNGGK------LADVAVWLLKQ--TGVKNDHCRTK  1120
Query_THRIP  1051  EAHIDKSLQHV    LRVLM--ECSNKFN--LDS--PS-RRIPRELTGRN------LWHVVRWLLKY--CTALNSQCRHS  1111
Query_TERMI  1085  QEQVVTAVNHI    ERVLK--EKASLFN--KED--NI-RRKPPEFEGIL------LKDAVSWLLLK--CGTVQVECRHK  1145
Query_FLEA_  1028  IQQLEETLNHL    MRVLV--EKSDLFN--RDS--GH-RNIPQVLNGCL------LKHYVTWLFKH--CGFVENHCRRK  1088
Query_MIDGE  1074  FEQITATMSHL    TRVFV--ERSDIFH--RAS--TE-RIRPNGFEGKT------LIDVLKWLFKEsgCHQHHSRYRHN  1136
Query_LAODE  1085  -REVFAALDHI[ 1]-RVLV--ERPEGFNvvSEA-----RRRPPEFEGVL------LKHAVKWLLAK--CNSSNERCREK  1144
Query_LOTTI  1112  QEQCAAVLKHL    QRIII--FHEDILK--KPS--KI-RQEPREWS-----N-kTLDIAVRWLSRQ--CGRPQTECRHT  1172
Query_APLYS  1129  QEQCSKALDHL    ERIIR--VKVDMLK--EDPglKV-RPEPSAWS-----S-rKLEFAVRWLTRQ--CGRPQTECRHV  1191
Query_SCALL  1142  QEQACKALEHL    EKIIC--KKADILR--KES--KL-RTEPSQWS-----K-iTIDICVRWFMRQ--CGRPQTECRHM  1202
Query_CRASS  1120  QTQCKMVLDHM    ERIIK--VKANVLS--KED--KN-RTEPKQWS-----V-vKLDIAVRWLMRQ--CGRPQTDCRHA  1180
Query_OCTOP  1095  NSQCCKALDHL    ERIIL--KKADMLK--EVS--SA-RKQPSQWN--Q--I--TLLYAIRWLVRQ--CGRPQTSCRHA  1155
Query_POMAC  1124  QEQCKKALEHV    EKIVC--KRVDILR--KPS--PH-RKEPSEWS-----K-qTLDIAVRWLVRQ--CGRPQTECRHA  1184
Query_WHIPW  1109  EEECVKALGRL    ERMLV--RYVALFQ--APSs--V-RRKPRDWKAAT------IYCCLDWLLEE--MGRPVSSFRTV  1169
Query_TRICH  1108  VDECKQAIKHI    RRIIL--HNLDMLQ--IHNv--Q-RRTPNEWSQST------VHCCCDWLLNT--MTWSNYQFRSE  1168
Query_AASTA  1030  ----PNAVLPL    VHALA--HMEKILV--RRA--ADlQHDDDTRVVLGgtDclNLNAFTTWLFSN--VASSAPTFRSN  1093
Query_AEUTE  1007  ----SDSALSL    HHAVD--HMEKIIL--KSA--PTlQHENPDRMILGgsGclNLESFTSWLFSN--IANASPDFRNR  1070
Query_PLASM  1092  REIMLNAVMKI    ERGIC--RFPHLFL--SDS--STtYSIDND---------yTLHQMTQWLFQQ--TGAREQSFRRL  1150
Query_BREMI  1083  HDVITKAVLKI    ERGIC--RFSHLFL--NEN--ISsKGT-------------SLQQTTEWLFRH--STAPEVLFRQL  1137
Query_OLIGA  1089  QEIIAKAILKL    QRAIS--RFSSLFL--SSQ--NEeENL---------------KSLTTWLFHQ--CSRPESSYRRM  1141
Query_LIGNO  1173  VEQCNLALDHL    ERIMV--ERAATLK-vESQg----RRTPLGSKWTSa----TLEVAVRWLVGQ--CGQPESACRAA  1235
Query_CESTO  1058  -DLLIALIKHS[ 3]---LTfeDYSPALK--RPRl-TRtRQIKSNIDSLKp---------VTWESASi----------SS  1110
Query_OVI2_  1232  LLSCFKVVVRT[32]KRMKLdeSQSSDLK---ASelNK-RIRPKRWKDANi-------EGCVWALLS-cIGQLQNPQSSE  1327
Query_PRIAP  1140  QEQCAQALDHI    ERILR--GKAHLLN--KDK--KS-RRRPKEFT-----D-aTLFTTVGWLLQQ--CGRPQTECRHK  1200
Query_ROTIF  1138  QEQCKLGLDHV    ERIIK--FKSDILN--QID--PK-RVKPPGWSEPV--L----EVAVRWLMRQ--CGRIETECRHK  1198
Query_SPONG  1130  VDQCSLSLDHM    SRIIK--AKSSTCL-sKAN--KN-RRVPRGLKSAD--L----SILTAWLLAQ--CGSIETEARKM  1191
Query_TARDI  1140  TKAVELCLLHL    ARIVE--KYAKTLI-sTAS---R-RRVPVELDEGKleKeaTLEKLTTWMLMQ--ACRPELACRMS  1206
Query_TRICH  1109  QEQLSKALAHI    ERIIK--AKASTLN-kESK---N-RRTPTGFMT-K--T-vTLHHVVNWLVRQ--CGRPQTACRHQ  1171
Query_OSTRE  1124  DVAMAVLIQEI    LHIIQ--ANANRLL---SS--------PGSKATESv--elFLQRTLELTMSS------DEFTRRE  1178
Query_NITEN  1190  VRQMSRVVHCL    QKLML--KFVDPLK--KPR--KDgRPMF-----AT------LDQFLGWLFTQ--TGRPQQRCSQQ  1246
Query_CLUBM  1076  IKQAYKALDGI    EKLIV--RNISLLQ--KNK--KD-RVLF-----CT------LEDFMTWLFEQ--TKLSHERCRTK  1131
Query_MOSS_  1140  MKQMNKTVELL    QRLVT--RNITLLE--KKR--KD-RPLF-----LA------LDDFLQWVFEQ--TRLLHGRSRTK  1195
Query_LIVER  1155  AKQVNKTLDEY    MKLVE--RNISALQ--RSR--KD-RPMF-----LT------LDDFLEWLLHQ--SGQLQEACRSR  1210
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1181  CMRLVCQ    LA---PCVH-EINGARELF   SVLLSQG---NSEYF---IQRFEFG-MYS[2]RMGLNQKPTMS-RVQ   1240
Query_LINGU  1190  SMMLVYK    LA---PCVP-EFQSVQGFF   QKFLKMEs-gGPQYF---MERFEAY-RLG[2]RFTLKMCHTLG-DIG   1251
Query_LIMUL  1187  CMELVFK    LA---PCLP-GTHNTQEYF   QQLLKMK---GPQFF---INSFENGLKTG[2]L---------E--KT   1237
Query_SPIDE  1186  CMDLVIK    FA---PFSP-GCSQNLSSF   IITYENC---GKNII---EDVFEAILQDS[2]S---------A--LK   1236
Query_SCORP  1185  CMELVFQ    LA---PCLS-GFNGCKDYF   SHLLETK---GKKFF---LNCFESEM-SN[2]P---------S--FN   1234
Query_TICK_  1144  CMELVHA    LI---PKLP-GAPSHQSYM   SSLLQGN---SPGFF---VESFEGNLAAS   --------------E   1189
Query_MONOS  1113  AIDLFMH    LVtk---IP-GPDGHGAMR[1]QQYLTRPvg---GFTpglIENLKRGQVGD[1]--TIATSHLGSsAVL   1175
Query_ROSET  1094  AMAFFLH[ 9]-Tgk----P--YATAHAWV[3]----------GAPAL---VALVEQGSAPQ   -------PLLGvSVK   1148
Query_BELCH  1210  CMLLVCS    FL---PLLP-RKPSPQAWM   KERLKED---GPDFF---ITRFEGGGGKT[2]KSGIKKHPKIS-TMD   1270
Query_CIONA  1181  CMHLFHE    FIk---LLPtEERSATSWV   KRKLTKY---GAGNL---VKRFEWGLVNM   -Q--------PvLNK   1232
Query_CAECIL 1197  AIDLFCK    FI---PLLP-GNKSPSSWL   NDHLKEE---GISFI---ISRFEGGG-SD[3]QSGLIAQPTLC-DLK   1257
Query_HUMAN  1187  SIELFYK    FV---PLLP-GNRSPNLWL   KDVLKEE---GVSFL---INTFEGGG-CG[2]-SGILAQPTLL-YLR   1245
Query_MOUSE  1184  SMELFYK    FV---PLLP-GNKSPSLWL   KDLIKKK---GISFL---INTFEGGA-SS[4]-AGILAQPTLV-YLQ   1244
Query_XENOPU 1209  AMQLFFE    FV---PLLP-GNKPLTAWL   DDQVEKE---GIIFL---INRFEGAGHSD[3]-TGIFNIPALH-DLH   1269
Query_PLATY  1189  SIELFYK    FV---PLLP-GNKSPASWL   NDILKEE---GVSFL---ISRFEGAG-TG[3]LSGILVWPTLL-DLE   1249
Query_TASDE  1192  SIELFYK    FV---PLLP-GNRSPSCWL   TDILKKE---DISFL---INRFEGAG-NA[3]LSGILAQPTLF-DLQ   1252
Query_CANAR  1186  AMELFYE    FV---PLLP-GNKSPSSWL   ADTLKKR---TVTFL---INKFEGGG-TD[3]PSGILSQPTLH-GMQ   1246
Query_OSTRIC 1156  SMELFYE    FV---PLLP-GNKSPSSWL   ADVLKEK---GCSFL---INKFEGGG-RD[3]PSGIFSRPTLH-GIQ   1216
Query_TURTLE 1195  SMELFYE    FV---PLLP-GNQSPSLWL   GDLLKKQ---GISFL---INKFEGGG-ND[3]LSGILSQPTLC-DMQ   1255
Query_GECKO  1195  AMELFFE    FI---PLLP-GSKSPSLWL   DDILRKQ---EISFL---IKKFEGGG-DE[3]VSGILSHPTLS-GLQ   1255
Query_DANIO  1201  SMELFFE    FV---PLLP-GNSSPAMWL   DEQLKQR---GPGFL---ISCLEG-----   -GGLLSQPTLR-EIE   1253
Query_SNAKE  1194  VMELLFE    FV---PLLP-GNPSPSSWL   GDVLQKE---GIYFL---IKKFEGGG-CE[3]VSGILSHPTLY-ELQ   1254
Query_STERL  1207  CMELFYE    FV---PLLP-DNKSPASWL   DDILKKQ---GVAFL---ISRFESAG-RE[3]CSGILSQTTLH-DLQ   1267
Query_MILII  1191  CMVLFYE    LV---PLLP-GNKTPSCWL   ADVLKER---GVSFL---IGKFESAG-SD[3]CSGIIYQPTLK-DLQ   1251
Query_BAMBO  1193  CMELFYE    FV---PILP-GARTASLWL   EEILKRN---GVSFL---VNKFEGAG-SD[3]CSGIICHPTLQ-ALK   1253
Query_ALLIG  1155  SMELFYE    FV---PLLP-GNKSPSSWL   GDVLKKQ---DSSFL---INKFEGGG-NQ[3]LSGILSYPTLH-HLQ   1215
Query_GAR__  1194  CMELFYE    FV---PLLP-GGKNPAVWL   EDILKDE---GEGFL---IRRFEGAG-RC[3]SSGLLSQATLR-DLQ   1254
Query_TETRA  1190  SFELWQS    ILnr----EdSDISCKKFL   ---AKRKq--GDNLFg-----------DQ[1]-----------kQLK   1229
Query_PARAM  1134  TLSLWYL[ 1]VSktqKVIK---KNARELL   -----------VDQIa------EIGQKMK[1]---FPKYPINRlQGE   1182
Query_SLEMN  1163  SMKLWEA[13]--------P---DNPVKWI   FSYYLQRr--GEKAIf---KRFAKISFIE[5]QNDQLNHRLRSeAVR   1234
Query_STENT  1060  AQDYWKT[ 2]------------------I   ISKY-------------------------[4]WKSPENCQKDLeDVK   1092
Query_NEMAT  1191  CMKLVCQ    VT---PLLH-GIGSAAIWM   KNTHK-E---HDSFF---VRRFEGGG---[3]N-GILTYPTLA-DMG   1247
Query_PISTI  1204  CMRLISK    LA---PLLP-GVNSAADWM   KKTHE-D---HPSSF---VHRFEGGG---[3]QKGIEQYPTLA-HIG   1261
Query_CORAL  1185  CMRLVSV    LA---PLLP-ECSSVQTWM   KKTVESD---GEDYF---ILRFECGF---[3]S-GIQKHPTLS-SIN   1242
Query_HYDRA  1138  CMKLFIS    FMl---L----QEGVEEYM   KNAIADR---DGDDL---FRYFEGGGISR   -KGIHHYPTLPeKLK   1193
Query_LATIM  1195  AMELFYE    FV---PLLP-GKKSPAAWL   DEVLKKQ---GVSFL---IQKFEGGG-NT[3]LSGILSQPTLC-DLQ   1255
Query_SHRIM  1169  CMELLAS    LA---PQVK-GCASVRDFI   ----EKYv--GISKLseiVDILEGGGEAK[1]GISYVKNPTASsKHI   1229
Query_MCRAB  1166  SMELLES    LA---PHSK-DYTSVHKFI   GQQLEKPe--SIKKL---VDIIEGGGQAG[1]GIKHHTEPTSLaT-A   1226
Query_AMPHI  1211  VTASLPG    LA---SSVL-GCSNLQSLV   DQRFIKAn--RLDVL--vGEVLERGGCKA[1]APVIGSGDISPsP--   1271
Query_BARNA  1127  CYRLLEA    LT---PLAP-----AAGGL[1]------------GLVa------AGGGAVR[1]---------LEhALT   1163
Query_COPEP  1136  AMTLYHE[ 8]-Aaq-PILE-KFGTINKLI   ---LRVEha-----Lt----RTEAAG---   --SIVEVPGGSsGTS   1193
Query_DAPHN  1106  CLELCGI    LAr-----P---QSLAEVM   -----------IQFFs-------------   ------RQPLEsHLP   1137
Query_STARF  1210  CMILIYK    IT---PLLP-GHNTPQTWM   QNTLKEK---GPDYF---VSRFEGGGVKA[2]RSGISKYPTLS-TMG   1270
Query_STRON  1214  CMVLVYG    LT---SLLP-GRNTPATWM   QNTARER---GAEYF---VGRFEGGG---[1]KSGIQKQPTLA-SVE   1270
Query_SEACU  1202  CMALVFN    LS---SVLP-GSNSANVWM   KNTLKEK---GASYF---INRFEGGGGRD[2]KNGILVSPRID-LKD   1262
Query_FUNGI  1150  CMSTFTR    LS---VCLP-GIHTTEQCV   RSFSTDQegt--TMV---LRSIQPYDF--[2]L--RYQLENPQsS--   1206
Query_DPURP  1264  SMLMFSQ    LV---VLLP-SKKTPLQWV   KD-H--VntnGLSFI---ISVAE--PRTS[3]VAPKIQVGNH-----   1319
Query_HALBU  1279  SIALFGR    VV---QLLP-GQKRANIWI   QH-QVKQhgl--AAL---IKMAEPPQT-A[3]P-PKLDGNNH-----   1334
Query_GLOBO  1114  CIELFDK    FA---SRMT-GVVS---WM   KN--IQTnneR--YV---FDLFK--LP--   ---VNTAVLDDcQ--   1160
Query_PALUS  1162  CMQVFME    LA---PLIS-TPKE---FV   RS--RSSskpT--FI---PDILE--RT--   -NLVFPSETDLqS--   1210
Query_CALIF  1219  CMDLFLI    LN---PLLP-ENSKIG--L   NE--LFLnnnN--LF---NDIYKVDEL--   --VIIESNRNLlL--   1269
Query_GLOMU  1241  CRELFGE    FV---KLLP-DFKTGFRWV   KA--RIDrdpK--FL---ESLYKNSEL--   -SPSNVLHKGLaA--   1294
Query_LOBOS  1193  CRMLFDS    FV---KLLP-GSPSSGTWI   GA--KIThnaG--FV---ASLYQ------   -CPQYTSETFTqP--   1242
Query_RACEM  1174  CMDFFNE    FV---VTLP-GIHSAKDWL   QE--RYArddD--YI---TNILETQRL--   -RPTQVKKE--nA--   1225
Query_PVIOL  1239  SMQLFIQ    LV---VLLP-GIKTPIQWI   KK-Q--Qpld--QYL---ISIAE--SQ-A[3]I-PRVDTANSTsSSA   1295
Query_BOMBU  1097  CMELVVK    LS---EHVA-DCDSAKTII   NNYIDTY---GIETF---NSI--------   --------ILS-QLP   1139
Query_AEDES  1170  CMEVFPK    LT---SSVD-GCRSAAEFV   GEYLSDQ------EI---LDICDHVDTVH   --GIRSAPNLQ-FIA   1223
Query_ANT__  1117  CMELYVN    IS---QDIG-SY--AQETT   QNFVQIY---GINRL---NDI--------   --------ILK-GLE   1157
Query_WHITE  1134  TMVLFQR    FA---SFLP-GFISASNYL   NSLISTE---GEDFV---SNLVEEN----   ---LTVEPLRD-RIS   1185
Query_BEETL  1121  CMEIFCS    IS---SKCK-GYTNMQNFI   NSHFD------GSWM---RDIYETYLIQH   --PTLQCPMKN-EDF   1174
Query_THRIP  1112  SMKLVDC    LA---RYVS-DCRSTKDFI   ESVISIE---GNCSL---LLVAEG--ETS   --LRTTLPSSA-PWD   1166
Query_TERMI  1146  CMELVYV    LA---PLLP-VCKSGSEFV   SWYAATH---SWTSI---IRICEGSKESG   --GITVTPTLK-HLR   1202
Query_FLEA_  1089  CMEMFLK    LA---VLVD-EINTAHMFI[4]GNFV-IK------SV---IEIGECSTNTN   --GLNSKPTLE-HLK   1145
Query_MIDGE  1137  CMEMFLK    I------LP-VCGVKEEIC   GKILNVE------KI---LVVGER----N   --GIATTPDLK-HLS   1183
Query_LAODE  1145  CMQMVDK    LVvq---VQ-GISSLNDFI   ----ETE---GIDMLr--ICKMKTTSNIT[1]-------------MS   1189
Query_LOTTI  1173  CMKLVYN    LC---THVQ-G-----NYF   VSFDLIDivdKSTII---KKVFEGGGTAT[2]KSGISACSTLL-KIS   1231
Query_APLYS  1192  CMRLVYN    LC---PLLQ-GVKSQKTFF   SIIHKSE---GASYF---LKRFEGGGQAA[2]THGILGQPTLPgKDK   1253
Query_SCALL  1203  CMKLLHK    MT---PLIQ-GIKSTQMYF   QAFLKSE---KQKYF---ITRFEGGGHTS[2]KVGLVACPTLM-SVS   1263
Query_CRASS  1181  CMKLVHQ    LA---PCIT-GIKSTKDYF   QTFYKSE---KASYF---INRFEG----K[2]KYGLLSCPEMT-AVD   1237
Query_OCTOP  1156  CMKLVYN    LA---PQIQ-GIKQSEEFF   QIFYEKE---GHNYF---IERFEGGGEGS[2]SNSIMRCPQLE--IR   1215
Query_POMAC  1185  CMQLMYK    LA---PLIP-GISSPFDYF   KAALDIQ---NASYF---VVRFEGGGQ-V[2]RGGLCGHPTPS-KVA   1244
Query_WHIPW  1170  CLTAYSN    IT-a--SLK----------   MDHFQDYf--SSDMKr-iIEHFESRHLEK[1]-NN----PIEEq---   1216
Query_TRICH  1169  CFISYQE    LTda--EIE----------   KLHFLRHi--GTSIFg-fIDHFERSHRSV[1]-GSFSICNLQDrVIK   1223
Query_AASTA  1094  CLRLFEQ    LA---RMVN-MTGSCKAWL   TQFQKNH---SAAQLq------------Y[1]---LAPPALTDiSVR   1142
Query_AEUTE  1071  CLKLFVT    LS---RMVG-MTGSCKEWL   KQYQKSH---TVGQLi------------D[1]---LAPQELISlYF-   1118
Query_PLASM  1151  CRRIFLS    FS---ALIS--T-STTEWM   HNYETVQ---ERKSIl------------A[1]---MVPLS----ALA   1193
Query_BREMI  1138  CRRIFMS    FS---VLVD--K-SIPEWM   QSYELKH---SGESIi------------G[1]---LVPIS----ALA   1180
Query_OLIGA  1142  CRQLFMA    FA---AIVN-GS-SCKKWI   QAHASTE---GCDRIv------------E[1]---LAPMDNLVlGLR   1189
Query_LIGNO  1236  AMRLLVN[14]LArc--LLT---SSPEKRQ   QTAASPRq-----FFqalLKQFKANYFYD[1]--WERFQPHQN-QQQ   1307
Query_CESTO  1111  CVKSFLR[ 2]--------K---NSPSRKL   --------------Fq------------E[1]-------------IE   1133
Query_OVI2_  1328  LATRWSD[ 8]LIcc-----lGELKRYQWS[4]ESSISTShl----FFq---TALDTGNEEM[5]EAGLSVSVDEGqHNH   1402
Query_PRIAP  1201  CMELVCK    LSmhqSKLP-VKSAAE-FF   RLLLDLH---RAEYF---IERFEAGGIPG[2]TCLSNQQ---E-GFE   1260
Query_ROTIF  1199  AMELSHK    LA---PNIS-GVKDIKDYF   NIRLVNE---SEMYF---LARFEGSTLGK[3]KTSLSTYQSFL-DLT   1260
Query_SPONG  1192  CRKLLVQ    LA---PPSS-NVVNIRMWI   RKTIESE---GGNYF---VTRFESHG---[1]-SGILKAPTLA----   1244
Query_TARDI  1207  SWALVGR    FLl---ALK-NATSGGYFF   ENLARSK---GPGWP---SKAFFCGGFDP[3]AARALNAKEIS-SVF   1268
Query_TRICH  1172  CMQLVYG    LA---PQLS-GVSTIQDWV   KTIIGSS---SVTFF---ITRFEGGG---[1]KYGIGRYQTAL-DIS   1228
Query_OSTRE  1179  AQKVFDV    IA------PiSVSDTSEWF[2]RSHLVSSs---IDALv-------------   -------PEMS-AVE   1220
Query_NITEN  1247  CMILFSA    LA---PESA-EGSAAV-WF   RS--KLSpstPLEAL---LEPFDV-QH--   -----PLPPGAsS--   1296
Query_CLUBM  1132  CMALLSA    VC---SALP-GGSCMH-WF   RE--RFTgtqRLHML---AGLF----E--   -----KLPSNI-F--   1177
Query_MOSS_  1196  CMMLFCS    IC---LSIP-GCSSAH-WF   RA--NICsgdQVSST---VKQFDVGM---   -----DYPHHP-S--   1244
Query_LIVER  1211  CMRILSY    IC---RLLP-AGSLAS-WY   RS--KLDhgaDMESI---IGVFNYEEE--   -----DRPFDVwS--   1261
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1241   -EN-FSLN   AACTWFDLVLATLDCYTWVFGERLMSPNQIL-iASN   NK---GT-IIFQVLHHFMEELAQEDIS     1304
Query_LINGU  1252   -EQ-FTVK   NAVLWYDMVLAALECYKWTFDTNLLRPEDVF-t--G   GH---SS-CIYLVLSHFLSSLALKDIE     1313
Query_LIMUL  1238   -FC-ASLK   VAVDWMEEVLGILDCYTWIFGECLLEPSSVF-kASI   QS-kgT--KIFASLEFFLSH-------     1295
Query_SPIDE  1237   -NQ-VSYK   SLSRWLEALQTSMECYQWAFSKQLMLQNMLC-eEKN   AS-liSCiKVFNAIEHFLENIV---KE     1300
Query_SCORP  1235   -QN-DGLK   PIINWLEGLHTMLDCYYWVFGQSVLKLSDIF-nDA-   SN-reS--KFFPSIEFFLKTLCIDFDQ     1298
Query_TICK_  1190   -GPwATLS   AVCKFCDQLLGLLECYTWIFEQRLLEPKNLF--EVP   KK---PS-GLFGMVDHFLEHLAV--CE     1251
Query_MONOS  1176   ----RSLR[5]EAAHWLRALAACLEGYHLIIKHGLVSPETLAtmMSS   EQ-----cTIVPALHEFQEALLGSGTW     1242
Query_ROSET  1149   -----PLE[4]SIHEWLERLQATLEAYQLVLSHRLLTPRALLvdCEM[1]---------LWGALAHFLSTFPRRSEL[4]  1214
Query_BELCH  1271   -ST-FTVK   TAVSWFDLLLAALDSYCWVFGQRFLSPSDVL-sGSG   KK---PS-QLFTALEFFLTKLALSDID     1334
Query_CIONA  1233   -EK-LTIK   ILCGWFDHVIAAIEVYSWLFQQKILSPSIVW-sGEA   NT----S-CLFKYLTYFMVNISAKSAE     1295
Query_CAECIL 1258   -EP-FSLR   AVVQWMDMFLAALDCYNTFIWMRAIKPNEIL--STT   TQ----S-SFLRTCEFFLQKLCLHDIT     1319
Query_HUMAN  1246   -GP-FSLQ   ATLCWLDLLLAALECYNTFIGERTVGALQVL--GTE   AQ----S-SLLKAVAFFLESIAMHDII     1307
Query_MOUSE  1245   -GP-ISLR   GVLQWLDLLLAALECYNTFIEKETVQGQEVL--GAE   VQ----S-SLLKSVAFFLESIATHSAR     1306
Query_XENOPU 1270   -EP-FSMH   AVLQWLDMLLAALDCYNTFIGMRFLKANTVL--GKN   AE---KS-SFLKAAFFFITSLSMENIK     1332
Query_PLATY  1250   -GP-FSLR   AVLQWMDMLLASLECYNTFIEEKTVKAHEIL--GTQ   AK----S-SFLRAISFFLERIAFHDIT     1311
Query_TASDE  1253   -GP-FTLR   AVVQWMDMLLAALECYNTFIEERAIKAQEIL--GTK   AK----S-SFLKAVGFFLENIALHDIR     1314
Query_CANAR  1247   -EP-FSLQ   AVMQWMDMFLAALDCYNTFFELQMIKPDEVL--GVN   ES----S-LFLKAVQFFLDTIALHDIH     1308
Query_OSTRIC 1217   -EP-FSLR   TVMQWMDMFLAALDCYNTFFELRMIKPHEIL--GVN   ER----S-SFLKAVRFFLETIALHDIQ     1278
Query_TURTLE 1256   -EP-FSLR   AVVQWMDMFLAALDCYNTFIEQGMIKPNEIL--GTK   MR----S-SFLRAVEFFLKTIVLHDIS     1317
Query_GECKO  1256   -DN-FSLR   AVVQWMDMLLAALDCYNTFIEQGMIKPNEIV--ATN   TE----S-SFMTSVEFFLEKLALKDVS     1317
Query_DANIO  1254   -AP-FSIR   GTLQWMDLLLAALDCYNTFTNLRCLQLQRIL--GTC   EK----S-SFLPAVHFFLTELSMQDIQ     1315
Query_SNAKE  1255   -GS-FSLQ   AVVQWMDMLLAALDCYNTFIEPGMIKPSEIF--ASN   TG----S-SFLKSLEFFLGKIALYNIS     1316
Query_STERL  1268   -EP-FSLR   AVLQWMDMLLAALDCYNTFIGLHLVKPHEVL--GAK   VQ----S-GFLKAIEFFITELAVQDIS     1329
Query_MILII  1252   -EP-FSVR   AALQWMDMLLAALDCYNTFVRLAMVKPSELL--GTS   QE----T-GFMKAMNFFLKTLARHDIT     1313
Query_BAMBO  1254   -EP-FSVR   AALQWMDMLLAALDCYNTFIGLHLVKPSKIM--GTS   EN----S-SFLKALNFFLTNLARHDVT     1315
Query_ALLIG  1216   -EP-FSVR   AVVQWMDMLLAALDCYNTFIEQRMFKPENVL--GTN   ------S-SFLRAVEFFLETIALHDIH     1275
Query_GAR__  1255   -GS-FSVR   AALQWMDLLLAALDCYNTFIGLHFVKPQQIL--GTQ   EK----S-SFLKAIGFFIAELTMQDIS     1316
Query_TETRA  1230   ----LAIN[2]SYFREINVLLSQLDVVTFIVKNNFFSVQEELlqFED[2]-D----S--FCTNIENFFNSLFKIQES     1293
Query_PARAM  1183   -QNyTSLR[3]------ESLSWQCRIWQFIHDRNLFTLDEIIqfDDY[1]----iTTyNIHQFLRYFIVKNQKHLMD     1246
Query_SLEMN  1235   kTMlSEKQ[2]-----IEMLTAQIECFNFLLSRKIASFSDIKskLVD[5]GNlysFLaD--HIVIKFVKNERKSLDK[4]  1309
Query_STENT  1093   kST-TNIQ[2]-----IDFVLWSSEC--------------------G[5]GEel----E--KKISEFLENDKK----     1134
Query_NEMAT  1248   -GA-FSLK   ETLLWFHYLLAALDCYTWVFGEQLLPPADLL--RS-   ---kySS-VLFTSLSYFLSNLAMLGIQ     1309
Query_PISTI  1262   -GH-FSLK   ETILWFEFVLAALDCYTWSFSEHLLSPSWVF--TSS   EDnemPS-VLLVSLKFFLSHLAMFGIT     1327
Query_CORAL  1243   -ET-FSLT   TTKHWFDLVLAALDSYIWSFGQRLLTPSSIF--QG-   ---ktTS-VLWRTLLYFIEKLSLNGIE     1304
Query_HYDRA  1194   -AE-FSFS   STINWLENLLGALDCYSWIISARYLPAQKIF-kGNS   -T----S-KLFTAIVYFIDEISLVNCK     1255
Query_LATIM  1256   -EP-FSLR   AVVQWMDMLLAALDCYNTFIELRMINPREIF--GTT   SV----S-GFLRAVEFFLTRLATQDIT     1317
Query_SHRIM  1230   -GS--SLK   DINIFLEALSASLDGYCWLIGQGIVPAAKLL---NH   SE----L-QLFSATTYFINEMALLTVS     1289
Query_MCRAB  1227   -GV--PGN   LMNIYFEALLAGLDGYSWLIGKGIVPADKLL---KC   KE----S-RILLASSYFLQNIATKSLS     1286
Query_AMPHI  1272   --A--HPK   EWTSFLERLLACLEGNVWLIGHRLVPSTVLL---TH   RD----S-NLLAALQLFLLNV------     1324
Query_BARNA  1164   -GVyRDLR[3]SWSAWLDALLAALDGYSWALQVGALPTDRPI---SS   -D----S---------YLLDNIRFLMAE     1220
Query_COPEP  1194   ---gLSLN[1]-----LMEFQALFESYNWILGQNICSMDELEkiENS[2]KD---------------LTNLFIKKVY     1244
Query_DAPHN  1138   gNFpFSAR   ------------LEVYRHLLGLGVIDTDKCDsdVDF   ------S----DQFEIFAKSVERRLDQ[3]  1189
Query_STARF  1271   -QP-FSVK   LACSWFDHLLAALDCYTWVFGEKLLSPSVVY-tATA   RR---PS-VLFQSLQYFLSSLSLLDIT     1334
Query_STRON  1271   -QQ-FSVK   ATITWLEYLLAALDCYVWVLGESLLTPAAVFktIKE   GK---KS-VLFQTIDSFMK-VAQYDLF     1334
Query_SEACU  1263   -QP-FSLS   AICNWFDLLLGALDCYTWIFGQRLLTPTP-------   -S---SS-NLTKSLQYFLDDFAQHDLE     1319
Query_FUNGI  1207   ------VS   QIQAVLEQLYVALDTSIWALQSGFGELLIQE--GRD   SH-rsML----LCADLYLRHWATKSTD     1264
Query_DPURP  1320   -------K   EIEFWFKSLTTSLNIYLWFFSEGFIEPSALVntGGN   GS-tlLS----TIFNVFVKHYALVNDN     1378
Query_HALBU  1335   -------R   HVEFWMRRLQASLAVYIWYFGEGFTDPVAIS--KDG   SS-riIS----NAFYVFLSEYALVDQS     1391
Query_GLOBO  1161   -------K   SLT-WLLSLRTNLIMYRYLVDRGALTAQSIV----L   QS-diVT-----SIVTFLDAAADVE--[4]  1215
Query_PALUS  1211   -------K   TVSLWLRQLTTALSGYTFLMDAGIMEASAILd--QT   QS-qlLP-----AFVHFINVYADTDVD[5]  1271
Query_CALIF  1270   -------E   SGVKWANRLSICLEGFTWIIKNKLLESSLFFk--SR   LN-cfIS-----SIVYFIINMSKKEIR     1325
Query_GLOMU  1295   -------Q   QTKDWCSRFLSVLANYKWLIKEKFVPPLDLIt--SS   GI-mfRK-----AINYFLENLVQNDIE[5]  1355
Query_LOBOS  1243   -------T   AYRKWCTQLASILTNYIWVLNHPGSAELLLAq--LE   KG-aiMN-----TSTKFLEHCLLFPRE     1298
Query_RACEM  1226   -------P   SCLLWLRQLNCALDGYVWLIDRDMVTVMVVLn--QP   SS-vlLD-----ACRYLV-------IE     1274
Query_PVIOL  1296   -KAnLAIK   EIEFWYKRLSTSLNVYLWYFGQGFLDPQMII--GSS   TS-kiLT----TAFHIFPNEFALLETD     1358
Query_BOMBU  1140   -K-iENLP   -V-NGILPFLRSLDCYIWLIKGDLLNMQCLFdnS--   -N-sqKD-EIFNCARNFIHVMNRIKIE     1201
Query_AEDES  1224   -EKkFPII[5]----WLEYLASSLDMYYWLI-----KNSLLRnaE--   EF-lsAS-NLFAPVAYFLESVSNASMF     1285
Query_ANT__  1158   -SGiKDIS   -LASNVTPLLKALDYYMWLINKKLLPIEMLF--L--   -D-vdKQ-IIFFCIRSFACQFWQIITE     1219
Query_WHITE  1186   -PYvK---   -VISWLKNLHAALEFYSWLIGGIIIDPGLLLka---   -----DS-RFMQALKYFISNV---GAT     1239
Query_BEETL  1175   -TPiK---   ----WMQGLLCALDGFIFII------------kNKL   SS-tlSE-TCTTSIRYFLENV---STI     1220
Query_THRIP  1167   -QAlA---   ----WVFSLLASMESYIWLM-EVNIPPVTLF-----   -D-lnSS-KITPALMHFLESI------     1214
Query_TERMI  1203   -NEvALVY   -VVRWLRMVLTSFDCYVWLFGEELVKPSVLWgtEGL   DN-tsET-RIFEAAVYFIQHVCHRDAS     1269
Query_FLEA_  1146   -QTeF---[3]DILKWMKNLYASLDFYVWLF-----ENTLFTptTLF   SD-leES-KILVTIQYFINNIVNIDPA     1208
Query_MIDGE  1184   -HIkAGIY[5]EICKWMESFLSSLDFYHWIL-----TNDLVNseQKN   EL-fkHS-VVLKIIYFFLKGFINKDII     1251
Query_LAODE  1190   --------   DLLLWMDSLSTSIHMLKWLKENEFDVLSAVKnsENS[1]---------FLENLNDYISKISSCSAG     1244
Query_LOTTI  1232   -DT-FTLT   STMNWFDMLLAGLDCYTWVFGEDLLTPINIF-sGTG   SN---TS-KLFESLEYFLSNISLRSID     1295
Query_APLYS  1254   -KS-FSLS   ATTRWFDLVLAALDCSCWVFGEDLMTLSQLF-tGKG   SE---MS-CMFKSLEFFLRSVAFQDIE     1317
Query_SCALL  1264   -DT-FSLS   STLTWFDTLLAALDCYTTVFGENLLTPVDVF-gGVG   MN---QT-NVFKAATYFVDNIALADIE     1327
Query_CRASS  1238   -SS-FSVQ   NAVSWFDHYLAALDCYIWTFGESLLNPDDIF-sGEG   RK---QS-KLLQAMQHFLSQVVMEDMD     1301
Query_OCTOP  1216   -EH-FSLQ   KTCQWFDYFLAGLDCYIWVFENNLLSPEQML-qVA-   ----eSNsQIIKAVCYFLEYIAL--SD     1276
Query_POMAC  1245   -EN-FSLK   AVLSWFDMLLAVLDCYCWVFGEGLLSPAKMF-eSSG   FK---GT-RVFSSLQYFVTEMALGDLS     1308
Query_WHIPW  1217   --------   ---LWLSSFEAVLDLYTFGLRSKLISLEKWA-hCVH   GNenlGA-LFISTFSRFACVALSWPSS     1274
Query_TRICH  1224   -GT--SIY   ALVLWLETYEAMIDGYVWALNTNRISMEAWNdyLLN   EDprvKS-LL------FPCSLDLVNSI     1284
Query_AASTA  1143   -TTgVRHT[1]--VVWYETLAASTESYVWACAPTLGNASLIAdaV-L   LD---PV----VA-KRT----------[3]  1195
Query_AEUTE  1119   -IEgMKN-   ---VWYETLAASTECYTWACSITLGNSSLIDenV-L   LD---PS----ASMKRL----------[3]  1169
Query_PLASM  1194   -HTvIT--   --VDWLEQLSACIENYVWVVEMLGDKTRGIFeiCGS   HS---KK----KSVKRK-HTAEVSDTN[4]  1255
Query_BREMI  1181   -NTdIT--   --IVWMDQLSASIESYVWYVDILGEKTQVFLekF-I   KS---RR----ENLKRK-HAAEVNSTS[4]  1241
Query_OLIGA  1190   -DDqVSDP[4]DLANWMRYVTASLECYGWCSDVLGDDAKTII--FST   SG---GS----TALKRKLDELEPTGSK[4]  1258
Query_LIGNO  1308   -HP-----   --LGPLRALLACLECHLWSLGGGLLDPQQLLgcGGG   SR--lCQ----RLSDFFNETDPSLDPG     1364
Query_CESTO  1134   -QVlFDAN   --LQLIPQIVSDSESLISTLGGNTENPTGFIdfLDN[1]------S--------------------[2]  1178
Query_OVI2_  1403   --QtFQPR   ----WMESLTRTLSSYSRILSADLLNVRILRllLLS   AR----S-QLLKVLRCYVVWCRQQNVH     1462
Query_PRIAP  1261   -AP-ISMT   TALSWFDHMLAILDCYTWVFSAKLLQPTALFagGTG   KR---EErTCFAMVSDFMR--------     1318
Query_ROTIF  1261   -GPkFQFG   LLVTWLKLLNAPLDCYTWIFGQRLLNPKSLF---ST   KK----S-CIWSSVEIFLDKVASFSL-     1321
Query_SPONG  1245   -PA-PNIN   IAINWLQSLQAAVDNYRWLLKEQYVSPREIF---SS   TS---SS-NLFKSIDHFLKCIALVGIE     1306
Query_TARDI  1269   -AVqDSHK   SATETLKRLAAVVDCFTMSFDKVYVKADDMK---NR   PE--wKV--FGDCLSEFIDKVVLRLVT     1331
Query_TRICH  1229   -DT-FNLR   SMKNWYDYFLAALDCYSWVFSEKMMSPEQII--GDK   AQ-lgSC-KIFGSIEFFLSKLCLHSLD     1293
Query_OSTRE  1221   ----YSER[3]-------QFIAVIQWSNWMIRMRHAGGEFFPatRTM[1]----------FRAAGLFIERTAQNVEI     1274
Query_NITEN  1297   ------PH   ELQQGLSSLLRSLRWYEWALSQQLLTVAELVp-GQE   VE-etRPgHPVVCLRRFLSSTADETDD     1359
Query_CLUBM  1178   ------TA   KSMPELSTALHSLHWY---LGAQMIGLEEIEp-LRT   QI-fsR----INQFVEELPARCIPLKE     1233
Query_MOSS_  1245   ------LQ   QEQLWLTKFVRSIHWFEWALKENIISIKELLd-FER   DC-viHIaSSISAFLQSL------LKS     1301
Query_LIVER  1262   ------IE   EARHQLAVFTRSLHWYAWALRQEMITTRELLg-KTR   SS-fv-----VSSFLHRV------FTG     1313
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1305   AAVKLFPKVA--DGS   -sIFTPREREDYNQT--KCTVIVRTLSFIC-TMFGNYK-TES   LKV   IPTSFWSSNL-   1367
Query_LINGU  1314   AAADLFTKGL--D-R   -kIFTPWEVEEYNRS--KCTVIVHLLDFLC-STISKSG-SHL   EQV   VPSNIWCPDL-   1375
Query_LIMUL  1296   ----FISVEI--IAA   NdLLSPKEKGEFKRV--KCTVIVRLFNFLT-VLIQNFS-KEL   SSC   TPQSLWSKSL-   1355
Query_SPIDE  1301   PLTALISKHY--IEE   SdVFIPGETLKYNRL--KCTVLIRLLNFIT-SVFATDA-KWF   HDK   FLKVLWNSNL-   1364
Query_SCORP  1299   RVQAFFTHES--VK-   KqIFSPKEKDEYLRL--NCFIIVRLINFVT-TIVENYH-LQL   NTI   -PSTFWSINF-   1360
Query_TICK_  1252   SPQELLRDVE--SEP   ---LTSSKLSAFNLA--KCSVTVRLFNFLA-TWMETSH-GSM   ASV[6]------SRAF-   1312
Query_MONOS  1243   ---------------   --LYTHLDAALRDIV---CGVLVRNLRFIE-VVLRNPAlTAA   ---[1]IPRQFLNAQT-   1289
Query_ROSET  1215   ---------------   ----TPAEAAAYRYL--KCTVMVRVLDFVR-TVLRVDA--SA   ---   LPTKLWSGEC-   1257
Query_BELCH  1335   AAARCFESHG--AAT   --LFTPRETEDYNYG--KCTVIVRIMDFFT-VLLGAFP-QDA   YKV   VPDAVWKSPL-   1396
Query_CIONA  1296   Q-LTSFSAD---SKG   ---LTPKEISSFNDL--KFTVVIRIMEFFS-ALLKN------   YNQ[5]VPKSMWEQEYy   1355
Query_CAECI  1320   AAEQCFRTG---PSG   N-IASPQEREDYNYS--KCTIIVRIMEFVS-MILENCQ-QDF   WRL   LEKDLLNNNL-   1381
Query_HUMAN  1308   AAEKCFGTG---AAG   N-RTSPQEGERYNYS--KCTVVVRIMEFTT-TLLNT-S-PEG   WKL   LKKDLCNTHL-   1368
Query_MOUSE  1307   AVEQRFGSG---APG   P--PSLHEEEKYNYS--KCTVLVRIMEFTT-TLLIA-S-PED   CKL   LEKDLCNTNL-   1366
Query_XENOP  1333   AAEQCM--G---SKS   S-VFSPHEIEAYNYS--KCTIIVRIMEFIT-MFIDICQ-QDS   LKI   LENSVFNEPM-   1392
Query_PLATY  1312   AAEQCFGTG---SKG   S-MFSPQEREEYNYS--KCTIIVRIMEFAT-MLLNT-Y-QDT   WKL   LEKDLLNATF-   1372
Query_TASDE  1315   AAEKCFGTG---IKG   N-VFSPQEREEYNYS--KCTIVVRIMEFSS-MLLNT-C-QDV   WKL   LEKDLLNAYF-   1375
Query_CANAR  1309   AAEQCFDSS---SKG   S-MFSPQERDKYNYS--KCTIIVRILEFVT-VILEMCQ-QDF   WKL   LEKELLNASF-   1370
Query_OSTRI  1279   AAEQCFDCG---LKG   S-VFSPQEREVYNYS--KCTIIVRSMEFVT-MVLESCQ-QDF   WKL   LEKELLNASF-   1340
Query_TURTL  1318   AAERCFDTG---SKG   N-MFSPQEREEYNYS--KCTIIVRIMAFVS-MILETCQ-QDF   WKL   LEKELLNANL-   1379
Query_GECKO  1318   GAEQCFNTA---SKG   N-IFSPQEKEDYNYS--KCTIIVRIMVFVS-MILETCQ-QDF   WKL   LERELLNKNL-   1379
Query_DANIO  1316   AARACFRLG---NAG[1]S-HFSPRETEQYNYS--KCSIIVRMLEFST-MVLQKCP-QDL   WKL   MEKDVFNSSL-   1378
Query_SNAKE  1317   GAEQCFSSA---SKG   DmLLSPHEREEYNYS--KCTIIVRIMVFGS-MILETHQ-QHV   WKL   LEKELLNENL-   1379
Query_STERL  1330   AAELCFSVG---PKG   A-LFSPKEREDYNYS--KGTIIVRMMEFVS-MVLENCQ-PEF   WKL   LEQNILNPTF-   1391
Query_MILII  1314   AAEECFTLG---VKG   V-LFSPREREEYNYC--KCTIIVRIMEFAC-LILEGLQ-QDC   WKV   LEKDLLNSDL-   1375
Query_BAMBO  1316   AAEESFAYG---IKG   I-LFSPREREDYNFN--KCTIIVRIMEFVS-MVLEDCQ-QDF   LKM   LEKYLLNSTL-   1377
Query_ALLIG  1276   AAEQCFYTK---SKD   N-VFSPQEREEYSYS--KCTIVVRIIRFVS-MILETCQ-QDF   WKL   LEKELLNLNL-   1337
Query_GAR__  1317   AAERCFSSG---SEG   S-LFSPREREDYNYC--KGTIIVRLMEFVT-MVLDKCS-QDL   WKV   LERDILHLTF-   1378
Query_TETRA  1294   -----YLKSDglLME[4]DaVDIDIQENQYQAV---INLFIKICEFMS--LLSNDIlNFT[2]LDV[2]---NITLENK-   1358
Query_PARAM  1247   -ESPLFDMIMtlTQD[1]---FNPSQNFQNDAG--------NIDNWLA-TSLPEKLvEQL[7]YKF[4]---HIDVEQY-   1312
Query_SLEMN  1310   -KEDK----------   ------EILTKSRMV--LITQFLRYLD---------------   ---   ---QLDEESL-   1339
Query_STENT  1135   ---DE----------   ------DIIEALRTL--KV--LIKYVEFYP------------   ---   -HPQVNQDFL-   1165
Query_NEMAT  1310   AAIKCFPK----TTH[1]TdVATPKEMDEYNRV--KCTVIVRIMNFFT-VLLQAYP-DQG   FKI   VPTSLWSDDL-   1372
Query_PISTI  1328   EATSCFTSGP--SRG[1]GdIFTPKEIDDYFRA--KCTVIVRVLNFLT-ILLQDFP-QET   LKT   LPPEFWSEDL-   1392
Query_CORAL  1305   KAFECFPTTS--DGV[1]ThAFSPSEIQEYNRA--KCTVLVRIIAFIT-LLLQMNA-KEY   LEI   VPSELWTGKF-   1369
Query_HYDRA  1256   NMLMKFMQD------   ---ILPSECDTYDRL--KCTIIVRIIYIIS-LTLEDEL----   WKA[5]IGDKMFSNNF-   1314
Query_LATIM  1318   AAEQCFTAG---SKD   T-MFSPREREDYSYS--KCTIIVRILTFVS-MVLEKCQ-PDL   WKV   LEKDLLTPAL-   1379
Query_SHRIM  1290   EYMKNHDAFQ--N--   -aVATPYETEATAKL--KCTVIVRLLDFVS-VVLQEGP--KV   ASL[2]-QHSVWSATM-   1350
Query_MCRAB  1287   QYVKDNIRVS--ETG   -aIGSPLEYEASNKL--KCTVVVRLFDFVC-VLLQDGV--QS   VQL[2]-EYSIFTAEM-   1349
Query_AMPHI  1325   ---------------   --VPEPRHVLSSSKQvsNCTLIVRLMDFMT-VVMKDKE--TS   QSL[2]---SIFCPNF-   1373
Query_BARNA  1221   AFKPSFEDSA-----   ----EAGRWAPVRRL--RCTALVRALAVIE-ALLRAAPrAVT   ---   ---DIWSLPV-   1272
Query_COPEP  1245   GTEEIEPEGM--DES[4]KdPLSPGERRRLEKE--KGCTIIRLFTFLT-SVLKLEKgRNL   --L[2]------QKEFa   1308
Query_DAPHN  1190   TAEVPLRDTT-----   --------SESFLAW----------LSFLT-EYLTKSQnRKL   ---   ----IRTKTL-   1228
Query_STARF  1335   GAARGIEYGG--SKT[4]Q-VFTPRESEEYNQE--KCTVIVRLWDFLT-VLLENYA-SDA   VKV   IPKSIWQDSL-   1401
Query_STRON  1335   AFAKGC--GS--MET   --AFTPREEEDYNRT--KCTVIIRLFYFMG-AMLGSHT-KEA   LEV   IPSKLWTGLF-   1394
Query_SEACU  1320   HVALSL---------[3]RgIFTPWEQERYNQS--KCTVAVRVMDFLT-VLLGNFP-SEA   MKV   LPSSFWVGSF-   1379
Query_FUNGI  1265   -----DA--------[1]FaRSTPSEMIGYHQV--NYNILLSLCRFTS-LQMKLSTkDNQ   ---   -IIS---QHY-   1312
Query_DPURP  1379   -----KSYEQ-----[1]LsSMTPKEIDLFNRM--KCQVITNLFGLVI-LLLE-KYnFQE   ---   -VIEISGQPF-   1431
Query_HALBU  1392   -----HP--A-----[1]-aSLTPREIDQFMTL--KCQTLEHVYALFS-MLLE-KYnYSQ   ---   -VMETGGLQL-   1441
Query_GLOBO  1216   AFKSLD---------   ----------AFNNN--RAFVFLKVVEFVN-SVFGKT-l---[6]WSP[1]----FECENF-   1262
Query_PALUS  1272   AAKH-----------   ----------QYNQK--RSLTIAKAVDFLL-SVLQRDGi---[6]EEA[1]----IWSSKF-   1317
Query_CALIF  1326   ----MDDNNL-----   -iFYSVNEISKLNNM--RYDITFKIFNFLK-YLIENSMdI--[7]KEH[1]----VFSNDF-   1384
Query_GLOMU  1356   DAMEIDDGEY--FIK[4]EnILTPAQKLQWLTR--KIQIIEHLIGFNI-VLFKET----D[7]WNS[1]----INGKQF-   1424
Query_LOBOS  1299   -------------LK[4]VtTLTPSERQQIMDQ--NLMTARKAISFVS-IVASQSQsRGD[7]QAS[1]----LLDSRF-   1360
Query_RACEM  1275   -------------VS[4]SnDFGAFDRRRLNAL--YAYATYRLACFLT-CLASSNQs--N[7]VTD[1]----LFSEDL-   1334
Query_PVIOL  1359   -----KL--S-----[1]FsSMTLSELEHYNRL--KCSVIMNLFGLFI-LLME-KHnVEQ   ---   -VMSYGGISF-   1409
Query_BOMBU  1202   NKEDNLI--I-----   ----LSKEIEDLQTL--QCKAILNIFDFTQ-ILLNF------   -DG[3]-PDFFFNKDF-   1252
Query_AEDES  1286   ELMTVIRPSI-vEDS[4]EfRVCTDKIIKFNLL--KCTIIVRIVDLLS-LLLPLRCcRSV[2]-AF[3]AAVSQFV--V-   1357
Query_ANT__  1220   SSAETAATTI-----   ----KSRELEQLQVL--HCRVLMTTLNFIQ-VLL--------   -DV[3]LPESLWDEPL-   1271
Query_WHITE  1240   SCQHFIEILY-----[2]WsAEVPQLDSNLYSS--VATEVLAQLFKFL-TISHN------[1]-LK[3]FCEILKEP-L-   1299
Query_BEETL  1221   TLDEAFAIIE-----[2]MwMFTTFEKLHFDKQ--RSQAVIKILQLFT-IC---------   ---[2]ITAFYTSH-L-   1273
Query_THRIP  1215   -----VGTPI-----[2]PsLWRPKAMEQFNYA--KCTTIVRMIDFFT-AVLDI------[1]-QG[3]S-PFWSSN-F-   1268
Query_TERMI  1270   SVTCLLSTDT-----[2]VlEITPWEAESFDRT--KCFVTIRLLDMIS-TVLKS------[1]-EA[3]PHSLLNLQ-L-   1329
Query_FLEA_  1209   DLPAFLDASVgmNIS[4]E---------KFIEF--KYNTFIMILTFVS-KILFC------[1]-DA[3]ITSFSWLNRL-   1267
Query_MIDGE  1252   DIVELANESH-iNTS[4]EaRMYNRNVDNISVI--RSLILVKIIDLFT-ISIEI------[1]-EV[3]LTKYQFE--I-   1316
Query_LAODE  1245   SIDWLSGESSpfCPQ[1]------STTDAYDRA--RRQVTFNLLAFLA--ALEEDLiTIT   ---   ---DLLHEDV-   1299
Query_LOTTI  1296   EVSKLFK-ES--S-V   -eYFTPREIEDFNRN--KCTVIVRITNFLS-VLIGKYP-KDA   IKV   IPGTVWSDTL-   1356
Query_APLYS  1318   GVSKLFS-SS--K-V   -tVFTPWEKDEFNRM--KCTVIIRLWNFLS-ILLARHS-QEV   LKV[9]VPASLWCEDL-   1387
Query_SCALL  1328   TVANLFPNQP--E-T   -tFFTPREIDEFNRE--KCTVVVRLLNFLS-VVVATSGaKDL   NKA   VHPEIWSGAL-   1390
Query_CRASS  1302   EVGKMFRHEP--L-T   -dVYTPKERDEYNRT--KCTAVVRTLNFFC-VLVARSPsS-V   KKL   LPSEILGKDL-   1363
Query_OCTOP  1277   LKTLTKNADY--LP-   ---FTPKEEEVFNRN--KCTVIIRIFSFLT-TVLKKCKdAEN   SSV   IPNEVWTPHL-   1337
Query_POMAC  1309   VAAKQFQRKA--E-A   -vFFTPKEKDDYNKV--KCTVLIRMWNFLT-VLIARYP-SEA   FKV   IPQTLFCDDL-   1370
Query_WHIPW  1275   DINFEFTD------N   --------RSQLMKR--RRCAFLSSLDFLT-VFVQQYN-ENF   EDQ[2]--AVVEDRNI-   1326
Query_TRICH  1285   VIGRHFPNLS--KAQ   --------RNRLNKT--RCKSFVALCKLIA-ALISCQN-LNN   YQG[2]--KLLLYPNT-   1340
Query_AASTA  1196   EA-VEGPHRLfvAIA[4]SgSREAPPSMARNRA------FLGVCNLVA-LALEHPTvVPE[2]RHV[4]--TGTGD-RW-   1264
Query_AEUTE  1170   NADVQNSHRVfvAAK[4]NvDE---PTQARNRA------FEGVCKLVT-LAINSETsLPL[2]KQV[2]--SGLVDNKF-   1235
Query_PLASM  1256   QQTLSWAILTflNCD[4]YrSKRSRWTEAYLSV------LVSLCNFVK-AIMSHGNiLLR[2]LDL[2]--QSL------   1319
Query_BREMI  1242   EETLSWAISTflSYD[4]DsSKRSDEIAAYMSV------VISICTYVK-SSMENENmILC[2]IDV[2]--RSF------   1305
Query_OLIGA  1259   HTTISWAVKRflEIH[4]SkTTTIERDEQYKTR------HQLLEAFIQ-SAEALSNlLLK[2]RST[2]--SNVVSEFQ-   1327
Query_LIGNO  1365   RLALSASVSA--SSA[3]----SISEIRASAVT-iRCQLIVRTAALLQ-LLAGKQP--DC   LKD[1]LQGKENSNGL-   1428
Query_CESTO  1179   --------IS-----[1]NlLSVSKFLETIHSA--NCSVIAKLEAYCA----SPERrTSG[9]LSL[8]LSSIFTESHSv   1248
Query_OVI2_  1463   -----IDVSQ-----   ----------EFKTIfeRTSCTLAVCEFAVk-MTPNSApEHS   -EF[4]------EHDIw   1509
Query_PRIAP  1319   --PSGEERRR--RRR[4]R-HVHPGEMEKYNRS--KCMVVIRAMDLFS-ILLN----MAS   LQV   IPDSTWSMEL-   1380
Query_ROTIF  1322   --QNFLSHNY--TEA[2]--LFSPNEIDEYNGL--KCTVMIRLMDFLE-TIFANYS-KES   LDV   IPMSLWNSNF-   1383
Query_SPONG  1307   GVVNSPEVAM--ATS[1]------HERGEYQRN--KWCTIELLLTFIE-LLLTKHT-SET   VSF[2]VPDSLWSQEL-   1367
Query_TARDI  1332   DPASLLATES--LRT[3]-------------------GILIRLTDLIM-TVHTKYG-RDG   IAV   LPPVVASKSF-   1381
Query_TRICH  1294   ETVKLLYS-----KS[1]LqLFTPREIEEYNMC--KCTTIVRLINFLA-VLLQMSP-TSI   IEG   IPRQIWSQNL-   1355
Query_OSTRE  1275   ---------------   -----SREKNAVQNQ--HLYVTREIFALLD-TILKSALrGAE[4]LAL[4]------DNGY-   1323
Query_NITEN  1360   -----SALEA-----[1]LrSLTPSERESFHRT--RCSTIMGVLSFIH-QVVKRSLgEST[9]SAQ[7]DVVPVEDTSF-   1433
Query_CLUBM  1234   -----NDS-------   --------------------AILHLMALFK-DIIQMGIgASE[9]DDK[5]EILPLHYDKL-   1284
Query_MOSS_  1302   -----QDSIH-----[1]FlDLTPGEQDELQKL--RYEATTHLMKFCS-QVLNAQSgYAE[2]---[6]IFMPDQSRYF-   1364
Query_LIVER  1314   -----HDSNS-----[1]WsSVTPGERESHRKL--VCGAIVHLLGYIR-QIVDSELgTRE[8]VDK[9]GLLPFSEPDF-   1388
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1368   W---DTLLLAVLQPT-QLGFNMGQ-VEVMN   HLPEQT   LGTLQILM--RSLPRNLAQE   LIESVSRHLNMS-sC   1430
Query_LINGU  1376   W---ILICDCVLNPS-ALGFNMGD-VEVMM   KLPEQM   EQVLRIFS--RYLPPKALDM   FRRTLKSKLDSG-sY   1438
Query_LIMUL  1356   W---KCISLCVLDPT-SIGFNMTD-LEIMT   QLPHET   THLLKAFY--K-LPSPVLED   FAHSLERHITSE-eS   1417
Query_SPIDE  1365   W---ELILLLVVNPT-SLGFEMTD-PEIAT   KLPMEV   SIVVSKLS--K------KSE   FHAFCLEYLKSK-nG   1421
Query_SCORP  1361   W---NCIYSSVLNPK-FLGFDMTD-TYIFN   QFPKET   RTLLKVFY--EYMPKPLYNN   FNEHLYEYL--E-sK   1421
Query_TICK_  1313   W---RCTVVCVLSPS-DAGFDLTD-LEVAN   QLPNEM   LRFLNCVS--RTFPKLLADD   FVGAHADLVDHR-pE   1375
Query_MONOS  1290   M---KMISTLLFNPA-HLGFNAAS-QSIMA   QLQQVL[1]-QTLDAVA--------HAST   -DSAIRQELEQDlaT   1346
Query_ROSET  1258   L---DFICKAVLMPI-MLDFDPSD-PEIEE   KLPKKT[2]--VLASLC----LSGEPRTN   LVAAMRTTLASD--E   1317
Query_BELCH  1397   Y---DMVVTAALEPT-ALGFNMGD-VEIME   KLPQET   KQLLLLMC--RSLPSQQGAL   LKSQVAAKLSRR-sS   1459
Query_CIONA  1356   I---GLLTYAVVSPY-VLGFTSTK-ETDLE   NLQ---[3]TSALQAAL--THLPQNKNNL   LVASIAERLAGI---   1416
Query_CAECIL 1382   V---ELTVLTVCDPS-SIGFNVAD-VQVMK   NLPDVC   IKLLKGLM--K---SPFKND   LEFSIKKRITSE---   1439
Query_HUMAN  1369   M---RVLVQTLCEPA-SIGFNIGD-VQVMA   HLPDVC   VNLMKALK--M---SPYKDI   LETHLREKITAQ---   1426
Query_MOUSE  1367   M---QVLVKMICEPM-SLGFNIGD-VQVMN   HLPSIC   VNLLKALR--K---SPYRDM   LETHLKEKVTVQ---   1424
Query_XENOP  1393   W---ELIAITVCDPS-SIGFNTAD-VEVIN   NLPNIC   IKLMKALG--N---TSYRSS   LEVSLKKRVTLQ---   1450
Query_PLATY  1373   M---ELVVEIVCDPS-SIGFNVAD-VQVMS   NLPDVC   VSVMKALK--K---SPYKDA   LEGNIKKRVTAQ---   1430
Query_TASDE  1376   M---ELLVKTLCDPS-SIGFNIAD-VQVMN   NLPNVC   VNLLKALK--K---SPYKES   LEINIKKRITPQ---   1433
Query_CANAR  1371   I---ELVVMTVCDPS-HVGFNTAD-VQVMK   KLPDIS   VRLMKALM--K---SPYQES   LKLCLKERITPQ---   1428
Query_OSTRIC 1341   I---ELLVMTVCDPS-HVGFNTAD-VQVMK   NLPDIS   VRLMKALM--K---SPYKES   LKLSLKERLTPQ---   1398
Query_TURTLE 1380   I---ELLVRIVCDPA-NIGFNTAD-VQVMK   NLPDVS   VRLIKALM--K---SPYKES   LELSIKDRITSQ---   1437
Query_GECKO  1380   I---ELTVRTVCDPK-SIGFNTAD-VQVME   KLPDIC   VRLIKALM--R---SPYKES   LVTSMKKVITPE---   1437
Query_DANIO  1379   F---TLVVLAVCEPS-SIGFNMAD-LEVMT   HLPEVC   FPLLKALA--S---APYRTQ   LESCIRMRITKQ---   1436
Query_SNAKE  1380   I---KLVVKTICDPE-SVGFNIAD-VQVME   NLPSIC   VRLVKALM--K---TPYNES   LENSIKEIITMK---   1437
Query_STERL  1392   F---ELITTTVCDPS-SIGFNMAD-VEVMK   NLPDIC   VRLLKALA--K---SPYKGQ   LEISMKRRISYK---   1449
Query_MILII  1376   F---ELIAMVVIDPT-SIGFNMAD-VQVMK   NLPNVC   VKLLNAVM--K---SPFKET   LVDSVRKRITAQ---   1433
Query_BAMBO  1378   F---ELVATVVIDPP-SIGFNMAD-VQVMK   GLPDIC   VRLMKAAL--R---SPYQES   LECSMKKRLKSQ---   1435
Query_ALLIG  1338   I---ELLVVTVCDPS-AVGFNTAD-VQVMK   NLPDNT   VRLMKALA--K---TPYKSS   LELSIKKRITSH---   1395
Query_GAR__  1379   F---ELIATTVCEPS-SIGFNMAD-MEVMT   NLPEVC   VHILKALL--K---SPYKTQ   LEQSIRRRVSQQ---   1436
Query_TETRA  1359   Y---SIFLQMLLKPNhLIRGMNID-IFLTK   --SELY[4]TPVIKALQeiTQEGVDFFQD[1]VLKALEDIFKGNkpQ   1428
Query_PARAM  1313   Y---NLSIGIIIRPS----YFLKG-IQLDL[3]QVDSVV[3]QRLLQGTIk----FTEFRQV   ----LNEIFDNQfvS   1373
Query_SLEMN  1340   -----LAHADLFQKNcNFNLIHFC-VQIM-   ----VI   -PLIHQLQ-------VLVE-   DQELERQLRNQSelL   1390
Query_STENT  1166   -----LKF--LLTPN-----------KIG-   ----FF   -PGAESQE-------FIVL-   YDKLSAKVIN---fL   1201
Query_NEMAT  1373   F---TVVFSCALEPA-LIGFNVGD-VEVIE   KLPEQV   SMT-----------------   -EAAWHISMFLGfsR   1420
Query_PISTI  1393   F---EVVLSCVLEPG-VVGFHMGD-VEVVE   KLPEQT   GVLCKLLV--QRLSGECLER   LKGTLEKKIALG-sH   1455
Query_CORAL  1370   F---EALCTSVLEPQ-SLGFNVAD-LEVLE   KLPEQT   FELCKSLA--SKLPKQFLIH   LTTTLMSKISLG-gR   1432
Query_HYDRA  1315   L---KVIVYCMVYPQ-CLGFQVSD-LEVLE   KLPEEM[4]EKMRKFAP--DHFLNSFYNV   LSSVLCEKQY-----   1377
Query_LATIM  1380   L---ELTVRTVCEPA-SIGFNTAD-VQVMK   NLPEIC   VRLLKALM--K---SPYKDT   LEASIRKRITAR---   1437
Query_SHRIM  1351   F---SLLTCCTLDPA-SVGFDIKD-TEVSK   HLPSRL   EEILGILA--RRGPSEVIGE   YSDYLEKQLQKP--E   1412
Query_MCRAB  1350   Y---GLILQCVLNPS-SVGFDVKD-TEVTK   HLPSRM   EEVLKVMQ--KKVPNVVSAA   FIQTLEKYLGSP--Q   1411
Query_AMPHI  1374   R---KIFVTAVLEPA-SVGFDLRD-TEVNE   KLLPRL   VDTLQTFQ--RYLSQEQKLV   LRKSIDSLMSGD--N   1435
Query_BARNA  1273   ---cDTVAAAVLAPA-ELGFDVSE-AEVVR   ELPAAL[1]-RLLRSVA--RYADAETRRL   LTDSLARRAAGR---   1333
Query_COPEP  1309   V----LLTQTIMEPE-KLGFNLKS-----N[2]--EDQF[4]IELLRAVR---SSWPEYFSK   LENFFTKVYQED-eL   1369
Query_DAPHN  1229   -------IRTCLLPS-RLGFSDAS------   GSQQQF[4]ETILPFLA-------PDKEV   ----IESVLNSE---   1275
Query_STARF  1402   F---DLVVTCILQPT-VVGFNMAD-VEIMN   NLPKTT   RKLLRLMS--THLPANLRNN   LSKAVTRSVTKS-sH   1464
Query_STRON  1395   A---EILCGCVLKPA-SLGFNMAD-TVIMN   ALPQTT   NRVLDLLY--TKLPTEQLKK   LETVIHKRITAS-rS   1457
Query_SEACU  1380   S---NLLCSCFLNPV-SVGFNMAD-VEVMN   NLPKTT   SQLLELMA--ANLPPKCCEE   FKTSVAKRTSSN-rN   1442
Query_FUNGI  1313   L---TLVLTSILSPH-MLG-STIS-DINQI   ---DRM[4]VQVIVDI--hRCNNSDVLEL   LKRSLEEVL--SqpK   1374
Query_DPURP  1432   I---QLLTKSLMDPQ-SIGFTNYI-DEKVL   ----EL[4]REEIRNLMsrSFKFPPLIEI   IDHIFSLFIA-VrpK   1496
Query_HALBU  1442   L---RVVGVSLLTPA-VVGYRTID-NNN--   NASNSG[4]VNEL-----------PLNLI   VDRVCQLIV-NQgdK   1497
Query_GLOBO  1263   Y---RMLALTMFAPE-KLGFDVGL-DA-TK[5]RLVEVT[4]RC-------lEQPF------   -KMRLMNTFA-----   1317
Query_PALUS  1318   F---HIIATCVFRPH-ELGFAIES-QE-VK[5]RMESVL[4]TN-------lPKMW------   -QKVMVQSLA-----   1372
Query_CALIF  1385   I---DLLNCSLYNIR-LLGNDLEE-KD-TQ[5]TLQKLL[4]NNELFIDYkvLISS------   -LAEQESIYS-----   1446
Query_GLOMU  1425   Y---KLLGNCLLRPN-IIGLDQLP-DNIVK[5]---QVE[4]LF-----KklSINN------   -VTQLAQDLA-----   1479
Query_LOBOS  1361   F---AVLAACLFNPD-SIGSDQLL-KG--S[5]LRDELK[4]VI-----SklPMSI------   -TRSCGEMLM-----   1416
Query_RACEM  1335   S---RLLAHIVLLPEqLVEVVQAE-QGALS[5]--EKVL[4]LFLHAMVGkgPPKF------   -LRTFAKSVN-----   1396
Query_PVIOL  1410   I---HLLVLSMMDPK-KLGFDTIS-PDNMI[5]KSSEDV[4]KQEISNLLgrQFKFPPLSEI   TARIIKILIENGgaK   1484
Query_BOMBU  1253   F---ELIAKCIMCPQ-VIGFDLKN-LEITE   ELFVIM   GNLLQSIM--HKNNISLLHL   LKCELLIYVEKH---   1313
Query_AEDES  1358   --------DLVYQPQ-KLGFDFKScRETME   QVPKRV   MQLLDKIE--GFENLRFKET   VYQKLSEKLSSS---   1414
Query_ANT__  1272   S---ALTMKSIMNPR-TVGFDVKN-IQMTD   KLPHVL   ETLLNSMR--SKYSYALPDT   FTKCLVNSVREH---   1332
Query_WHITE  1300   W---KLIVLCTLRPS-TTVARGVC-SETSL   KFSEVV   AELLRSLD--KYLTQKLQAE   LHDAFSTEVSVV---   1360
Query_BEETL  1274   W---NLLVNFVFNRQ-SLGFECFQ-PQDDD   NES---   --VLKGFL--NSLPQSQVLQ   VGNVLSKYVTDN---   1329
Query_THRIP  1269   W---KLLINCCVWPS-NLGFDSTS-KNVEE   NLSITV   QRFLRSLN--THAPASAVND   LKCILKDYLSSV---   1329
Query_TERMI  1330   W---QLIVACVLHPG-SLGFDLYN-TEVTD   GLPKNL   IQLLGVLQ--ETLPSSQLEE   LNSELGKAVSSM---   1390
Query_FLEA_  1268   S---NVVVNCLFCPQ-DLHFNVEN-NYIGQ   NYLIKL   NEFIENLR--QYKNDTYYHQ   IEGSIIEKLNEC---   1328
Query_MIDGE  1317   F---DITKQLIFLPQ-RLGYSYKT-KNSLI   KLPERI   IKFISSIN--DHAPNEFKES   LIEVLRTEVITN---   1377
Query_LAODE  1300   W---QLLSRCLSSPA-FIGFTTVSqNQLNS   -LRSAL[1]-QLLPKLQ--KHLGDNILKE[1]FVRPIITKLASSlpD   1364
Query_LOTTI  1357   W---ELLIKCVVQPQ-SVGFNMGD-VEVMN   NLPSEV   PTKLMFVG------------   --------I-TN-fS   1400
Query_APLYS  1388   W---RLLCDCVLRPA-SVGFNLAD-TEIML   NLPKEM   SQVLAVVH--KHLPAQLLKQ   FKVTLRRLL-ET-eC   1449
Query_SCALL  1391   W---EMVCMCVVEPG-SVGFNLGE-VEVMK   NLPREM   EQVMMVMS--KKLPTKLFDN   MKHQFTKTL-SG-rR   1452
Query_CRASS  1364   W---NLISMCVLQPS-AVGFNLGD-VEIIK   NLPTEM   EQLLKVLN--AKLQPDVLAA   MCKEFGRLL-QG-dR   1425
Query_OCTOP  1338   W---RLICECVLQPE-YVGFNLTD-LEVVN   KLPDEL   QNLLNIFL--KKLPAEFKQQ   MRESIKNQLV---nQ   1398
Query_POMAC  1371   W---RLCIRCVVQPS-SVGFNMGD-IEIIT   KLPHEM   MQVLKVLS--KQLPLPERKE   LVKILEQEL-KG-kG   1432
Query_WHIPW  1327   W---CTVIKGIFEPE-SMGFSDHA-VETED   LYSAKI   -PALFGAL--KA-NHHLCAS   ISPIFVDHLKSRhpW   1388
Query_TRICH  1341   W---NAFLKVACDPK-SLGFRLYS-KQFFE   CISQDV   -RLLFLAL--KN-NTFLVQQ   VSTSMKEYLTQFypW   1402
Query_AASTA  1265   W---DLVVLAAVDPK-----AATL-FELST   --PEAV[3]VRMCKASLehQ-LPFRRV--   ----VAKHLESS---   1317
Query_AEUTE  1236   L---EFVVQARIEPG-----RVGL-KEFSD   --TQV-[1]MNLCQSLLrrQNLLFRNV--   ----LLQVLSNS---   1286
Query_PLASM  1320   ---rDSVVGKLVEV----LFHRAD-LQLHN   --ASSF[3]EEFCSYVTmhSPQWAHQMQK[4]MLIKLHQSLAD----   1383
Query_BREMI  1306   ---rRAVTERILQA----LLRSTS-MSLHD   --TSLV[3]EDFCSYLTvhSSDWANQFQK[4]ILLNLKQCLAH----   1369
Query_OLIGA  1328   LegrDDVIADLILSV-IGLLVSGE-VS-GS   --LKLL[3]KSFCADVAerRSSLATSLRK[4]ITSKLLKTIQDD---   1397
Query_LIGNO  1429   W---TALARATLRPQ-LLGFDAQD-PSLMA   KLLDCL[4]HAVARNLA--P-VGSSYRSS   IVSAIAAELELAgqS   1495
Query_CESTO  1249   W---SFLVLCLLNPV-KAGFQTDE--STAG   DMLKLF[4]ITFVKGSI------SKFLES[4]-----SDLFDIGssS   1310
Query_OVI2_  1510   WllsRCLVQCVLHSDiREECSVSC-QRYLE[3]SLPPPH[4]IRNFEALV----IGDSHAQN   LSSYVHQRIRDV---   1579
Query_PRIAP  1381   W---DVVLLSVLDPT-RVGFNMAD-IEVVE   KLPEQT   ERLLIALS--WKLPNAPKKE   LHSRVARSLSEK-k-   1442
Query_ROTIF  1384   F---GLVINMTLNPT-SLGFSLKN-LEIFS   NLSPKI   RQFYKLVI--QNLPTDKQNK   FRQQCQQIILDS-rE   1446
Query_SPONG  1368   Y---TIVMDSLFYSS-LVNTAPDD-PSTAK   RISEAA   VSCLTAMC--TSLPIDVKQN   LIAYIASNINTG-qT   1430
Query_TARDI  1382   A---DIIVRMVLDPK-AVGFDLHE-VD---   RLPRQM   SEFLKATT--TKVDALWTVM   IRQAVTLLLEKP---   1439
Query_TRICH  1356   F---EMIFLCILTPH-QLGFNVGD-IEVLE   KLPQQM   KDVCGLFQ--MKLPKDKREL   MIITLKELLSRN-rS   1418
Query_OSTRE  1324   VnacKVMSAAIFFPQ-----NLGNeQEASK   ---KSI[4]IQLLKTIS---------VDN[1]--SSKAKRLYED-aV   1379
Query_NITEN  1434   H---RLMCLLLLRPR-RLGFVTSD-LKQ--   --SDRL[4]MNLMREL---AEASSGFKGG   VGIALRSLMEEPggS   1496
Query_CLUBM  1285   Q---HLLCLLLFKPS-EMTTFTPE-SSG--   --AEEL[4]SSVLWSL---TRISSGFKDG   LVQVLREMLN-TssE   1346
Query_MOSS_  1365   Q---QLLCLVLLKPR-QLGLVISE-LQR--   --AQDL[4]SKVLTRF---TQLSSNLKSE   MVELLRRVIEGDdpA   1427
Query_LIVER  1389   F---ELMCLALLRPR-DLGFVATD-SHH--   --AEEL[4]LGVIQRL---SVVSSDFKEG   LTKRLMVLLA--dnK   1449
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1431  -D-LSSKLP-IP--LLEPN    QDYI--QLRLLVSGFEQLNSAGLLSN------TL--VSKSSEKLAEDLVNTV[ 1]M  1488
Query_LINGU  1439  -N-LLKKLP-HI--LSGPS    EDHM--TDSHLVAGYQLLHKTDFLLPALKDggHS--AENVSNNLLQTLFQAV    -  1500
Query_LIMUL  1418  sN-LMAMLPLDLRK---KY    IDHL--MLRHLLNGYEQLINCGFWPVGGRK-------DSSAEFISRELLNAV[ 1]K  1477
Query_SPIDE  1422  cN-LEKLLPLKLSD---PR    CNSL--QMTELLNGYDLLLKNGFHAND--L-------FNNPDKSLNNLLDSI[ 2]D  1480
Query_SCORP  1422  cN-IYNLFPLNLND---KD    LKIL--EIQEMLLGYEQLIENKLWPK--QN-------CNKSKKLACDMLSSV[ 1]E  1479
Query_TICK_  1376  cG-LDKVLAMPLTNLDCLR    NDLL--KWIGLASGYRLLAELRIIRVPDLL---------AARALVRSLAELL[ 2]R  1437
Query_MONOS  1347  wDeIKTLFPLKLAQNMGAG[ 3]SLTLlpLARRLATGYQLLQATGLLKGLLPKpaDQ---------VAKELVQSL[ 1]-  1412
Query_ROSET  1318  cD-----LLRMVRDVLPGR    -DIPapLIKNVVSGVQQLLAADLLDECFFT--DR-------DAFGRELFAAL[ 3]-  1376
Query_BELCH  1460  hD-LFSHLPEVSAGVAEL-    -DHV--RLTLMVSGYQQLHNAGLLLPA-----LG--TN-SGKSLASELVSAV[ 1]K  1519
Query_CIONA  1417  gS-LFESLPFPFGSTASHT    NNAI-----HLITGHQTLNSCEILLRIFQEfpNKipGANNVTSASEALIASV[ 4]-  1485
Query_CAECIL 1440  -S-IKDLCDVDLYN-QDAR    HYHA--RIGSILSACKQLHKAGLLNSVLQ---SQ--TAGLQFTVGSKLLSVV[ 1]K  1502
Query_HUMAN  1427  -S-IEELCAVNLYG-PDAQ    VDRS--RLAAVVSACKQLHRAGLLHNILP---SQ--STDLHHSVGTELLSLV[ 1]K  1489
Query_MOUSE  1425  -S-VEELCSINLCS-SGAR    QERS--KLLSILSACKQLHKAGFSHVISP---SQ--STALNHSVGMRLLSLV[ 1]K  1487
Query_XENOPU 1451  -S-IEELCSVDLYD-PGAR    FNRV--KLGSVLSACKQLYKAEFFNSIVP---EQ--VGGQRF--GSKLLSVV[ 1]K  1511
Query_PLATY  1431  -S-IEELCTVDLYK-PDAR    FDRP--RLASILSACKQLQRAGLLHAAVH---PQ--ASHLQRPGSTKLLSVV[ 1]K  1493
Query_TASDE  1434  -S-IEVICAVNLYN-SDAR    FDRA--NLASLLSACKQLHKAGLLHSTVE---SQ--ATDLPHPIGTKLLSVV[ 1]K  1496
Query_CANAR  1429  -S-LEDLCSVDLFN-SEAR    LDQV--RFGAVLSACKQLQKSGLLHSVLH---GQ--DEGSHFSVGSKLLSVV[ 1]K  1491
Query_OSTRIC 1399  -S-LEDLCSVDLFN-SDAR    FDQV--KYSAILSACKQLQKAGLLHSVLH---NQ--DEVPHFSLGCKLLSVI[ 1]K  1461
Query_TURTLE 1438  -S-IEELCSVELYN-PDAR    FDHV--RLSSILSACRQLHKAGLLHSTLQ---SQ--GSGLHFTVGSKLLSVV[ 1]K  1500
Query_GECKO  1438  -S-IEEFYGVELYN-PDSR    LDHV--RLGAILSACKQLHRVGLLHCTLQ---SQ--GVGLHCEAGSKLLLVI[ 1]K  1500
Query_DANIO  1437  -S-VEELCAIDLYE-TDTR    NSHA--SMNLLLSACRQLHQSGLLNSVLH---SQ--DASYGCSLGSKLLTSV[ 1]K  1499
Query_SNAKE  1438  -S-IEELCTVDLYN-PEAR    LDHV--RYIAILSACRQLHKAGVLHCILQ---SK--GADLQFETGSKLILVV[ 1]K  1500
Query_STERL  1450  -S-VEELCGVDLCN-VDAR    SDRA--RLVFVLSACKQLHKAGLLNSALQ---SQ--AASFGLSLGSKLLSVV[ 1]K  1512
Query_MILII  1434  -S-IEALCNVDLYS-PGAR    TDHA--GLRSILSACQQLHKADLLNSAID---CE--AGSLSSSLSLKLLSVV[ 1]K  1496
Query_BAMBO  1436  -S-IEDLCNVDLYS----Q    SDHA--KLNSILLACKQLHKAGLLNSVLK---CQ--IGSIVPPLNLKLLSTV[ 1]K  1495
Query_ALLIG  1396  -S-IQELCSVDLYD-PDAR    FDRV--RISSILSACKQLHKAGVLHSVLQ---SQ--GSGLHFTVGSQLLSVV[ 1]K  1458
Query_GAR__  1437  -N-VEQLCAVDLYG-SEAR    SEHA--RLALVLSACKQLHRAGLLNSTLQ---SQ--AASFGLSLGSKLLSVV[ 1]K  1499
Query_TETRA  1429  kF---EEIISDEQKLKENE    -------LKIFLEGIKLIFKLCSQSSRYTI------------RYIEKAISGI[ 1]N  1479
Query_PARAM  1374  fA----------QTITNKD[ 1]--------------IKDINELKFFLKGLKWisNQcsDDQDSNNFFQECIDQL[ 3]-  1424
Query_SLEMN  1391  lRkLVAIKNLDLAKIIKKD[ 4]MRC--iNLNHMTVDNVQNYHPKMLMSMCLG--YLklA--SIDPNIKEYLISF[ 7]E  1467
Query_STENT  1202  iNqGVKAENI--LKIWETA[ 4]ERLQktGLKGLTA---------LLESQTL----LktP--VLQKVTEEFAKNM[ 2]-  1261
Query_NEMAT  1421  cN-LFEKLPIPLAG-EDSR    VDHL--YLSHLVTGYKQLHAAGLLMDVML---TR--GSTKPLEYGSRLLNAV[ 1]E  1484
Query_PISTI  1456  cN-LFDQLPFPLSNPQEHQ    VDCL--RLAYLVQGYQQLHCAGLLLPALK---AQ--ARNASEEYAVKLFTAV[ 1]D  1520
Query_CORAL  1433  cN-LFELLPLPLSGEECDQ    MDHT--RLQLLVRGYQQLHRAEILLPALA---KE--GEWSAEKCSEELLKSV[ 1]Q  1497
Query_HYDRA  1378  -N-LLSHLPLDI----CNT    NNDIcdMLCQLASGYKLLHKCELFAHLLSV---------DVTDLKKKLLESL[ 3]-  1436
Query_LATIM  1438  -S-IKELCSVDLYN-PDAR    LDHI--RLGSILSACKQLYDAGLLNSAVQ---YQ--SGDFNLQLGSTLLSVV[ 1]K  1500
Query_SHRIM  1413  nD-IIKMLPLSLSE-----    TAHInlRSRHVIEGIGILEKSGWLAKCNVV--SY--------SVAKELVDWL[ 1]-  1468
Query_MCRAB  1412  yD-VLETFTSPLAE-----    GVNVplKSQHVIEGLQILQRSGWLAQCSVI--RK--------SLASELIMWV[ 1]-  1467
Query_AMPHI  1436  lR-LSTLLPLHSSA-----    VTYVapPTRHFLQGVAALHQTEWLDSLGST--D---------DLARHVTRWA[ 1]-  1490
Query_BARNA  1334  -S-VCQALGEALSSQHDQH    -------VAQQLRGLRTLQSSRLLSSAAPA---------ELRVSGEELFAVV[ 1]E  1388
Query_COPEP  1370  fS--PNHLLVGIIDGNPNM    ---------HAVKTYLRLTKHKMLTSPWNG-----------VQLLNELLSKL    -  1418
Query_DAPHN  1276  fR-------LEYLNLAETR    -------SLDIIRGLILIVKSVYMGAHFPC-----------CKSSSALLKLA[14]S  1336
Query_STARF  1465  hD-LTKLLQSMFGRDEGLT    TDHG--KLIHLVKGYEQLHQANLLSAA-----TT--GHGSGSSFAVKLFITV[ 1]D  1527
Query_STRON  1458  eD-LFSILPMLFSNESDLG    LDHV--QLQHLVNGYKQLQKAKLLTRQ-----TT--NQ------ALSILDFV[ 1]R  1514
Query_SEACU  1443  eD-LLAMLPSVFENVRSGQ    YHAN--SLEHLISGYKQLQRAKFWPLR-----SA-----DGSEIVEE----V[ 1]K  1498
Query_FUNGI  1375  yD-ITAWCSMKEKHFVEM-[10]DEATg-LITQVLPRAKGQTLKRISHDLLSQ-v------FDSSQVTNPLEQN-    -  1444
Query_DPURP  1497  yQ-D-FFKLSLYEIILDKK[12]LEGTd-KLIQLIHSYQILYKFKFLDSILLE-nNKtiINNSSSNNNNTLPE--[10]-  1583
Query_HALBU  1498  yR-DTLIGYLGEMISSSS-[11]TEGTd-RLVSLINGYNILQKQSILD-VTLK-----fAKLSLPAMTSTL----[ 1]-  1567
Query_GLOBO  1318  -----------EYLLSSDF[10]R-QTt-LYLQVTNGLQSLRKCGILEAVLM---------------AKSVSSR-[ 7]-  1376
Query_PALUS  1373  -----------DVYFNDQI[10]Q-RRg-TLLRTVAGLQQLCRADMFDAVMD---------------TRAMDTS-[ 4]-  1428
Query_CALIF  1447  -----------KKFIDTYQ[10]F----------LNSIYILNASGILQDIIKY-n--------------NISEY-[ 7]-  1497
Query_GLOMU  1480  -----------EIAFSDEV[10]--DPt-YYDQITHGYLVLSSSGLFPDLLEH-eM-------TSYMNTSSKLY-[ 7]-  1544
Query_LOBOS  1417  -----------EAISSGDL[10]GQDPi-RFKAVIDGLELVREC----DVLE-----------SMFDGSSHKAC-[ 7]-  1476
Query_RACEM  1397  -----------AIIRTDDL[10]A--Li-EMLQTLQGFKVLQSVKLFDMVCRQ-nK-------TAGHASSERA--[ 6]-  1459
Query_PVIOL  1485  yK-DTIYQVVREYILLDQ-[11]QQGSd-RIIQLIHCYKNLYTNGMLDTLLLE-----iSNSTAQTAQSTLISKC[11]-  1569
Query_BOMBU  1314  ---IYDFVVL---DNIIYG    -KKNftELKQYVRGLSFLDRR------NIL--NQ-lDKMKELINRQDTVTRI[ 1]K  1370
Query_AEDES  1415  ---MEDFSDR--IEDLMSS    -ETIglEDGNCAKGI-LMVAQRYTTFKRYF--SSgtIDFLETISNK-LLHNI[ 1]N  1477
Query_ANT__  1333  ---VIKLLNL---CDIVEN    -D-ScdDLIQHVNGLSLLKRC------DIL--DQtiLETNGFANAEHMVDQI[ 1]R  1389
Query_WHITE  1361  ---YENLINKPLMYQLLHQ    -----aSNNEVAEGLVLLQQIQSCRSKNSF--LMkvLKDKSFSTVN-LYQKI[ 1]N  1422
Query_BEETL  1330  ---FNDEFDL--------K    -DHVklKQRHIVRGIYILQTL--------------sLKQFTVGRVA-PLLNK[ 1]L  1375
Query_THRIP  1330  ---APDMLNR--FHEAIRR    -DRVlpHDKELVRGLSVLQ---S------V--FKetQMWSGLIGER-SLVAV[ 1]Q  1384
Query_TERMI  1391  ---QTDVLSN--LQDSLRN    -DKVtlEQRQLLKGLEILH---KCGMLSHI--VKelPNQYGLDGNQ-IVKHL[ 1]S  1451
Query_FLEA_  1329  ---IHEYFNE--TEVNFNS    -VSLsvAKNNCLNAISSICKHSYVTRSKFW--SSla-------NGQ-YITNI[ 1]N  1385
Query_MIDGE  1378  ---FNEVWQN--CDKSLNE    -KTIslLESNKLRGIDLMLIQLKHVLS--F--DAlaQQLLVKTAEV-LLKKL[ 1]Q  1439
Query_LAODE  1365  nE--HGQLAANLSDLIRRQ    SISVe---LNSLNTLIVLKQINLLDADDQG-----lIKDETLWNSENLIQAI[ 4]-  1429
Query_LOTTI  1401  -D-LFSLLTNLS--LNNPV    VGHT--VLQHLVSGYILLNKTGLLVN-----tNV--QQ-LVFDTVYKVCIYP[14]S  1472
Query_APLYS  1450  -N-LTSSLP-LD--LSSAS    LDYT--QLSQMVGGYRQLHAAKLLMDQLPAaaTA--GTTSKADLASLLLDNV[ 1]E  1513
Query_SCALL  1453  -D-LSGYLP-LT--LSDLK    LDSI--YLQQLLGGYHQLHTAGLMTS-----vHL--GSGKASDLAQKLFDNV[ 1]Q  1511
Query_CRASS  1426  -D-LSGLLS-RP--LDDPS    TDHV--FLQQLVSGYQLLHSVGLLSK-----vSM--AT-TLPDLANKLLEYV[ 1]S  1483
Query_OCTOP  1399  rD-LLTVLPIHLKS---PN    TNYF--KLQQLLTGYRILNKVKLLGDM------------NLESLAAKLFHNV[ 1]Q  1453
Query_POMAC  1433  -D-LFGQIP-IS--LTSSS    IDYT--NLQQTVEGYCQLHDVRILLE-----vLG--KHGQPEKLAELLFDSV[ 1]K  1491
Query_WHIPW  1389  pD-VNNVFSADSCD-----    ---Ms-EAQRTTRGYALLKRMGYLPYEIIL---N----------CDNVWEVL[12]R  1449
Query_TRICH  1403  pD-VESVVTCDSHL-----    ---Ai-EQQQLLNVYGMLVEVGLQSVVELP---R----------AEFLWTTV[ 6]K  1457
Query_AASTA  1318  --------TYDRTLLFNME    SDLAq----RLCMTYKQLDQIKYWDGESSS---R--RRDIACDFA-RLAIDK[ 1]-  1371
Query_AEUTE  1287  --------PYRRENLFDMD    SNLVq----WICATYSELHKINFW---SHD---Q--QEDIANDFA-NMAIAE[ 1]-  1337
Query_PLASM  1384  ---------VSSSESLSDH    SDT-i----TSLTLFGS--KVLSAELVKSS---A--ADNYAEIFA-LVARKI[ 1]-  1433
Query_BREMI  1370  ---------VDSSETFDKQ    SET-i----SSLSLFGS--KVLCAGIAGFP---E--ADKHAEMFA-LIASKV[ 1]-  1419
Query_OLIGA  1398  --------SAESSESPTVR    IDRQl----LSLRLVNE--LVVC-DVVELS---K--LTPHVDNFV-SLATRA[ 1]-  1448
Query_LIGNO  1496  vT-VQRLLPQSGLASDPPS    VAGAv----RLARGVRALRDCGLLADSTAD--EK-----------RQILDRC[ 4]-  1552
Query_CESTO  1311  iN----------------E    -DNTt-----------VLFNISLLIDSVKE--DRglQNLFSHYIDKKMCQDI[ 1]-  1352
Query_OVI2_  1580  fS---NTFPHDSETYSTQR[ 5]SDQSvpVLSRLLQYVTCLLHSVPQNLSWRQ-------QFNSAEFGASLLGLV[12]E  1658
Query_PRIAP  1443  ----YNILPRLPCLMTDTG    DDTR--YLSSLVCGYEQLYKAELLLPA-----LQ--EA--APALPASLLTSV[ 1]E  1500
Query_ROTIF  1447  lN-LDFLVLKDI-------    -DWI--NLSQTIAGYEILSDFKIFKITTDM---------------SKFVFDL    -  1491
Query_SPONG  1431  nN-VINRLPVSYNE-STTS    SGIL--LFNRLVQGHIYLAESNKCLSDVLK--SM--GISEEESLAMRLYTNA[ 4]R  1498
Query_TARDI  1440  -----EYDMVELLR-NSGK    LENVh----LLVAGFTVLHEASQTELVKDYlqTK---KLTSEQFATKLIQHV[ 8]A  1506
Query_TRICH  1419  yN-LPDLLPIRFDRDNELN    IDFE--RLTFLIKGYKQLLTMGILKTTPA---DH-------KDLIEKIMKSV[ 1]D  1478
Query_OSTRE  1380  aN-LQQMVSIPKFNLSEVD[ 2]STQGciAAQRLVDGYTGLASVSMLRRVLMR---------EGNTSA-KVLAEV[ 6]-  1447
Query_NITEN  1497  fD-LGAVSLDEVLGVSGRA[ 1]GQGAa-ELLSLVRGYQQLFHAGALTSILPP-kR-------LGELAEKLATA-[ 2]-  1559
Query_CLUBM  1347  fD-ISTLELHQI--TEGR-    ---Ls-QSMNLVKAYQHLHTVGALVPTLGK-eR-------SEQLSERLLHT-[ 2]-  1402
Query_MOSS_  1428  yD-LNNLSFSSK--ISGKP[ 1]ALSLe-DALSLVKGYQQLSQVKILPLVLNK-vR-------AGKLGERLILT-[ 2]-  1488
Query_LIVER  1450  fD-LSGIDLSLD--GSGNS[ 1]SLGVe-DAMSLLKGYQQLHLSGSLLSVLGK-kE-------GKQMGEKLLAR-[ 2]-  1510
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1489   GMvtAV-D-GKPTLVNLSP--VSQSLAEKMLTLAF   QMGYPEKN-LVELIF-nSNNIHGV   STS------VQC     1547
Query_LINGU  1501   ----SE-E-DQFRLLECSP--STLELSRKLLELAC   SMTDKPAQ-ILGAIL-sK------   TKT------KNQ     1549
Query_LIMUL  1478   GCvgNV-DGSTFVA-TLNP--PTVALMEKLLDITF   LLNTQPDY-LIEKLL-vTDNVLTT   KRDTYSRL----     1538
Query_SPIDE  1481   SSicCV-SINKNIA--LNP--ASLNLAEKILKVAY   KMDLQVQN-LAKALL-eTTEEEHS   KLVV--TK----     1538
Query_SCORP  1480   SCilET-EIGCCIP-ILNP--SVLNFGKKMLTVAL   KLDISSED-LLKKLL-dEQQLYVS   NLKTKTRR----     1540
Query_TICK_  1438   GA-------------DWTP--LGKELCLRVLELAA   ALGAQARD-VLEALL---------   ------------     1471
Query_MONOS  1413   -LqtQG---------RLHP--SSKETGLVLLELAF[4]-----CTDaLLACLLdeETVLLAT[5]DAPASATQgIQQ     1475
Query_ROSET  1377   SHtsTP-----------EP--TNLAMAGLLADLAL   SLSVPKSQ-LLAMLRddTPTPQAG[4]TMLASASKsATQ     1437
Query_BELCH  1520   SMvtRK-D-GHPTAASLPP--SFHQLTGGLLQLAC   QLHPQGHD-LLDCVM-nKTPLMTG   GGK------KSS     1578
Query_CIONA  1486   -VklDD-SEEALTQNEVTP--HNIQLGNKILQLSL   HMAEAAKKkLVNCIF-sNKQVATL   NFDVESHR----     1547
Query_CAECIL 1503   GI--AP-GDERKSLPSLDI--SSKRLADGLLQLAF   DLGGQCDE-LVSLLL--NTVKLSV   PLSGTSQRnIIS     1565
Query_HUMAN  1490   GI--AP-GDERQCLPSLDL--SCKQLASGLLELAF   AFGGLCER-LVSLLL--NPAVLST   ASLGSSQGsVIH     1552
Query_MOUSE  1488   GI--VP-AEERQCLQSLDP--SCKSLANGLLELAF   GFGGLCDH-LVSLLL--NSAMLST   QYLGSSQRn-IS     1549
Query_XENOPU 1512   GI--AP-TNERKSLPSLDI--SSKRLAEGLLELAF   MFGGQCEE-LVSLLL--NTVILSV   PLPGTSQRnIIN     1574
Query_PLATY  1494   GI--AP-GDERQSLPSLEI--SSKQLASGLLELAF   AFGGLCEE-LVSLLL--NTVVLSV   PLAGTSQRnLIS     1556
Query_TASDE  1497   GI--AP-GDERTALPSLEI--SSKRLASGLLELAF   AFGGLCEE-LVNLLL--NTVVLSV   PLSGTCQKnLIN     1559
Query_CANAR  1492   GI--AP-GGERMSLPCLDL--SRKRLADSLLQLAF   AIDEQCEE-LVSLLL--NTVVLSV   PLSEASQNqLMN     1554
Query_OSTRIC 1462   GI--AP-GAERLCLPSLDI--SKKRLADGLLQLAF   AIDGQCEE-LVSLLL--NTVVLTV   PLLGASRKnLVN     1524
Query_TURTLE 1501   SI--AP-GDERKSLPSLDI--SSKRLADGLLQLAF   AFGGQCEE-LVSLLL--NTMMLSV   PLSGTSQRsLIS     1563
Query_GECKO  1501   GI--AP-GDERKSLPSLDM--SSKRLAEGLLQLAF   AFGVQCKE-LVSLLL--NSLMLSV   PLSGVSQRnLIS     1563
Query_DANIO  1500   SI--AP-GTDRKSLPSMDV--GSRKLADRLVQLSF   CLGDQSEQ-TVGLLL--NTITLSV   PLSGSLNPhFLS     1562
Query_SNAKE  1501   GI--AP-GDEGRSLPSLDI--GSKRLAEGLLHLAF   AFGVRCTE-LVSLLL--NTLMLFV   PLSGAAQEsLIN     1563
Query_STERL  1513   GI--AP-GDERKSLPSMDV--SSRKLADGLLQLAF   SLGGQCKE-LVDLLL--NSIMLSV   PLSDTSQKnFIT     1575
Query_MILII  1497   GI--AP-GDERKSLPSIDV--SSKQLADGLLQLSF   SLGGQCEE-LVSLLL--NTMELSV   PLSGATQKnFIT     1559
Query_BAMBO  1496   GI--AP-GDERKSLPTMDV--SSKMLADGLLQLSF   SLGGQCEE-LVSLLL--NTVELSV   PLSGASQRhFIS     1558
Query_ALLIG  1459   GI--AP-GDERKCLPSLDI--SSKRLADGLLQLAF   AFGCQSEE-LVSLML--NTAMLSV   PLSEAAPRnLLS     1521
Query_GAR__  1500   GI--AP-GDEKKSLPSMDV--SSRRLADGLLQLAF   SLGGQSEQ-LVSLLL--NTIMLSV   PLSGTSDKnFIS     1562
Query_TETRA  1480   GVksSE----ENILNKQDSl-FRESRTQQIFQFAL   DISSESTNeFILQII---------   ------------     1524
Query_PARAM  1425   -VinFE--QGNDQINNIEY-------VTHLFEFLL[3]----RAKNfVISNLI-------KA   -----------S[1]  1467
Query_SLEMN  1468   AWfdID-----------TPisNES--ARVLVKFLS   --KLEIPN---ETLC---EWLTKK   -----------H     1506
Query_STENT  1262   ---sLI-----------SKdsHEARKGKNQLEFFL   --KTKIEE---KQLR---TVLANS   -----------K     1299
Query_NEMAT  1485   GAcrAA-VRTSGSIIVLSP--ASHQLACKLLELAF   DLGIEVIT-IFRRAV--NILLIQM   ICFLFSD----G     1545
Query_PISTI  1521   GV-kPS-LNVGLQTVILDP--ASQHLASKLLDLAY   DLGLQFN------------YLMES   IMEDSNDNp--G     1573
Query_CORAL  1498   GFivQD-SDV-FRIVNLTP--ASRELAKKLLDLAF   DLGLKIES-V----L--HCILNKS   CLADAKGS---S     1554
Query_HYDRA  1437   ---lSN-KEKV---SFLLP--AQMLLANSLFRLYL   EINPSAA---LTDLF-vNEKFS--   ----QNTK----     1484
Query_LATIM  1501   GI--AP-GDERKSLPSIDI--SSKRLADGLLQLAF   SFGLQSGE-LVCLLL--NRVTLSV   PLSGTSQRnFIT     1563
Query_SHRIM  1469   GSifKE-QNLSVSGISPHP--SLLKFLYTVMQLAI   TLDPKVKHhLLSWAT--NSDTVSG   TSGQEVEK----     1530
Query_MCRAB  1468   NSlfMY-GSAS----SPHP--TQLPFLHVVLDFAV   TLDSQVLElLVKQVT--NMEDVKG   VTGQQIEK----     1525
Query_AMPHI  1491   DFliTN-PSRSVAA-------AFAPISHQITSFAS   TA------lHVALQL--REETVSE   VLELVIAA----     1541
Query_BARNA  1389   SCgrQP---------PTDPm--TVRLADTALETAL   ALGVPARR-LLSALLseEQVRAGS[1]-----------S     1437
Query_COPEP  1419   -----P----EKTLNKISPd--ILKNLEVFIQIIL   -VDESDSEhIAARLM---EIISSD[1]------------     1463
Query_DAPHN  1337   GStgDS----------VSSa-VRLRASSLLVELAL   ---------LLGLSM---EVLSNE[1]-------RnPAN[5]  1383
Query_STARF  1528   SLttSD-EAGKLLQANLTP--LSVELAESLLDLAF   LLGIEAKM-LVSCLL-dKTALFSG   R---------SS     1584
Query_STRON  1515   GLvkRG-ENGQEAR-KMTP--SEIDLGQRLLELAF   QLGVQTNS-LIDCLL-eTTPVVSS   TRSLTSKLsASQ     1579
Query_SEACU  1499   ALrtSE-D-GRDVAVSLRP--AEIELGTKLLELAF   QMGVKVPI-LVDLLTkeDRKVWNS   T---------TE     1555
Query_FUNGI  1445   --------------------------YGSLVLQLC   NRFDVDTSdVVDKLFdaT-IIQGG[6]GTMELGSVnVEA[7]  1501
Query_DPURP  1584   -------NYLFNHCENLSP--SKLLISKELIQLVF   TIG-IEPQkLLSMIIne---S-NT[6]DANNQGDG-EDS[4]  1649
Query_HALBU  1568   -------DYLNEHVDNITP--AHLIVYKEILVFIL   STGHIQPSkLLEYLLppIKVLTKS[6]TSSTITDKqLEE[8]  1643
Query_GLOBO  1377   -------RSGITISKSSDL--LQIQLARNMVWECI   CNN-ESQEdAFSLILn--------   -PPVSGTD----[1]  1425
Query_PALUS  1429   -------KLVCSLREETDP--STILVVGGLLQVCL   MHR-SSRErYFSEILc--------   ---LENEK----[1]  1475
Query_CALIF  1498   -------NCLEKLKDNENP--SYDLIIKQLLDICI   INS-EIQEeCWGKLLgl-------   --STMTDL----[1]  1546
Query_GLOMU  1545   -------KKFNNLRDVSDP--LWVELIRKLLSFAL   SLS-KYNGaINQWLLsdI----IG   FPSDNSTF----[1]  1598
Query_LOBOS  1477   -------EVFMSMQSTRDP--LWINYCSALLQFSL   GER-QCRQrLWDYLL---------   ----DSEH----[1]  1521
Query_RACEM  1460   -------DKFLEMRQSSDP--LRIKLVGDMVRVAG   NSL-CSESpYCRKLLsnD----LS   DP----------     1506
Query_PVIOL  1570   -------EYLFEHSTKLSP--SRLLVFKQLLELVF   TIG-ISSDkLLSYIIdp---TVKD[6]DNNNNSDK-MDT[8]  1640
Query_BOMBU  1371   ALvrE--SLGELIYVTIKP--SAKDYLEILMEF--   SLIHYEFS-MTRMLI---EFIEND   TMLGSGD---KK[1]  1429
Query_AEDES  1478   GLkrE--RLDELFVVDLTP--SVKRFFSCLLHI--   VLLFKSTG-TMELLV---EFMLNE   QKLRLYQQsVNS[1]  1539
Query_ANT__  1390   FLvsE--RIGELVCRDLSM--RMIEYLRALMEFQF[5]TIIGNILS-LTKKLE---ELIIND   TPVTSVD---CT[1]  1455
Query_WHITE  1423   SLydS-----NGTTVHLSE--KDTDFLRSLFSL--   ATL--LDE-EVELLS---GFIFKT   DQARLKD--SLR[1]  1477
Query_BEETL  1376   QFaeD----KIIFLDDLQD--TTLSYCITVLNY--   SLE--DQN-EFKVFV---DYLYKD   YKVQTTN--CKQ[1]  1431
Query_THRIP  1385   AIvePVlNSEHKAPVSLNP--TARMYAQSMLEF--   YFS--TCP-ELNQLI---NNILDR   SPLSQRD--LQC[1]  1444
Query_TERMI  1452   ALveG--QGDMKRTVVLQP--AAFSYAESLLKF--   AFL--LNK-EVTTVI---CCIKEQ   ASIQNPA--THS[1]  1509
Query_FLEA_  1386   GLieE--NTDTDTKIGIFPehNVINYYGKILSV--   IF--EHVD-CKDLLI---ESVLND   VHIIGNI--TES[1]  1445
Query_MIDGE  1440   CVveE--KLNVKRPKLLTP--SAKAYASNIIQL--   CL--KIND-FLSTTV---TFSFND   MQLMISQ---SS[1]  1496
Query_LAODE  1430   -VgqSD--------KAIKPnkSQTLYLELLLDFAF[3]--NKSNITkLIECMCdeTQLKYED   ---DDQSKnVTK     1489
Query_LOTTI  1473   GIyeEK-G-KLKKIITLTP--AASILGKSLLDLSF   KLGLKDNV-LITALL-nDDMIQGK   TPS------VNL     1531
Query_APLYS  1514   GLvqTE-D-RTRTAATLSP--TAQTLANALLELSL   QLSFDVDK-LVAKIL-dKSSVQGA   RKN------IEI     1572
Query_SCALL  1512   GVvcRN-G-GGLEAITLTP--VVHTLMKDMLALSF   HLGLQGSV-LVEAIL-dQTLVSGS   SKG------LEQ     1570
Query_CRASS  1484   SLicVE-G-GTKKSQTPSP--TAQVTATRMLELSF   CLGVQIKK-VVMCIL-dESMVEGT   AKG------YSQ     1542
Query_OCTOP  1454   NIvtHS-DDNQLTSRTLTP--TELSLTQTLLELAI   QLSMPINE-FM-AVM-fTEEAHQQ   NTSC--------     1510
Query_POMAC  1492   GIteKK-D-GKDKTVSLSP--AAHRFASQLLKLAF   QLGLPAEK-LVEGIL-dIRIIQGL   SLG------IEY     1550
Query_WHIPW  1450   ST-------------SLLS--TKVGLAKRILDIVP   QSAAQIKS-LIHILAdcQPVRYLN   FSSIATKR----     1500
Query_TRICH  1458   RI-------------GLPP--SKLGMVKQIFRVAA   LTGSAVDV-LIKLLMntESVRLAE   CTSLCSTK----     1508
Query_AASTA  1372   ----AS--------GLLTP--MQDQLTHRAMETAM   ----DMGWdVATFIL---------   ------QPtPAN     1409
Query_AEUTE  1338   ----KN--------NTRSP--IRDQIARTAMQTAI   ----EMGWdLASFIL---------   ------KStDEN     1375
Query_PLASM  1434   ----EY--------GHGSP--QDRLVVVAALQAAA   ----ASEWnITDIFL---------   ------KSaDRR     1471
Query_BREMI  1420   ----DH--------GNGAP--HDRLVIAAALKGAA   ----ACGWrITDIFL---------   ------KPyDRQ     1457
Query_OLIGA  1449   ----GG--------ASVDP--NDRILALEAIKAGC   ----FGGWdMTRVFD---------   ------DRdLRR     1486
Query_LIGNO  1553   ---------ESAQTTQQHP--VLPQLIDAALTYAL   NDSALLQP-LVTLLT-eHSAGLDL   LNS---YKpCLT     1607
Query_CESTO  1353   NMlkKE--------LTIPPd-----AMSKLIHSVF   ---------TISTLF---QYQVSF   ------DKeLIF[3]  1395
Query_OVI2_  1659   AVhqVLlDVIGLALASLDK----TTIIEYLLDQLL[4]--NGSFQQhKLSSGLsdKTVCQLL   -LWGRTEMpTRK[8]  1734
Query_PRIAP  1501   GCtlSP-RDGVSSLTGMSP--SVRELGGRMLRLAF   TMKVESDS-LLSLLF-dDTRLLGV   SAK--------Q     1558
Query_ROTIF  1492   -------SQEETESQTTTE--AKRKLFSLCLSIN-   ------RN-FLTKIL---------   -------QnYLE     1529
Query_SPONG  1499   GIssSP-GDGTPGY-YMPP--LHKLTGELLLKLGL   ELSDSVEP-FLLLVL-------SP   AQSGAGAS--FA     1555
Query_TARDI  1507   NS-----GSRKEAWISTDV--SQEK---ALFSLAL   QCQSSGAS-AAAFCI---------   -SSHSTQVsIYS     1556
Query_TRICH  1479   CF--RG-REMNGDRTALSP--AQQKVAAELLDLAF   ELNETIDI-FFR------YLLDET   KLYGDSNSa--T     1535
Query_OSTRE  1448   GLghHA-----------NPp--QRVVGQAIVKLCL[4]-----APNvLLDFLL---------   -----DRRgPNS[4]  1494
Query_NITEN  1560   -------NLKRE-C---PP--PLRPVATAMLELAT   ALG-LSVGrLLDLVSdeSEVVSGG[8]TPPGMVRR----     1620
Query_CLUBM  1403   -------TISKEPV---SP--LAEPLYVEIIRLSV   ALR-VVVDqLLRIVLdk-------[3]TVDRGTDS----     1452
Query_MOSS_  1489   -------GIEQE-C---PP--SLQPITAEMVNLAM   SLG-VNATfLVNLIMde-------[3]GISTN-RK----     1536
Query_LIVER  1511   -------QFGRE-F---PP--ILETIASEMVSLSL   ALG-TSVQrLVELCLde-------[3]SLPLQTRK----     1559
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1548  SPHGTLFYAAFKSVVNLNFVKSV-EKFVPLLMR   HL-EANG---ETLSSVLIG---IAESVIS--DRPLRKSH---GK  1611
Query_LINGU  1550  N-SSQNFYAMFQSTINLYLMKEA-KYFIPELVN   YA-SKES---SVVSSVLSS---LLEQISR--DRQLRGGE---GK  1612
Query_LIMUL  1539  ---GLVVFGLFQAKISRYLAKTA-NTSLSLLVE   NA-KKKP---KVVPSIILG---VLDFISS--NKNIRKEV---GT  1599
Query_SPIDE  1539  ---GLLVFNLFTNSISLFAVESA-ENFVKTIAS   NL-NHQL---FNEISILLK---FIDCVAC--SKILRKKH---GK  1599
Query_SCORP  1541  ---GLIMFSTYQQEISLFAGINS-KHFISSLMN   NI-NKSK---TTVLKILVG---ILDSIGK--VKEFRYRY---GK  1601
Query_TICK_  1472  ---PSQAWEVLGAQLCRHFAERA-SDYVPAALA   LA-EGAP---KMLTTVMSH---ML----Q--DKELTSRR---GS  1528
Query_MONOS  1476  -SKGLRFYETFRTCIAGFVARHT-----TVLMP   ----DLA---------LTA----AADLAR-----YRREA-----  1519
Query_ROSET  1438  TTRGALFYDTYRTVLAPYIARHF-KVFAEQLLP[8]RLvHDAAr------ALVTG-----------DTEMRRYAd----  1499
Query_BELCH  1579  TFQGMVFFSTFKSCICQQLIKTC-AESLPVLLQ   QL-TAEP---VWVGSILHG---LLDHLAK--DKTLRKKE---GA  1642
Query_CIONA  1548  -SKGILFLSTYSLAIYKELCQHP-ALYIPLITN   HTcKADA---STCNFILTG---VLDYLAT--NRHARKQH---AA  1611
Query_CAECI  1566  FSHGEYFYSLFSETINTELLKNL-NTAILQLMK   SS-VDNP---RMVGTILNG---MLDQCFR--DRTVRKQQ---GV  1629
Query_HUMAN  1553  FSHGEYFYSLFSETINTELLKNL-DLAVLELMQ   SS-VDNT---KMVSAVLNG---MLDQSFR--ERANQKHQ---GL  1616
Query_MOUSE  1550  FSHGEYFYSLFSEVINSELLKNL-DIAVSRLME   SS-SDNP---KMVSTVLNG---MLDTSFR--DRAVQKHQ---GL  1613
Query_XENOP  1575  FSHGGYFYTLFAETINTELLNNL-DTIVVELMK   SS-LEDP---KMVSCVLNG---MLDQSFR--QRTIRKQQ---GV  1638
Query_PLATY  1557  FSHGEYFYNLFSATINTELLKNI-DTAVLELMK   SS-VDNA---KMVGTILNG---MLDQSFR--DRIVRKQQ---GE  1620
Query_TASDE  1560  FSHGEYFYSLFSEVINTELLKNL-DTAILELMK   SS-VDNA---KMVGAILNG---MLDQSFR--DRAVRKQQ---GM  1623
Query_CANAR  1555  FSHGQYFYSLFSETINQQLLKNL-DVFIVRLMK   SS-ISNP---QMVGSILNG---MLDQSFR--DRAVRKQQ---GA  1618
Query_OSTRI  1525  FSHGEYFYSLFSETINEQLLKNL-DVIILQLME   SS-VSNP---QMVGSILNG---MLDQSFR--ERTIRKQQ---GV  1588
Query_TURTL  1564  FSHGEYFYSLFSETINQELLKDL-DVAILQLMK   SS-VDSP---KMVGTILNG---MLDQSFR--ERAVRKQQ---GI  1627
Query_GECKO  1564  FSHGEYFYSLFSDTINRELLKNL-DVAVLHLLK   SS-LDSP---KMVGTILNG---MLDQSFR--DRAVHKQQ---GV  1627
Query_DANIO  1563  FSHGEYFYSLFQTSLNTELLRSV-DRSVPLLLS   SA-NQNP---SMVSVLLNG---MLDHSFR--ERSVRKSQ---GS  1626
Query_SNAKE  1564  FSHGEYFYSLFLEIINRELLKNL-EIAVLELMK   SA-LDSP---KMVSTVLNG---MLDQSFK--DRATCKQQ---GV  1627
Query_STERL  1576  FSHGEYFYSLFQETINSELLKNL-DTTVAQLMK   SA-AENP---KMVSMVLNG---MLDQSFR--ERSVRKTQ---GS  1639
Query_MILII  1560  FSHGEYFYSLFSETINAELLKSL-KSAVTQLMR   SA-SENP---KMVSTVLNG---MLDQSFR--DHTVRKTQ---GV  1623
Query_BAMBO  1559  FSHGEYFYNLFPETINSELLRNL-NSAVAQLMT   SA-RENP---KMVSAVLNG---MLDHSFR--DHTVRKKQ---GI  1622
Query_ALLIG  1522  FCHGEYFYNLFSETINHELLKNL-DVAILQLMK   SA-VDNP---RMVGTILNG---MLDQSFR--ERTVRKQQ---GM  1585
Query_GAR__  1563  FSHGEYFYSLFQQTINIELLGSL-DSTVPQLMS   LA-AENP---KMVSAVLNG---MLDQSFR--DRSVRESQ---GR  1626
Query_TETRA  1525  -SQQSQFFQKYENQLLTYVKSKW------NLIG[4]MLaKDNI-------SLLKLtipLLQSIIQsfNRKEKEESll---  1588
Query_PARAM  1468  ------FRQKYEEQLIAFVARNF--ESIGQIYF   -----------------------QYICK--ETHLFRQGil---  1507
Query_SLEMN  1507  -----LLYQKSSDLLLSYIAENW-NHFKLNVGE[3]IKaKDEP-------SLYAVfipILEKLSQ---KTPNYYKi----  1566
Query_STENT  1300  -----L-FQSNSEILYEYMSRN----MESGFGV[3]SMaGKPE-------KMFEFlvpFLEK-AK---RLIEAKEv----  1354
Query_NEMAT  1546  SSVGAVFYTTFQASINTYLIKNA-QSHIPQLVS   MA-TSYS---SHVCSILNS---ILDQFNR--DKTSRKRY---GS  1609
Query_PISTI  1574  SSCGCTFYRIFQTSINGFFIERA-NIHMANILM   KA-ASSP---FLVLSLLNG---ILDQVIR--DKILRTRN---GS  1637
Query_CORAL  1555  STVGMVFYQTFHNTIEHLLIKNS-SEHLPKILE   YN-RDEA---QWVSSTVIA---LLDSMVK--DKELRKMN---GR  1618
Query_HYDRA  1485  -SRGLVLYDLLRSVIINYYIDYP-LSLQKGFDN   ------S---ETLSYVHLI---ILQDILA--EYVSQKKS---SV  1542
Query_LATIM  1564  FFHGEYFYNLFMETTNSELLRNL-STAVSQLMN   SA-AENP---KMVGAILNG---LLDQSFR--NRTVRKQQ---GS  1627
Query_SHRIM  1531  ---GSHVLSILGPKLIPQFLESP-DDALSRCVE   MI-REGR--fLPAMGFASA---LLTCLIR--NADIRKKY---SV  1592
Query_MCRAB  1526  ---GFHILNVLGNKIIPHILPCC-DLIVTTLLE   LV-KEGK--fLEVMSFNGM---LLNTLIR--NSELRKKH---SV  1587
Query_AMPHI  1542  ---GDEPL-VHADAVMTVLVLRC-SYAVQRCVQ   LA-QAGG--wKKAVLLLLS---LARTLLK--DSALRKKY---SE  1602
Query_BARNA  1438  AERGWLFFSRFSAVLVPHLLQPA---AREAVLS[3]SRaAAHPrg----VQLLTA---VVRTTAA--DRQLRRQH---GA  1502
Query_COPEP  1464  -EKDVWFIRTFSMLCHKPILLNM-----QTLVS   KGvQSGD--------ILQSvqmLVNIILN------RKQPg-lGK  1519
Query_DAPHN  1384  ITYGQYLRSKLGHLVFPVVINQM-TSVLPFIPQ   EGnGDSQqw-----ALLLE---LLQAANQ-tQRFSDKRKw----  1446
Query_STARF  1585  TSKGAVFYTTFKTKLNGEMAVQA-QKVIPLLVR   EG-GKDV---TVVSSVLNG---LLDYLTS--SREVRRTH---GM  1648
Query_STRON  1580  TSLGVLFLTTYKQPIMDQLVKHG-QDTVYELSN   RV-REDR---GTIGMIIYG---LLEHAVG--NREIRKRY---GM  1643
Query_SEACU  1556  IPHCEVFYAAYKSALNDELACQS-QHSVRALLK   EG-G------TRVMGILNG---LLDQILS--NRNQRKRY---GM  1616
Query_FUNGI  1502  MTRGEEFYVKYHDALDKYTLEHR-AVMARVLAS   KIaQYSL-----AVRIVLSvshIASKLTH--PFLHELFR----H  1566
Query_DPURP  1650  DIN-DIFYKSLFTEINTYISSNF-ELFLPLLAN   RIiHTNH-----IHKVLNDvciLKQQP--------NEKS-----  1706
Query_HALBU  1644  IEHQSNFYKTYHSELIEYISKNF-ESFSPLFID   RIvQTNH-----IDTILNDiciYKQTIQF--NNTTDKQS----K  1708
Query_GLOBO  1426  AMVNLKNLAPTVSHAVC--------SYPSSFI-   KSmREGV-----EASFAIV---LIDNFLNwiETAESKSE----V  1481
Query_PALUS  1476  AETAIKFYQRYASYINSLLARNF-DQYVDFFE-   KHlQHPI-----VHTILLG---IPDYLI--vNRTSKADE----C  1536
Query_CALIF  1547  YSDVISFYSKYKFKINYLFITNF-EIFNKYLSF   -CiSENI-----VTDIIND---CIDLIND-yKSIINEKI----K  1608
Query_GLOMU  1599  YQNGTIYCQKFSSVINEHMSRNF-REFAKIFF-[8]DQmNIDI-----IKDLILN---LIDYLI---SKKDTQQS----V  1666
Query_LOBOS  1522  EQAKMIYL-KYADVINRQIALHF-VELSPVISC[9]AIwN-------------D---FLDYMY---SHPELSTE----K  1581
Query_RACEM  1507  -ERVTVYE-KFHPYINACMALHF-KAFAPILL-   EDiEGDF-----MPRVIIG---LVEYLKD--NRKTLAKE----I  1565
Query_PVIOL  1641  DSRYDIFYKYYIIEINTFVSNHF-ESFLTLLAN   IViQTNS-----IHKILNDiciLKQSL--------SSQK----L  1699
Query_BOMBU  1430  -KHGVYFFSTFKCEIFRYILKDI-EKSIEIFND   LLqRDPS-------TLLMImeqLFLFVQR--YKK-EFHD---CV  1491
Query_AEDES  1540  ILHGEHFVNTFDQVIFEYVIQRL-PESLNLILD   RLsANNF------IAIVKWmcnFLEFLYQ--RKR-AEVP---AL  1603
Query_ANT__  1456  IKYGEHFLNTFESVIFKYMLTHE-RMITVVCND   MWrDNSS-------FLLKWiedLLLFLRQ--HKR-ELQA---HV  1518
Query_WHITE  1478  VTQGDHFKYLFFRQIISRFSDSF-SKAVHHL--   -ItPDKI-----HADILLAimkEHECSPP--NINQRNSE---LV  1540
Query_BEETL  1432  ISFGTYLFTIFEDTIIRKVVTNF-DVFFTYCLE   --yKDKI---DKTVEFIFYilr---CIGL--NKHFKTQH---GI  1494
Query_THRIP  1445  ITHGEHFVACFEAVLVPRLLGVH-HQTVSALAE   HLtPEVS---PQHLTILSSilrWRCSVSS--TTRYHLMQ---TV  1512
Query_TERMI  1510  VLHGDYFFVTFRSAILEHMLACF-EHSVSCLIK   DIsANNV---QWVISTVTEilqYYAKIRK--SK--HQDS---TT  1575
Query_FLEA_  1446  LAHGQHFYSIFGNQIIKYLMRFP-IDAVKKTIT   NLdGASP---IKRHIIVDFliqIILYIYR--NEK-ENEF---LI  1512
Query_MIDGE  1497  IRHGEHFIETFKQPIFKLFTSNI-KETIQLLVN   EMkTTND---TNRLRIVAVlseLNDYIFK--HCT-DNKE---LL  1563
Query_LAODE  1490  ---GQHFKNTLFEKIFAHICQFP-AESVTELMK[2]---SENF---EFISKCLCK---LLDRCRN-cEDSVKSSI-----  1549
Query_LOTTI  1532  Q-HGEAFFLMFKSVICQYIAENG-RTIIPILTN   QA-KTHG---RKVNSILGA---TVDYIAR--DRTLRKKE---GP  1594
Query_APLYS  1573  G-HGDFFLTTFSSTLASHIMTSSaREIVRSLMT   HA-SKEP---RCVSRVLLA---VLDHALK--DRSLRKKE---GV  1636
Query_SCALL  1571  C-HGDLVFSMFKSVIAGHMAVES-DMCVKSLAT   RA-GADG---RKVSRILVA---MVDHVTR--DKALRKRI---GQ  1633
Query_CRASS  1543  S-HGELMFTAFKTTVSVHLAKAM-AQAVPVLCE   HA-QKQC---YRVMSILVT---MVDTVSK--DKNLRKQI---SA  1605
Query_OCTOP  1511  ---GKILLCIFKQEVSGYIIEHA-AELIPLIMA   KS-DSNS---QIVGHILIS---VLD-ILK--HKSFRWKH---AS  1570
Query_POMAC  1551  M-LGDHFFTLFKTPISSHVAASS-D-VVSLLMN   KA-TVDP---HRVCNVLVA---MLENVAQ--DRRIRKSN---GP  1612
Query_WHIPW  1501  ---NDAFVSLFETEIAYVLCRNA-TEFIEEGIR   AF-EDND---EEFIQLLCS---LLSCVVR--SADLRKCY---GS  1561
Query_TRICH  1509  ---GELLIANFSDEIALILCTNA-ELFLKHLKE   AS-RCQP---DVVGRALFH---TVSVVVH--SSDLRKKY---GR  1569
Query_AASTA  1410  -------YTLFSAAFDKVFCDRD--DVWLQVVR[3]SHvQTHP------S-FIHTlhhVLELTSS------NNKNl--SK  1465
Query_AEUTE  1376  -------YDLFASTFDSSL-HKS--SIWLNVAV[3]SNiENYE------SCALNTfqrVLEANPK------SSDG-----  1428
Query_PLASM  1472  -------MYASIYSDVIHMLPTL--AVWKRCAS[3]SLsLEND------F-VVNVlidILRQVAC------SKVYl--RK  1527
Query_BREMI  1458  -------KYAFIYSDMTHVIASL--AVWKRCAS[3]SLsLDND------F-VVNAfldILRQVTS------FKVYt--PR  1513
Query_OLIGA  1487  -------KLSSIYRDIILLLPHR--HIWSRVAK[3]TDdITTQ------R-LVG-lveLLTEVAS------LKVHp--AG  1541
Query_LIGNO  1608  AYFSANCQSLLTILAASPLLKAA---TAPATAA   ATiARLP----ALAAILES---LAASVKR--TALSKQGCaddTP  1672
Query_CESTO  1396  ------------SQINKCDIHPI---ALLLILG[1]KQeSDTPl--ESPNTVIATllkLFSIYSR-aSNKFRTNQlceGI  1455
Query_OVI2_  1735  VSKDSPYSSCFIDRVLNRVRLAI-DTFDPAVLN   -----RAv--TLLTTCLEL-clSLQGSAK-iNIAVETSPkqqGD  1801
Query_PRIAP  1559  TTKGSMFYSTFGHEINNHLVGMA-TVLIPKVLD   RI-NNRE---SVVMTILTD---LLDQVGK--DRETRKRY---GS  1622
Query_ROTIF  1530  EQSANVTLSYFRAEIFTQCLNFY-SDILDFLFS   KL-SSKF---SKILQFLIN---FVEFIAM--DKHLRKTN---GQ  1593
Query_SPONG  1556  AQDGLLLYITYSGLINNKLLLHI-KMAIDVLLK   NA-KTHP---QQVSVILNL---LLEYVQK--DFKINNNN---NN  1619
Query_TARDI  1557  KLRSVVFTYLFPDTALRSSDQRQ-RLFFYGLIG   GDaEEHAgtwRSRHAVLKD---ALEYVCR--NRPARKRY---ET  1624
Query_TRICH  1536  ATLGMTLYYTFNSNINAFLAKVS-NRCLPMILD   LA-ASKS---FMISGILSG---LLDFILR--DRGLRQSK---GI  1599
Query_OSTRE  1495  ---GEALYRNYRESFIRTCEFQFdDYATDLLIT[1]SAdVTSAg-gRVATDIILG---VVDRATStlQCGGRMSLngvGI  1565
Query_NITEN  1621  VTTGALFLSNFSDVIVRRCVFQP-RAALERIFS   QTpRNGT-----MRALLG--aiLSHYLAG--E---DKLS----R  1680
Query_CLUBM  1453  VTPGGRFYLSFRKVVVQQVKFFY-KKCLPVLMK   ESvRNET-----ARLILH--avLDEY-------QGETIT----K  1510
Query_MOSS_  1537  LTSGEVFYMNFRDTIVHQVKFFY-KDCLRLILE   AAaEKQA-----ARMLLI--alLDSYLKP--SEMGASIT----K  1599
Query_LIVER  1560  AMPGGIFYLNFRRLIVQQVRFFF-RDCLAFILK   AAvGNRT-----ARLLIN--avLDDEIKS--IEQGSSLS----R  1622
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1612  M   VSRAIV-SQWSSLSV-wWSES-TASLD--TKMVAITLLTKLMQID---STVIIDVKNKSFY-Y   VFKMFLG     1673
Query_LINGU  1613  T   VVDVVL-SCWSQLQLqqWWAP-DSPQE--VKMSALNLLRKMLLIN---GKVVSDPTHPSFS-C   VFSMYQS     1675
Query_LIMUL  1600  A   VVESLF-SCWGKLNH--WWSNAADEDD--LNNTVVLLLTKLLIID---PAVVTSCSLPYDQ-S   LCSVFIS     1661
Query_SPIDE  1600  Q   VIDAFL-NNWNSVCV--WWRDHPDIDS--IYS-VLSLMTKMLLID---SKVALNESNQFSM-S   ILPTYLT     1660
Query_SCORP  1602  D   IINSIM-CQWSRLQQ--CLEESCDID---LQNDIITLITKMFLID---SSFIKG-NVNYDSnS   IFFSFIT     1662
Query_TICK_  1529  E   LSLEVM-RCWPSMER--FWSE-GTSQD--VVNGALHMLEKLVLLA---PEVG-SPKGPHFE-R   LFGVFLK     1588
Query_MONOS  1520  -   -LAAFM-RAFPTLHE--------ALSDq--PRLLVRLVEQASRCGlplAEVVAGLHPTAFRhV[1]------A[1]  1573
Query_ROSET  1500  -[1]FFTAFL-DGFLHVRT--WINSAGAGSEt--HVQGMRLVSAVVRCGvr---PETAQRHANFE--[1]VLEVYKR     1561
Query_BELCH  1643  A   LCDSLL-QSWPKLAP--WWEE-GALQD--SKSTALTLLRKILIID---SQVSRDPAHPAFQ-P   VFTMYLD     1703
Query_CIONA  1612  S[4]ILCQFD-K---NLKP---WINTDEDTD--NLNNLLTLLRKIIMLD---SSVCTKPNHPCHS-T   VFTIYTS     1673
Query_CAECI  1630  K   MVIAVL-QNWKKLDS--WWDK-NSAPE--SKMAVLTLLAKVLQID---SSVSSNAHHEAFS-A   VFNTYTS     1690
Query_HUMAN  1617  K   LATTIL-QHWKKCDS--WWAK-DSPLE--TKMAVLALLAKILQID---SSVSFNTSHGSFP-E   VFTTYIS     1677
Query_MOUSE  1614  K   LATAIL-QNWRKCDS--WWAP-DSAPE--SKTTVLSLLAKMLQID---SALSFDTNHSSFS-E   IFTTYAS     1674
Query_XENOP  1639  K   LVNAVL-ENWRRLDS--WWYK-DSPSE--SKMAVLTLLAKVLQID---SSVCFDINHSAFA-E   VFKTYTS     1699
Query_PLATY  1621  K   LVDAIL-QNWKRLDT--WWAE-DSTPE--SKIAVLTLLAKVLQID---SSVSSNTNHAAFS-A   VFNAYTS     1681
Query_TASDE  1624  K   LVATIL-QNWKKLDT--WWAK-DSVPE--SKMAVLTLMAKILQID---SSVSFSINHGAFL-D   VFTTYTS     1684
Query_CANAR  1619  K   LVTAVL-RNWEKLDG--WWAK-GSSAE--SKMALLTLLAKVLQID---SSVSFNTNHEAFT-A   VFNTYTS     1679
Query_OSTRI  1589  K   LVTAVL-RNWKRLDS--WWAK-GSAPE--SKMAVLTLLAKVLQID---SAVSFNTNHEAFT-A   VFSTYTS     1649
Query_TURTL  1628  K   LVTAVL-RNWMRLDS--WWAK-DSAPE--SKMAVLTLLAKVLQID---SSVSFNTNHEAFS-A   VFNTYTS     1688
Query_GECKO  1628  K   LVNAIL-KNWVHLDN--WRTR-DSAPE--SKMAVLTLLAKVLQID---SSVSFSTQHEAFS-D   VFNTYTS     1688
Query_DANIO  1627  Q   LAEQVL-KGWDLLRP--WWDGPAATPE--SKTSVLSLLAKVLQID---SSVCSNTSHPAFN-A   VFTTFTA     1688
Query_SNAKE  1628  K   LVAAIL-KNWGHLDH--WWAK-DASPE--SKMSVLTLLAKVLQID---SSVSF-THHEAFP-N   VFNTYTC     1687
Query_STERL  1640  Q   LVGGVL-GSWTSLQS--WWAS-SSSAE--SKMAALTLLSKVLQID---SSVSFNTSHLAFA-V   VFSTYTT     1700
Query_MILII  1624  L   LVSAVL-TNWTRLES--WWSE-NSSAE--SKMAVLTLLSKVLQID---SSVSFSSNHEAFA-M   VFHTYTT     1684
Query_BAMBO  1623  Y   LVSAVL-SNWAKLDS--WWSV-KSSPE--SKMAVLTLLSKVLQID---SSVSFGNSHDAFS-M   VFQTYTT     1683
Query_ALLIG  1586  K   LVTAVL-RNWMRLDT--WWAK-DSAPE--SKMAVLTLLAKVLQVD---SSVSFNTKHEAFF-A   VFNTYTS     1646
Query_GAR__  1627  Q   LVKRVT-QAWSSLGS--WWSSPASTSE--SKMATLTLLSKVLQID---SSVCFNVNHEAFE-A   VFTTYTA     1688
Query_TETRA  1589  -   ----ILiQNLPVINI--------ERKEd-------TFLEIELKLNwl-----IMSSIEGLViE[2]-YSDYVE[1]  1636
Query_PARAM  1508  -   LVNKVP---QNQVHH-fFKFQFKVNREn----------SALLLLEiqiNNIVLNQTKNINR--   ---EYVK[1]  1559
Query_SLEMN  1567  -   LLRSVP-ENWNEQER-tDFTEFEVATSk--------VKIYT-FIAnacQVQLRDRSNHYAI--[1]-FIDHLC[1]  1624
Query_STENT  1355  -   NLTIFL-KNFIKQTT-pLFATLALVPSp--EFSMGIIRAYCIFMEmakDLSIFNEGLHELQnN[1]-FFNHLV[1]  1421
Query_NEMAT  1610  V   VCASIV-NSWEVFSS--LWDM-SSSSE--LKQALLEIMKKLLLSD---PKCLIDPAKPAFQ-I   VLKMFTS     1670
Query_PISTI  1638  S   FVAVLL-NNWTSFEI--FWRT-HSAMD--LRHSILMILKKMLILD---PAKVTDCNRLTFP-S   LFEMYLL     1698
Query_CORAL  1619  K   MTTVLL-ESWERLSH--WWE--NGGEE--LKRSTAILIRNILIVD---PKCLSEVK--SSR-G   MFKTYNG     1676
Query_HYDRA  1543  K[4]FTFQFL-KYPEKLAT---GIRFEKTQI--QVLDIF-----ILILN---VFEGTLKDLPFLN-A   LIEV---     1599
Query_LATIM  1628  Q   LVDTVL-QNWTKLSS--WWIR-SSSPE--SKMAALTLLSKVLQID---SSVSFNPNHAAFP-A   VFNTYTT     1688
Query_SHRIM  1593  K   VVQSFL-KHWPTLAA---QAGKTSL----HQDSVLSLLSSLFLID---REGTIKSHCTSSR--[2]MVVFYNG     1652
Query_MCRAB  1588  K   VVQAFL-HHWSTYST---EAGKSSI----HQESVLDLLSSLFLVD---REGTMKAHCSAED--[2]VLKFYRR     1647
Query_AMPHI  1603  E   LVRGVE-TTWGTLAT---YGSTSDLVSncSTSDLVSLMSLLMLVD---QGRVVAERCNAAPpA[4]LLEFWLR     1670
Query_BARNA  1503  P   -LARLLaEGWQQLES---------ASEp----ALVQLLGLLLLVHlp----LVGEDQAASPvA[1]----YTR     1552
Query_COPEP  1520  E   LISHVM-QHWNSVKG-----SINSRTDi---LRTIFLLRSLSKVA-------------EFKdE[4]AVSWFCN     1572
Query_DAPHN  1447  -   IVQNVL-DSWKVLKK-----HAESYRQh---FNLQLLINCLASIDid----VVGRNSDVKQ--[1]VIGCWNQ[1]  1503
Query_STARF  1649  S   VHSAIL-TQWDILMG--GLKK-GCSTD--LQSFALLMLRKLLQID---AKFVSDTTHPVFS-S   VLEAYSF     1709
Query_STRON  1644  A   IYSAVL-TQWDTL---------SNLGQ--EDDFMLNLMKKVLQID---FKFATDPTHPTFR-P   VFNQYLS     1698
Query_SEACU  1617  V   VHSALL-THW----G--ALAR-DKSPD--AQRAALLILKKLLQID---SKFVMDVSHVAFQ-G   TFTSYLS     1673
Query_FUNGI  1567  S[5]MLS------------------SPQSTH---RENAVQLARQIIQNG----FVK--DSSEGLQwI[2]LLNKMMS     1617
Query_DPURP  1707  -[5]IQSIVN-FFKEISEW--TLNPNLKIVE---KENIVEIIKNLIKIN----PLQFFENKQCYDfV[2]IITNYLN     1773
Query_HALBU  1709  K[5]IM--IP-FLIKISSH--LLKDVGKTTD---KENIIEFTKNLINID----TPTFFETKEGYDfV[2]LLINYLN     1774
Query_GLOBO  1482  E[5]AIADSS-FTTSIISF--YYRTGEHSL-------TRFW-RKFLKIS----SVTL-AISEENAlS[2]FWSEFFL     1543
Query_PALUS  1537  R[5]FVKHSA-FLTKVVRE--Y---NSGNT-------LQLL-KKIVQLD----PKVL-ARTATSDfC[2]FTDAYVT     1595
Query_CALIF  1609  S[5]ISNQID-FIKSYLSF--HLDKKNFYI-------IIDFFKKLFKLD----NQFI-DKYASSIcK[2]FCKIYLS     1671
Query_GLOMU  1667  E[4]LNSEFY-FLKSFVKI--CTYDRHYL--------LKLW-NGLFMLE----PNLL-RRFSLEEfQ[2]FRENYEA     1726
Query_LOBOS  1582  D[5]LTKDYS-TLESIVNY--LGKDQISMT-------VAIW-KRLVALS----PRIL-RNSKTDTfV[2]FFKIFQS     1643
Query_RACEM  1566  N[5]LVSS-S-FLKDLIAF--WKSQPKLLN-------LAAFLKGLFGAD----PGLLYRTANKGHtF[2]FFDAIRA     1628
Query_PVIOL  1700  I[5]ILSIIG-FLEKIKQW--SLSNNLKIVE---KENIVEFTKNLIKIE----PIVFFETKQGYDyV[2]IVTNYLT     1767
Query_BOMBU  1492  E[1]LVDTVI-KKFTIFES--VINNLEDRRQ-----KLINILGIAVHLK----HKPIEVSSLSKD--   -FYMWIL[1]  1549
Query_AEDES  1604  K[1]VSEFLL-AKWAPICD--RASEIDNKFGs-TDLRLVEFMSNLAMAS----PFALSEIGR----K[2]GFERWLL[1]  1666
Query_ANT__  1519  N[1]TVDIIL-QQFTCLNN--AVFNVDSRKE-----RLLNIYSNAVHLK----SKPTEATQ-NHE--   -FYQWIR[1]  1575
Query_WHITE  1541  K[1]IFQK-----WDIIQI--YGNKYDNSY-----ARMSTVLDILMALV----KTFKNRVGCFRSpE[2]GIIKWLM[1]  1599
Query_BEETL  1495  V[1]IVLTN----WDIFEE--FFNES------------LSNIQLGYKLL----KLFNPTIKSSRL-Q[2]NLKKWLL[1]  1546
Query_THRIP  1513  A[1]LFSK-----WDKFIF--WANADPCKE-----AWLISLINQIVEL-----DIDFHSVE-----N[2]HICSWII[1]  1565
Query_TERMI  1576  L[1]AFERSIvNHWNKLSS--FVEGNSEQV-----YQLLGLVVNLARIM----PQPDRGLA-----R[2]ALQKWVI[1]  1634
Query_FLEA_  1513  S[1]ISETML-NHWNDIVQ--HVTESD----------AVDLVTYLGIIV----PV--SNVVG----P[2]YLAEWCM[1]  1564
Query_MIDGE  1564  D[1]NLQTMI-KHWPIIYQ--ISQEMDNNLNs-VDLAIINLITKMALTS----PIELHELGN----R[2]NLQIWLF[1]  1626
Query_LAODE  1550  -   ---ATL------IDKd---FPTDVSREq--LDSAITLIYSTRNFLpg-NNIRMSKLNKIFCfL[4]RYYEYIS     1608
Query_LOTTI  1595  N   VASTIL-SEWKKFS--tWWTN-NAPPE--LQNLAISLLTKVLLID---SRVTSHTSHPSFS-T   IFEMYQT     1655
Query_APLYS  1637  A   VCREMV-SSWREVG--sWWGA-DSSAE--TQGLALLVLTKVFLID---SQFIKELG-PGVE-S   VFEMFLS     1696
Query_SCALL  1634  N   IVKSVL-MQWKTIA--gWCDS-KSSID--LQSMTLLLLTKLLLID---SKFVSTPSSPSFE-D   VFGTYLQ     1694
Query_CRASS  1606  N   VVTSVV-EKWKSIS--sWC---LATPE--LHSMAMLLLTKLVLLD---SKVVTDSSSGHFP-E   VFETYLS     1664
Query_OCTOP  1571  Q   VCDIIL-KNLNMIEN--WWRDDASED---TQRITFDILLGILRLD---SKVLSDGQNPTCE-L   LHKFFYI     1631
Query_POMAC  1613  R   IVDMVA-SKWHHLE--sWW-Q-AGSSD--LQGLAVLLLTKLLLVD---SKAVLNVEKSHFT-T   VFTSYCT     1672
Query_WHIPW  1562  R   LVDDVV-SLCNLKMN--LWADVSCSSEn-LKLLWVKLVMKLLLIN---SKI---DNWPCAS-S   IFRCYFK     1621
Query_TRICH  1570  E   LALQIL-QIFGQELL--YWADVVVSNEt-LKMLWLKLMMKLILID---SKI---TEWPEFA-P   MASAYFM     1629
Query_AASTA  1466  R[5]LVAPFA----KFV--------LHATTEk--EQP--LLMSTLLLFV----RSILDKSTL-----[1]LVDCTAG[2]  1519
Query_AEUTE  1429  -[3]YVNDFA----KYA--------L-STAAk--DQS--SLLKILDQVF----NNFQDVKSL-----[1]LFDCALG[2]  1478
Query_PLASM  1528  S[5]FTATLV----AAVKQ--FINKLETVKRe--TQPALSLLQILKHFV----ELSQNCSSTYVEaL[4]ILDIQNA[2]  1597
Query_BREMI  1514  S[5]FTKTLL----PNAKH--FVDSLDAFKAn--TQRALSNLQVLYYIL----ELWLHCSASFIQaL[4]ILDIQNA[2]  1583
Query_OLIGA  1542  S[5]FVEDFL----PHLKS--MIGKMADEGNd--AQRTRLCLESIRQTV----DIFKHCSRRMADvL[4]FPDMRDA[2]  1611
Query_LIGNO  1673  K[1]LASQIL-AKWTLVQQ--LWDY--GNDDc--RFSVLGILTNLLAVD---AELVKSND----Q--   VFVLFVG     1728
Query_CESTO  1456  V[5]HIEPLL--RFKDFDT------------c----LFVDLLEAIISLSqv-------------ElD[3]FISSFLH     1503
Query_OVI2_  1802  V[3]IVSRIL-DSWTQLFEr---LVCASPYQqwSNQVMLGFMERLLMLGp--------------KgT[4]LVPTVVF     1861
Query_PRIAP  1623  G   VVNTVL-QCWGHLRD--LWVA-TSTPD--RKDMLLMLLRKLLLID---SSVATKGSHSAHR-D   VYDTYCC     1683
Query_ROTIF  1594  K   LILKVF-ESWPKFEP--YWSF---ESQ--FKILMISLLTKCLMIE---SPQPKNEK-------   VFQMYFE     1646
Query_SPONG  1620  K[5]LCTELI-SHWQDLSL--WWED--GSKD--LKSAAVTLLQKIIALK---PKLLLKSADVSKP-L   V-DMYTA     1683
Query_TARDI  1625  T   LEKLVM-QSYGSFVE-aRPDEVKKDMD--TVRSVLGVIKLLLQIS---QKTLKTGDAEVKK--   LYDFYIS     1686
Query_TRICH  1600  H   ISKKIL-SSWNSLKC--WWLT-PDAQD--LLQISVNILNKCLTLD---PQFATQSNDSSFE-T   VFNFFLS     1660
Query_OSTRE  1566  R[5]NISHLR-HLYEDHDS-----ATAEMRNt-CHQRFLTIIRKVMTLD---ANLRIDESIASND--   ---TIVS[1]  1627
Query_NITEN  1681  A[5]MAA-----NLRYLQR--WTEPEASLEE---KRAFLAIVGKLLVVD-g-GDPPRALAQPAARvV[2]AYEAFLS[3]  1749
Query_CLUBM  1511  G[5]FLL-----HLTELRP--CYLDGSSVTD---KQFIAGIIHKLLVLD---KASVVGGSQPAFQfI[2]VLFAFLS     1575
Query_MOSS_  1600  S[5]FLQ-----HIEVLAP--LCRAGSSLTL---RLFFLDILHKLLMLDtg-KDVVLNIKYSSFDfI[2]TFLSFLS[7]  1673
Query_LIVER  1623  D[5]MIL-----HLKCLHP--CCASYAGLAD---KQFFLEILQKVLLLDlr-SENLLQHDQSSRSfI[2]SFISCID[6]  1695
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1674  LI   SDKNNT   LQFKTRVLDLL-I-FF   CN   L--PAE-ENKQLESCLNQFVAD---CFPMS-SLEFHA-GSQQ  1731
Query_LINGU  1676  ML   TDTKSS   LLFKCKVLDLL-V-FF   AK   L--PEK-ETQQLKACLDRIVAD---HFPLR-SSEYAV-GSPK  1733
Query_LIMUL  1662  LL   TRKGIS   LSFMGHVLDLL-A-FF   LK   NndTNK-EVL-IRNALDELSAN---TFPLQ-STEFVV-GSLR  1720
Query_SPIDE  1661  ML   SKENNT   MDFIVHALDIL-P-FF   CE   SstSSE-NSTNLRSHLEEMMVN---HFPLQ-STEFVP-GSSA  1720
Query_SCORP  1663  LL   QKSEMS   LEFKICVLDLL-P-FF   LC   DafLCK-----VQAALNDLFTN---HFPLY-SSELKP-GSSQ  1718
Query_TICK_  1589  LL   ERSGAT   LADKVLVLDLL-P-VF   AS   Di-RPD-TGLKLRDALEEMVAQ---HFPLD-SSTLLP-GAPE  1647
Query_MONOS  1574  -L   -ESNST   LAEKADSRQLI--gVL   CS[5]LagTEDlMPNVAADSVRKMCGQ----FPFQ-SSEYGP-GTTK  1636
Query_ROSET  1562  FL[1]-DRAVD   MTSKRVALQLLgP--L   CK[1]----TQhVRTECAASLNAFCQA---SFPLR-SREYAV-DSTQ  1619
Query_BELCH  1704  ML   TDSRST   LPFKSQVLDLL-P-FF   TV   L--PQT-QLAELKARLDRMVAD---NFPLK-SAEFPA-GSSK  1761
Query_CIONA  1674  TI   SSNKLD   LSFKNQFLDLL-S--F   FT   Ls--ST-HSEEIKETLDKFVSN---HFPLR-SSEFVA-GSEK  1730
Query_CAECI  1691  IL   TDQKLG   LNLKSQAIIIL-P-FF   TN   L--TGD-SLNDLKNALDQLVAF---NFPLK-SDEFPK-GTLK  1748
Query_HUMAN  1678  LL   ADTKLD   LHLKGQAVTLL-P-FF   TS   L--TGG-SLEELRRVLEQLIVA---HFPMQ-SREFPP-GTPR  1735
Query_MOUSE  1675  LL   ADTKLG   LHLKGQAIILL-P-FF   TS   L--REG-SLENLKHILEKLIVC---NFPMK-SDEFPP-DSLK  1732
Query_XENOP  1700  IL   TDQKLG   LNLKSQAIIIL-P-FF   TK   L--TGE-KLTELKNTLDQFVAS---NFPMK-SDEFPK-GTLK  1757
Query_PLATY  1682  LL   IDKKLG   LNLKSQAVLIL-P-FF   TN   L--TGD-RLNDLKNALEQFVAF---NFPMN-SDEFPK-GTLR  1739
Query_TASDE  1685  LL   DDPKLG   LNLKGQALIIL-P-FF   TN   L--TGE-RLEDLKLALEKLVAS---NFPMT-SHEFVQ-GTLK  1742
Query_CANAR  1680  LL   TDQNLG   LNLKGQAVIIL-P-FF   TN   L--TGE-KLNDLKNALDQLVAF---NFPMR-SDEFPK-GTLK  1737
Query_OSTRI  1650  LL   IDQNLD   LNLKSQAMIIL-P-FF   TN   L--TGE-KLNDLKNALDQLIAF---NFPMR-SDEFPK-GTLK  1707
Query_TURTL  1689  LL   IDRNLG   LNLKSQAVIIL-P-FF   TN   L--TGD-RISELKNALDQLVAF---NFPMK-SDEFPK-GTLR  1746
Query_GECKO  1689  LL   TDQNLG   LNLKSQAVIIL-P-FF   TN   L--VGE-RLVDLKNVLDQFVAF---NFPLR-SNEFPK-GTLR  1746
Query_DANIO  1689  LL   TDVSMP   LNLKSQALIML-P-FF   TA   L--PSM-PLEELRRALESLVAT---HFPMQ-SDEFPR-GSLQ  1746
Query_SNAKE  1688  LL   MDQKLG   LHLKSQAIIIL-P-FF   TE   L--VGE-GLHKLKHALDQLVAY---NFPLT-SDEFPK-GTLK  1745
Query_STERL  1701  LV   TDSTLP   LNLKSQALIIL-P-FF   TN   L--LGS-TLADLHNALEVLVAS---HFPMQ-SDEFPK-GTLR  1758
Query_MILII  1685  IL   TDSTLA   LNLKSQAMIIL-P-FF   AH   L--PEE-NLAVLRHALDQFVAA---HFPMK-SDEFPK-GTLK  1742
Query_BAMBO  1684  IL   TDSTLA   LNLKTKAVIIL-P-FF   TQ   L--PEE-NLVTLKNALNHFIAS---HFPMK-SDEFPK-GTLR  1741
Query_ALLIG  1647  LL   VDQNLG   LNLKSQAVVIL-P-FF   TN   L--TGE-RLKELKDALDQLVAF---NFPMK-SDEFPK-GTLR  1704
Query_GAR__  1689  LI   TDSALP   LNLKNHGLIIL-P-FF   TN   L--PES-TLAELQRALEMLVAS---HFPIS-SDEFPK-GTLR  1746
Query_TETRA  1637  LI   -EVSLK[4]LFIHEQGFNLL--gFY   LK[1]InnADFfVKRGLRDKVNSVRSKy---FPMK-VKELKR-DSKQ  1701
Query_PARAM  1560  LL[3]LNTGRS[4]-------IDAL--iLF[3]IK[2]-----D---TDLPNYLNDLANKy---FPIK-ISDLKK-GSAE  1617
Query_SLEMN  1625  KV   -SKYLK[4]VHIKREVLNLV--gM-[5]TQ[5]VhdKDN-----INFSLRTLQSQf---FPIK-TRELVR-NSNV  1692
Query_STENT  1422  LL   -EKFMA[4]FNVKKESFHLY--aIS[7]IY[5]LffE--------------IQNQy---FLPS-TQNL-Q-PSKE  1482
Query_NEMAT  1671  VF   DDKNLT   LTFKSRAVELF-P-YF   KN   L--PDN-YQSQLKSCIDTLVTD---NFPLK-STEFIP-GSPR  1728
Query_PISTI  1699  YL   KDKSTT   LAFKGQVLELL-P-YF   TK   S--TEE-YLRRLKEGLNAFVTN---NIPLK-STEFVT-GGPQ  1756
Query_CORAL  1677  LL   KDKTIT   LTVKIQCLELL-P-CF   TS   AtiPSE-QTEALRSALKTLIAD---NFPLS-SKEFNE-NNPK  1736
Query_HYDRA  1600  -L   FDSKNS   LAVKEKVISCS-PhIF   KK   Gs--SQ-QCADFKVALEHVVAN---DFPVN-SDELQI-RSLK  1657
Query_LATIM  1689  LL   TDQTLA   LNLKSQAVIVL-P-FF   TN   L--PED-HLADLKNALDQLVAF---NFPMK-SDEFPK-GTLK  1746
Query_SHRIM  1653  LL   ADAKVE   LQAKIKVLKLL-P-FF   FT   Mg--ED-IAKSVSEALEMLSLN---HFPLH-STEFSE-GGAR  1710
Query_MCRAB  1648  LL   IEDKAT   LSMKNKALKLL-P-FF   LT   Ms--KQ-VSKSISEDLEMVSFK---HFPMQ-SSEFPV-GGSR  1705
Query_AMPHI  1671  TV   ADPKTE   LQVKITCLQLL-P-LF   CA   Ln--DH-VQQKVVEALEFFSVH---HMPLK-SDEFAA-GSAR  1728
Query_BARNA  1553  LL[1]-GDRLT   LEQKRPALRLL-TaML[2]AT[1]-----S---ANICSLLERLESR---HFPLE-SSELAA-GSPA  1609
Query_COPEP  1573  LL   KNPSIE   LKVKYKGFQIL--gLM   LQ[1]----TDg--NDIKDALLSFSSL---HFPAT-STELVP-NSAA  1628
Query_DAPHN  1504  --   TDVTLA   ----SDLVQLL--nVF[3]CS[1]---SEE---KTLFASLQQFFTH---NFPSK-STELEK-DSDK  1556
Query_STARF  1710  ML   TDASTS   LTFKAQVMDVL-F-FF   TS   L--PDG-EQTKLRSALDRLVAD---NFPLN-SSEFRV-GSPK  1767
Query_STRON  1699  ML   RDSKKS   LAWKTQVLDVL-Y-FF   AN   V--PEK-EEKELKSALDLLVAN---HFPLK-STDLEA-GSPR  1756
Query_SEACU  1674  ML   TDPSTN   LTFKTQALEIL-Y-FF   VQ   V--PDS-ESQKVREALDQLVVD---NFPLK-SSEFGV-GSAK  1731
Query_FUNGI  1618  --   --RSIP   LSLKMQVIELL-P-GI   CV   NqkMRE-G-A--KEMLLDMVVY---DFPVS-SKELPV-GSAA  1670
Query_DPURP  1774  --   --RSNP   LPFKNKVLVLL-P-YL   LS   FpkHHN-F-SLIKEKLNEIIVY---DFPLH-SKDLTV-KSPI  1828
Query_HALBU  1775  --   --RSNS   LAFKNKALVLL-P-DL   LL   FpkHHN-Y-NLIQEKLNEIVVY---DFPLS-SKDLTK-GSPI  1829
Query_GLOBO  1544  --[1]FNHRSS   LKALSDLFDLL-P--L   VM   VanKFQ----KLEESLSSIVVE---KFPLV-SELTED-SNEQ  1598
Query_PALUS  1596  --[1]LDVANP   LATITDALNLL-P-AF   LL   VpsNAA----KLERRLNDIVVN---QFPIS-PKDLVE-GTAQ  1651
Query_CALIF  1672  --[1]LEPSIN   LKVKSESLKLF-P-IF[1]SL   LtnNYK-D--QLEKSIEFIIYN---QFPIS-SNDIIK-ETLQ  1729
Query_GLOMU  1727  --[1]LDKQID   LEFKNEALDLL-P-FY   FK   SdlPKE----KIESCIEFILYQ---QLPTD-SKDYSKfGTKK  1783
Query_LOBOS  1644  --[1]FERDPD[6]LPVMSEAFPIL-P-VF   LT   YpgQRT-V--EFESVFNRAVLN---LMPMS-STEYER-GSSK  1706
Query_RACEM  1629  --[1]LDRDAP   LSVKMEAFDVL-P-VY   LT   VqaEHV-D--RTIALVKETASW---QLGIR-SRDLRV-GSEE  1685
Query_PVIOL  1768  --   --RSNP   LSFKNKVLELL-P-YL   LS   ChkHLD-Y-SIIKEKLNEIVVY---DFPLN-SKDLTQ-RSPI  1822
Query_BOMBU  1550  QL   MDNLD-   IEYKIYILHNF-L-VC   LT   --dMTD-TKPEVLVILRNLKNN---RLDVC-PNDFSQ-KNVK  1606
Query_AEDES  1667  LI   ANAQN-[1]LDLKAKAMFLL-P-AL   VG   --dSDF-EKASVAEALQKLQNQ---HFPLR-SSEFIA-NSAE  1724
Query_ANT__  1576  QL   MNNND-   LEYKIKILKNF-F-VC   LT   --dVTD---PDLQMNLQILKIN---RRSLC--SDLSE-TSVN  1629
Query_WHITE  1600  QL   -GKDI-[1]VNLKFSILDFL-V-AL   FN   --pGSQ-RDATYDNVIEPAIHA---YTAAL-ETHILG-ESSF  1656
Query_BEETL  1547  VL   TYETNV[5]FELKIETIDLL-S-DI   LD   --nLDE-NT-DLRRSLDCFKAK---LISET-SPIPAE-ERPF  1608
Query_THRIP  1566  LL   TDSSR-[1]LSLKTDLLYCL-P-IV   CS   --iSEP-S--DVKNSLLSLKNQ---HFPVQ-CSEFRP-GS--  1619
Query_TERMI  1635  QL   TSRDL-[1]LHYKSRALDLL-V-CL   AG   --pQED-TNQELRSALQTLREQ---HFPQL-STELHS-GS--  1690
Query_FLEA_  1565  LL   MNTTV-[1]LDTKSKVLDVL-P-FI   TG   --pKDI-ENPELTQALKTFQSL---HFPLE-SKEFRE-GSIE  1622
Query_MIDGE  1627  FL   QNRNY-[1]FELKSKGVFLL-P-CV   AS   --sND----TALNEKLMNDLNSmqyNMPLR-SREFPE-GSLA  1684
Query_LAODE  1609  --   ------   -SLRPRILKTL----L[7]SS[1]-heESD---AYLKETLEKFIYYnr----IYiSSENWG-GNSD  1662
Query_LOTTI  1656  ML   VDNKTN   LAFKNKVLDLL-P-FF   AM   A--QET-YITKLKSSLDRFVAD---NFPIK-SNEFTK-GSPR  1713
Query_APLYS  1697  ML   TDCHTT   LAFKNQVLDLL-P-FF   VT   L--KGE-QSSKVIDGLNRFIAD---NFPLK-GNEFSK-GTHK  1754
Query_SCALL  1695  CL   TDTRTN   LAFKCRVLDLL-P-FF   IN   V--PKP-HVTKLKTSLDRLVAD---NFPLK-SSEFPK-GSSQ  1752
Query_CRASS  1665  ML   TEKKTT   LSVKCKILELL-P-FF   AT   L--PEP-HLKKLKRSLDRLVAD---HFPIK-STDLPE-SSPH  1722
Query_OCTOP  1632  LL   SKKTVA   SSFKVSLLDLL-P-FL   SV   SpePHN-SRL--RTNLKKFLMD-----------KSQP-GTLL  1682
Query_POMAC  1673  IL   AEKKTN   LAFKNRALDVL-P-FF   AA   V--PD----AKLKDVLDGLVAD---NFPLR-SSELTK-GTPR  1727
Query_WHIPW  1622  LL   NDQSLC   LGFKNQMLALL-A-AF   CC   C---EK-EKENLESELKLFVVN---NFPLK-SSEFAS-EAAK  1678
Query_TRICH  1630  LL   NDQSLT   LDFKNRMLSLL-A-GF   CT   F---DR-HACELGRSLKLFMVN---NFPLK-SDEFAK-GTAK  1686
Query_AASTA  1520  QL[1]CDPSTS   VGVKVAALQLI-P-VL   IQ   -anATE-FGPPLVDAVGHLVVH---GFPVH-STDVAR-GSLD  1579
Query_AEUTE  1479  QV[1]CDKSKA   ASLKVDAMHII-P-HM   LK[2]-pdCTE-I---MTEGIAHLVVH---DFPVN-SNDITK-GTVA  1537
Query_PLASM  1598  LL[1]KQRGCS   YLVKADLLQIL-G-LL   GP[4]AtkERQ-PSDPILNALIAFVFD---EFPIV-SVDVSR-GSRD  1662
Query_BREMI  1584  RL[1]KERGCS   YLVKANSLRLL-R-LL   GP[4]VssEEQ-VLKMTLDALVEFVFD---EFPIV-STDVIR-DSKD  1648
Query_OLIGA  1612  TV[1]KRREIS   SSAKGDCLSIL-A-GL   G-   TalSAE-GIKQIAEVLAAFVVD---EFPIS-SADVHR-GTRD  1671
Query_LIGNO  1729  VL   -NSRVN   FDLIVNGLALL-G-YM   CR   Np-STE-RAERLSSQLKLLCSE---TVLMRcSNQLGI-NSHA  1787
Query_CESTO  1504  LL   KDETLN   VPTKVKNLDML--aFF[7]TN[2]LsnERRy---EIVEAVKYLFAS---NLPLD-QHEIAS-SPVK  1570
Query_OVI2_  1862  LL   RDPGLS   LGQKICVLDML--sIF[6]AS[5]CasETDi---NILQALKEFLAT---QLPIN-QEEFKRdGSKS  1931
Query_PRIAP  1684  LL   VDAGTS   LAFKTRVMELL-H-FY   CE   L--PEQ-ETTKLRENLDRMVAD---FFPLR-SSDFIV-GSPK  1741
Query_ROTIF  1647  FL   TDEKSM   LNFKAKVLDLL-H-FF   TN   Sp--------KFKVSLSQFMS----QLPLR-SQDLAK-SDDI  1698
Query_SPONG  1684  MI   GDEKLE   LNFKAVMIDLL-P-SF   LL   Ls-SPE-HQSQLKGSLNRLVSL---QFPLT-SSELPA-GGPL  1742
Query_TARDI  1687  VL[4]ASDPQG   AATKNRALTLL-AaFL   DG   A------NGAVAETAMHEMCVN---DFPLK-AAELPK-GSPQ  1746
Query_TRICH  1661  LL   SGTSTS   LAFKKYVLDLL-H-FF   LV   S--SDK-RIRELREHLDRMIVD---KFPLK-SAEFVV-GSPD  1718
Query_OSTRE  1628  LM[3]RNDCIA   LSSHVDAITITcD-VL   -Q[4]LraQKDaLEKSLGDSLKNMMVQe------R-CDAFTF----Q  1689
Query_NITEN  1750  AP[4]SERGLR   VSLKREALGLL-P-FF   LG   PgvDGE-TRRRVLDSVRRLVAE---HLLIQ-SRDLPQ-GSPQ  1813
Query_CLUBM  1576  AT[4]WSE--C   AMLRKSALLLV-P-FF   LE   EsmDSS-LRTRLIGSLTSLVTS---HLLVR-EADLPD-GSME  1637
Query_MOSS_  1674  NS[4]S---AA   FTLRTTSMRLL-P-FF   FS   AnvDSG-VHSKLIRILSGIVTD---YLLVR-EADIPQ-GSIQ  1734
Query_LIVER  1696  DS[4]SDKEAS   ITLKTHVFVLV-P-LF   FK   PtfGQE-FQSRLCRSLTKVVSD---NLLVR-EADLPK-GSIQ  1759
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1732  ---QKCYITALDRLLLAM-ELSGSLM   LLNLIISVICKEV-K-HTHENAIQSAFQR---FIKRL    PKDKH--KLA  1793
Query_LINGU  1734  ---YQEYIMALSKILSAL-ELSGSLM   LLELLISIMCREK-N-HVYEDRMQESLVT---FIRRL    PTGHQ--KAA  1795
Query_LIMUL  1721  ---YNSYVTVFRKILTAL-ELCGSLT   LLTFVVGIMCREA-C-HLLEEEIQRSISR---SVCRL    KPEEQ--VNS  1782
Query_SPIDE  1721  ---YHTYISALKNMLSAL-SLTGCGF   LFEILTSFFCKES-E-HVLEVNFLQCITD---MLKSK    SSSHH--QLY  1782
Query_SCORP  1719  ---YNDYHLAFTKLLIAL-ELTGNEM   LLKFVIQIMCRES-T-HLLEENIQKCLKK---LCKRT    DSNKE--RLI  1780
Query_TICK_  1648  ---YLCYVSAMRKMVAGL-ERSGSLV   LLELLIRVLYRERgS-HFLEAEVLLALQR---TLTRC    GETKQ--QQL  1710
Query_MONOS  1637  r---RELVATLNDVFRGL-VVSGAEP   LFKVVLELYVKEG-EaHVHHEGFVQAFSS---FAAQT[ 6]----A---SV  1700
Query_ROSET  1620  ---HREYIAALDDILDAL-VLSGSLE   LLEVVVTVHQRED-RtHVHEDHIQRKLRDf--IVQEH    TDAPQ----A  1681
Query_BELCH  1762  ---YNDYVGAVTKLLSAL-ELSSSLM   LLELLISVMCREA-K-HVMEDDIQTSLSR---FIKGM    GADRQ--KPA  1823
Query_CIONA  1731  ---YNSYITAMRKLLDCM-ELSGSLE   ILNTMVTLYCRDQ-S-HVMSDEIMSSLEK---CVASI    PVSKQ--CPA  1792
Query_CAECI  1749  ---YNNYVDCVKKFLDAL-ELSQSTM   LLKLMTEILCRDT-K-HFMEELFLTSFKR---IARRG    NCDKQ--VAL  1810
Query_HUMAN  1736  ---FNNYVDCMKKFLDAL-ELSQSPM   LLELMTEVLCREQ-Q-HVMEELFQSSFRR---IARRG    SCVTQ--VGL  1797
Query_MOUSE  1733  ---YNNYVDCMKKFLDAL-ELSQSPM   LFQLMTDILCREQ-R-HIMEELFQTTFKR---IARQS    PCVTQ--LNL  1794
Query_XENOPU 1758  ---FNNYVDCIKKFLDAL-ELSQSPM   LLQLMTEILCRDE-R-HFMEELFQSSFKK---VIKRS    SCDTQ--VIL  1819
Query_PLATY  1740  ---YNNYVDCTKKFLDAL-ELSQSPM   LLQLMTEILCRDQ-R-HFMEELFQMSFKR---IARRS    GSAKQ--VAL  1801
Query_TASDE  1743  ---YNNYVDCTKKFLDAL-ELSQSPM   LLQFMTEILCRDQ-R-HVIEELFQTTFQR---ISKRS    SSATQ--VAL  1804
Query_CANAR  1738  ---HNNYVDCTKKFLDAL-ELSQSPM   LLQLMTEILCRDN-R-HFMEDLFQASFKR---ISRRS    PTDKQ--VLL  1799
Query_OSTRIC 1708  ---HNNYVDCTKKFLDAL-ELSQSPM   LLQLMTEVLCRDH-R-HFMEGLFQASFKR---ISRRS    NSDRQ--VVL  1769
Query_TURTLE 1747  ---YNNYVDCTKKFLDAL-ELSQSPM   LLQLMTEILCRDH-R-HFMEDLFQTSFKR---ISRRS    SYEKQ--VVL  1808
Query_GECKO  1747  ---YNNFVDCTKKFLDAL-ELSQSPM   LLQLMTEILCRDH-R-HIMEEQFEISFIR---IARRG    SCYKQ--VVL  1808
Query_DANIO  1747  ---CNNYMDCIRKFLEAL-QLSQSPL   LLKLMARVLCRDK-K-HIMEELFQACFQK---IAHQS    YLGKQ--VLL  1808
Query_SNAKE  1746  ---YNNYVDCVKKFLDAL-ELSQNST   LLELMTEILCRDN-K-HIMEELFEINLKR---IAKRG    SCERQ--VLL  1807
Query_STERL  1759  ---YNNYIDSIKKFLDAL-ELSQSPM   LLKLMTGILCRDN-K-HVMEELFQTCFKR---IARRG    SCEKQ--VSL  1820
Query_MILII  1743  ---YNNYVDAIKKLLDAL-ELSQSPM   LLKLMIAILCRDS-R-HVMEEGFQVCFQR---LAKTG    SCEKQ--VAL  1804
Query_BAMBO  1742  ---YNNYVDCVKKLLDAL-ELSQSPL   LLQILTEVLCRDN-R-HVMEEAFQICFQN---IAKRG    TCEKQ--VAL  1803
Query_ALLIG  1705  ---YNNYVDCTKKFLDAL-ELSQSPL   LLQLMTEILCRDH-R-HCMEDLFQTSFKR---VIKRS    DCNKQ--VVL  1766
Query_GAR__  1747  ---YNNYVDTIKKFLDAL-ELSQNPM   LLKLMAEILCRDK-K-HVTEELFQACFKR---IARRS    GCERQ--VVL  1808
Query_TETRA  1702  ei---DFKSILGCMINMI-KLAGQIE   LLEILFPVIRESL-E-EPYHTKIITFIQQ---FIKEN[ 4]TVNKQ--LYF  1767
Query_PARAM  1618  el---NFKTILKKLTEMI-KLCGQIQ   YIEVLFPLLRQDL-N---FDLFIQEFLQIn----QES[ 1]---QQ-----  1671
Query_SLEMN  1693  ---YNNFHIIVEGFLSLL-ILSKNLK   VLRMIYNSMREWK-T-T-YEPSLKRTLKI---IIEQS[ 4]-PLHQ-fKQT  1757
Query_STENT  1483  ---TNDFEIMSRSFFELF-NRIPTLN   NLELLHTLIREDK-S-L-YEKELNQITMV---LIK--[ 2]-VAEG-yETT  1543
Query_NEMAT  1729  ---YNDYIATLDKLLVTM-VTSGSLV   LLQVITAVLCREE-R-HVHEEEIQKSLAA---FIQRL[ 2]SSSLV--KGA  1792
Query_PISTI  1757  ---YNDYVQTIDKLLAAF-VNSRSLI   LLEVLIAIFCRER-Q-HVHEDQIQQSFIA---FIQRL[ 2]EPDQA--KAA  1820
Query_CORAL  1737  ---YSDYIAALDKILSAL-VISGSLM   LFDLLISIVCREE-S-HSHEEEIQIQFAK---FIKRL    KLQQA--RKA  1798
Query_HYDRA  1658  ---YAEYTSCLSQFLRVL-VNSSSPL   LFEVLVSVFVREV-K-SVYEKVILPALQM---FAKNL    KVNDV--DIV  1719
Query_LATIM  1747  ---YNNYVDCIKKFLDAL-ELSQSPM   LLRLMTEILCRDS-K-HCLEELFLRCFKK---IVRRG    SCDRQ--VTL  1808
Query_SHRIM  1711  ---DREYHQAICEILTAL-ELSLSPI   LLQFVVNLFCRED-K-HRYEEVMHESLEK---FMKKL    PLLKQ--QET  1772
Query_MCRAB  1706  ---DRDYHQALSEILTAL-ELSLSLE   LLKFILQVFCNED-K-HRYEEAFNESIEK---FIKRQ    QVSKQ--IDT  1767
Query_AMPHI  1729  ---DRDVRRCLKQLLTAV-ELTARPV   LLLHLAKIFCHES-N-HRYKDLLESSMSK---VMAGV    GTEAQ--FSL  1790
Query_BARNA  1610  ---LATYRAVLASLCEAL-EMTAHPA   LARYMCVLMCREA-R-HPCSGLLRAALTAa---AGRH    GTAQQ--LAL  1671
Query_COPEP  1629  ---YNLYMSVFRQLLRTL-ESTKSPT   LLYFIISIACREP--cHVCEDDIQSTLSN---LISRS    SSDQQ--FEL  1690
Query_DAPHN  1557  ----SAYFSIFNRMIGLI----CVPV   VFHFLVDVFALEP-N-HPCLEPFIQSVQNen-WTESI[ 3]SLDKV-----  1616
Query_STARF  1768  ---YIDYAGALNKILTAM-VLSGNLM   LLEIVISIMCRDE-K-HAYEDEIQESLAA---FIKRL    PEEKQ--TAA  1829
Query_STRON  1757  ---YNDYIMALNKILTSL-VMSGSLM   LLELIISILCRDD-K-HSYEDKIQGALSS---FAKRL    PSEKQ--SAA  1818
Query_SEACU  1732  ---YNDYIAALNKLLTAM-VLSGSMN   LLEIVIGILCREE-K-HTYEAEIIKSLES---FIRRM    DKSKQ--VSA  1793
Query_FUNGI  1671  ---YTDYISAIDKILSSL-EISISDR[8]FLEALFPALS-EM-N-HSHFGSMSVSIEK---AITSC    DSSSA--CQL  1739
Query_DPURP  1829  ---YNEYVTCIERLLETL-EITKNP-[1]IIDILLHVLK-ET-D-HSHVHYINLSIER---TILTT    NDSEA--KEI  1889
Query_HALBU  1830  ---YNEYITTIERFLNTL-EITCNP-[1]IIDTLLQVLK-EQ-D-HSHITHINQSIER---FIVRT    NDSQA--KEV  1890
Query_GLOBO  1599  ---MNDYRVCMNKLIEAMaRCSNSAM   VERILMLHMCRNP-S-HPFNESVMDCISA---KASSV    SFEVF--DEL  1661
Query_PALUS  1652  ---YSDYVDAMNSLMKAL-ELSDGVV   IPKVLARHVCRES-D-HMFAVSLKKAMGI---AARRL    SASDI--DML  1713
Query_CALIF  1730  ---YIEYKNTVTAFINLL--STGNFI   IIKNLILLTCKEE-N-FKYQEKIKEFIAI---IINNY    SEEQF--NEF  1790
Query_GLOMU  1784  ---LNEYLSVLDKMFKAM-VRSRSVS   LFKILIPTFCRES-D-HMYGEEFQRQVSE---FLSKL    SRSKY--VDT  1845
Query_LOBOS  1707  ---FKDYINALDLLLKAL-VASSSFS   LFKTLMGIAIRES-N-HPHMEQMQQHISS---FALKQ    PLSKV--LEI  1768
Query_RACEM  1686  ---YNQYIAMLDKIIEMM-VSFRSTP   LFNILLPLYIQEE-D-HIHAEAFKEGFAK---FAKTL    DMKRF--HEL  1747
Query_PVIOL  1823  ---YNEYVTTIERFLETL-EITKNP-[1]MIDVLLHVLK-ET-D-HSHIQFINLSIQK---IISST    NDQQA--KDL  1883
Query_BOMBU  1607  ---ALKITNCFHKLVMLL-PVTKSVI   VFETIILFATGIA-K-YLCNETTNEYLQK---YFDSI    TTDYA--VKS  1668
Query_AEDES  1725  ---RVSFENCLTALLDAM-VVSRSPV   LLKAVIDATAADP-E-HIAEHKIRIALEK---YMSSQ    NSNQQ-cFNL  1787
Query_ANT__  1630  ---AMKVIDCFETLLVLL-STTKSMI   MLKCVIHFAAGAG-N-RLINEKLEEHLRQ---YYYGA    SLEHV--LES  1691
Query_WHITE  1657  dknYILVKTSLNKLLAGL-ELTGNAF   LLKRIMIRTSTDL-N-QNIRNDIKKTLAN---LTKKK    NDII--kKDL  1721
Query_BEETL  1609  ---------ILDKLLSLI-PKIRCFE   LLVFILEIYAKMS-T-GNSISDVKHYMTE---NI--M    NNQQEhkLLL  1664
Query_THRIP  1620  -eeLATFKRAFEALLQSL-VITGSSS   ILNIVIRTMTVDP-N-HNCRHLLPSTLQK---YMAVL    QKSQDhqVES  1685
Query_TERMI  1691  -vgYSGFVNAFRKILRGL-EESGSCV   ILLSVVAMAASES-Q-HVLEDQIQKSLQI---FIKRL    EDKQ--qVKA  1754
Query_FLEA_  1623  ---RASYMSAFSALLNSL-CKSQSPI   LLERIISIIAADL-D-HICEDEIQEKLCR---TIENC    APENQ--FNL  1684
Query_MIDGE  1685  ---RVGLVRFTEELFQAL-LTSRSPA   VYRFIINITISDD-A-YIMESKLQRVQVN---LMERL    SSTEQ-eVIM  1747
Query_LAODE  1663  f-----YKEVFNILLEMI-RRTGNAD   LVMLVTKALADDP-S-HLYRDKFKSILGQssvFLKNN    ---SA---AI  1721
Query_LOTTI  1714  ---YQDYISAINKILLAL-ELSGCSM   LLELLISIFCRES-K-HVYEDEIQDVIAR---YTKRL    VIYKQ--KLA  1775
Query_APLYS  1755  ---YNDYISAIDKLLVAM-EMTGSAQ   LLETVISVFCREP-R-HPHEDAIQAGIAR---FVQRL    TPDGL--ATA  1816
Query_SCALL  1753  ---YRDYITALDKILSGL-ELSGSLV   LLELLISITCREQ-K-HTYEDTIQQFLNR---FTKRL    PADKQ--KLA  1814
Query_CRASS  1723  ---FKDYIAAIDRLLTSL-ELSGSEM   LLELLISIFCREP-K-HVYETKILESINR---FTRRV    PAAKQ--KSM  1784
Query_OCTOP  1683  ---YDTYISLLDKILFAL-KKSKSKM   LLELVIEIYCQEE-K-HCHEKDIHNCLET---FIKRV    SLLDQ--KSY  1744
Query_POMAC  1728  ---FVDYISAIDKLLSAL-ELSSSLV   LLELLISIFCREEgR-HVHEDAIQQSLNT---FTARL    PAEKQ--KDA  1790
Query_WHIPW  1679  ---FEQYIIACERFCDAF-ASTGSLV   ILNLLLAVVCREE-Q-HPIVSYLHSSFHY---LVQSV    RPASV--NAA  1740
Query_TRICH  1687  ---YDQHVAACQKILFAF-GCTGSPI   ILELLLSMVCREQ-S-HPIASDLKHQLQK---FSSRA    NSAVL--ENI  1748
Query_AASTA  1580  ---FESFELLFHTYLSML-VQSGHVG   LLKTVFRSL-NEK-NqHVFSNDMYDMLTG---FCDDL    PGPRI--PSV  1641
Query_AEUTE  1538  ---FDAFEMLFQSYLFVL-VKSSQIV   LLKTLFQSL-KEK-KsHAFYSAIQDALES---FALEV    PMCRI--YPI  1599
Query_PLASM  1663  ---FDVFHMLFTEVLHVV-EQSKSIA   YLKVIYPSL-KEA-EsHLFRVEIKQMLAR---FSLSL[10]SSEMI--QRQ  1734
Query_BREMI  1649  ---FNTFHMLFLELLGVI-ERSKSVA   YLKIIYPSL-KEA-EkHVFYTEIKQMLAR---FCITL[10]HADNM--RLQ  1720
Query_OLIGA  1672  ---YEVFQLLMDKLLLVV-EECRSVE   LLKVLYSSL-KEG-SkHLFGKTIQSSLGR---LATNV[ 9]HDYMM-----  1739
Query_LIGNO  1788  ---YCQYVRAFRGLLLAQ-ETSEHPL   LLEILSILVCKER-Q-HALEAEFQQCLNR---LLLRI    PIQRC--RDF  1849
Query_CESTO  1571  ---LAQCSTLLKSVLSLLkTLSCPEA   -ILMLTMTFCRQD-D-YFMDEELEESLRIa---MKRI[ 5]-AQEQ----L  1634
Query_OVI2_  1932  ----EEYRKVFTTLLSCL-ELTGSCV   ILEALLIPFCRED--lHPMDVSLESALAAt---MKRH[ 5]---TQ--LAF  1994
Query_PRIAP  1742  ---YHDYVAAMSQLLDVM-EVSGSTM   LLRLVISIVCRES-K-HAMHDQLQTSLSR---FIHRL    SADKQ--KLA  1803
Query_ROTIF  1699  ---YKDYLNVVHKLLVCL-ELSQNLD   MLSLLIGVVCREK-E-HLCENQFRESLER---FIKRL    SSEKQ--MEI  1760
Query_SPONG  1743  ---LNEYTNIIEKLCNSL-VASGSLV   LLELIINIMCREV-R-HVCEEKIQTSLHQ---FISKL    -SEDG--VDA  1803
Query_TARDI  1747  ---LAFYLQAVGFVLAAV-EECRSIG   GIEICIKLVCRDA-H-HPLLERCRGMFEA---VVAND[ 1]-SENV--QKI  1808
Query_TRICH  1719  ---YNDYVACIDKLLANL-IKSGNLM   LLEVLISVMCREK-K-HAYEEQIQKSLLL---FISSL[ 2]K--DA--LGA  1780
Query_OSTRE  1690  ----RRFPSTINAFKQAF-AMCRSFR   LISALVPFYDKWL------EPCVKGAVTSwngALAQS[ 2]--------LI  1745
Query_NITEN  1814  ---FNNYVRLLEGLLRAL-VASLSLG   LLEELFPVLQFQG-N-TLRSRRIEEALGD---FAGET    ARKDP--KPL  1875
Query_CLUBM  1638  ---RTDYLGTLDLLIGAI-VSSRNLD   ALEVLFPLLQ-NP-T-GLTTPPLSEALRA---YA-QN    IGDKR--VAA  1697
Query_MOSS_  1735  ---RSNYTQLLNQLLAAV-SVSHSLE   LLEVIFPLVQ-TP-N-KISSRAVSEAVEV---FA-ES    LEDNR--EAA  1794
Query_LIVER  1760  ---RAMYTRMLSRLLEAV-SLSCSVT   LLEILFPLLL-DP-P-KFAVRTLFDSLEL---LS-QN    IGEER--TAC  1819
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1794   LDVPYKIFS---QEG-R--FTNNQRRAAC   EKVSLSLMRL   VHSSA   LTEFY   CE---HIKEIM-K-II---   1847
Query_LINGU  1796   LDVPFNIFT---KEG-L--FPNEIRRSTL   SRVCVPLLRL   CHQSS   LSEFY   VD---HIKFIT-D-TL---   1849
Query_LIMUL  1783   MDTLFGMFL---CEDQ---LPAEMKLGTI   ERVCVPILRL   LSKPV   ITDFF   VK---NIRSIM-E-IV---   1836
Query_SPIDE  1783   ADIVFSLIM---KSGFD--SKSDVLLKIV   NKVLILVMQN   ISLTS   FINFF   LS---HIKEIM-K-TV---   1837
Query_SCORP  1781   AEVPLKMFL---EDEL---YPSDMRIHII   EKVFRVILGM   ISKIT   LKYFF   ID---HIKEIM-Q-IV---   1834
Query_TICK_  1711   IEAALLTNQ---STNQG--LLISARLSLA   SKLLPLLLKS   CSKPA   LHRFF   VD---NVTAMV-D-AV---   1765
Query_MONOS  1701   LRLTFRVAV-----DLTk--PEAARNALA   HHVLPALLEG   CDAAV   RRDFY   RE---EITAIT-A-ML---   1753
Query_ROSET  1682   IELMFQIYL---AETNPl----FVRRGVA   QRVLATFLQA   CKLSM   CSAFY   IK---RLPAIM---SL---   1733
Query_BELCH  1824   LDIPYGIFR---REG-S--FPNEIRRSAL   ERVCLPMLRL   VHPAA   LTEFF   TD---HIREIM-D-II---   1877
Query_CIONA  1793   LEIPYTIFT---TSTQR--S-IPAAQTAI   HN-VLVPMAQ   CCNVS   VIQFF   QT---NISHIM-S-VV---   1845
Query_CAECIL 1811   LETVHDMFK---TENLW--S-NTTRQTIV   DKSLLILIWH   CSLDA   LKEFF   CK---IIIHTM-D-TL---   1864
Query_HUMAN  1798   LESVYEMFR---KDD-P--RLSFTRQSFV   DRSLLTLLWH   CSLDA   LREFF   ST---IVVDAI-D-VL---   1851
Query_MOUSE  1795   LESVYTMFR---KAD-L--PSNVTRQAFV   DRSLLTLLWH   CDLDT   LKEFF   SR---IVVDAI-D-VL---   1848
Query_XENOPU 1820   LNTLHNMFK---SESLM--L-NGTLQSLI   DRCLLTLLWN   CSLDA   MISFF   TN---IISLAM-D-TL---   1873
Query_PLATY  1802   LEAVHGMFL---RDDHL--S-STTRQAFV   DRSLLTLLLH   CSLDA   LKEFF   YR---IIVGAM-D-ML---   1855
Query_TASDE  1805   LDTVHRMFQ---KDD-L--S-NIPRLACV   DRCLLTLLWH   CSLAA   LREFF   SS---IIVEAL-D-VL---   1857
Query_CANAR  1800   LDTVHKMFQ---SEELL--S-NAARQAFV   DRSLLTLLWH   CSLDA   LRVFF   GK---IIVEAM-D-TL---   1853
Query_OSTRIC 1770   LDTVHKMFQ---SEDLR--S-NVTRQAFV   DRSLLTLLLH   CSLDA   LREFF   GT---IIVQAM-D-TL---   1823
Query_TURTLE 1809   LDTVHKMFQ---SEALL--S-NATRQAFV   DRSFLTLLLH   CSLDA   LKDFF   SK---IVLEAM-D-TL---   1862
Query_GECKO  1809   LDTVHKMFQ---SETLL--S-SATRQAYV   DRSLLTLLLH   CSTDA   LDDFF   CK---IIVEAM-D-SL---   1862
Query_DANIO  1809   LSSTYQSFQ---AKEVP--S-NFMLMGLI   DRVLLPLASH   CSPQA   LSQFF   IS---NIADIM-T-TL---   1862
Query_SNAKE  1808   LDTVHQMFQ---SETLH--S-NITRQAYV   DRCLLILLLH   CSLDA   LKEFL   SK---IIIEAM-D-TL---   1861
Query_STERL  1821   LDTVYQMFQ---NDELH--S-NTARQAVL   DRSLLTLTLH   CSSDA   LKEFF   CT---TVGDIM-G-TL---   1874
Query_MILII  1805   LEAVYGMFE---NEGGL--S-NIARQSVM   DRSLLTLITH   CSLDA   LKEFI   SK---IIKDAM-G-IL---   1858
Query_BAMBO  1804   LEAAYEMFR---NEEML--P-NATRQFAM   DRSLLTLIGH   CSPDA   LREFF   CK---IVKDAM-E-IL---   1857
Query_ALLIG  1767   LDTVYKMFQ---SESLL--S-NVTRQAFL   DRSLLTLLLY   CSLDA   LKEFF   SK---IVVEAM-D-TL---   1820
Query_GAR__  1809   LETVYEMFQ---AGELA--S-GAMRQAVL   DRVLLTLVLH   CSSDA   LRGFF   CK---TIASVM-E-TL---   1862
Query_TETRA  1768   MEWALDQYK---DKDLDrsLLECIRYEII   QKVLLHLIEN   AKEEA[1]-VKFM   VKnspYLEKLFyN------   1826
Query_PARAM  1672   --LKWLMLQik-DEQLDtsTLNNVRIALG   NKLLIPLLSN   --SNL   LLEFY   SE---FTPYLL-Q-QL---   1726
Query_SLEMN  1758   FSQFMDEFM---NRQIDlsVDDNLRWAIM   KKIILKILES   CSQEF   LADIM   IE---FTPQLE-I-VL--k   1815
Query_STENT  1544   FQTLIEMFY---DPNTDkgIIDNLRWGIA   KRILIPMIDY   ADELM   IEKLM   IY---CTQQMI-G-KI--t   1601
Query_NEMAT  1793   VDVTYRIFT---N-ERD--HTSNVRLAVL   DRVSVTLLCN   ASKAG   VSEFF   SS---HVKEIM-E-II---   1846
Query_PISTI  1821   MDVCLSIFF---N-ERD--YQTHSRRAVI   ERICITFLLK   TSTTA   RRAFY   QD---QIQKLM-S-VL---   1874
Query_CORAL  1799   LDICFDVFT---KKDTE--YRNLTRRAAI   DKVCLVLLRA   SSTSV   FREFY   VD---HIKNIM-E-II---   1853
Query_HYDRA  1720   LDLSVKMFN---DESKY--A-DDIRAVII   DRVCIPLLNE   VSWIS   LKQFF   QR---NILVFT-S-HI---   1773
Query_LATIM  1809   LDTVYKMFQ---DEGLL--S-NTTRQAIM   DRTFLTIMLN   CSIDA   LKEFL   CK---NILDVM-G-TL---   1862
Query_SHRIM  1773   LKHIYDMYN---DPTCI--KQNL-RYNIM   DLVLTPLLQQ   AEAVP   TQNFL   LT---VISAVM-A-GM---   1826
Query_MCRAB  1768   LRFVYDMYA---NQSHL--KQNL-RYNIM   EKVLIPLLQL   SDLTV   TVSFS   ID---VIKPVM-A-GM---   1821
Query_AMPHI  1791   ASRVFQVVE---AQEAS--GLQLHRQLLN   ERVLLSLLRT   ADVVS   VERFF   ER---HLKSIM-N-LL---   1845
Query_BARNA  1672   AEELLALFWp--AGGLPa----AGRRTLL   VSFLCPLLRA   ADRPV   LLQFY   TShvvAIARPVqD-PE---   1729
Query_COPEP  1691   ITVPYTIFI---DDSGLf--SAQVRLSAV   SRFAVPSVQT[3]---AT   IRRFF   SE---NMESLL-L-NV---   1745
Query_DAPHN  1617   LNRIYHHAK---TRSINw----KTRFEAI   ETLYTSLLKA   --LPS   --SFV   -Er--HVNELLgD-II---   1666
Query_STARF  1830   FGVAYQIFC---QEN-S--FSKEIRRAAI   ERVCIPLMRL   AGVKA   LSSFF   KE---CIKNIG-G-II---   1883
Query_STRON  1819   ANIPFEIFK---QDK-R--YSKEIRKAAI   QRVCIPLLRQ   ASRDA   TVTFY   KG---HIHDII-S-VL---   1872
Query_SEACU  1794   IGTCLKIFQ---AEG-E--YTTQMRRAVM   EKVCLPMFRL   CQVDS   VIEVF   QR---NIRDFV-D-IL---   1847
Query_FUNGI  1740   FASCFSHFI---MESLQ----DQLRLHLV   DLFCIPLIKR[9]MVGQL   FAERI   ST---IIGILN-H-TI---   1801
Query_DPURP  1890   FSHCFELFL-----NHN----DELKITLI   DKFCVPLISH   MKEEI   LVQIF   SH---HLNSLM-S-II---   1940
Query_HALBU  1891   FSHCFAMFM---SNDYQ----DELKITLI   DKFCVPLIGH   MKQQV   LIQLF   AQ---HLSSLM-S-II---   1943
Query_GLOBO  1662   TKYGFEFSV---QSKNS----HDLRWNAA   KLLLLVTLQS   CQLDN   LVRFF   VQ---YIREIM-D-NL---   1714
Query_PALUS  1714   TSTLLGFFF---DSSLP----MEVRGNVV   LHMLVPALSV   APKAY   VRAAF   VR---HINQIM---TT---   1765
Query_CALIF  1791   TDYCMEVLL---NNKAL----PLVRVNVI   KQYLIPALIH   VNCQF   KIKFY   SK---HIKSLT---KV---   1842
Query_GLOMU  1846   MDIAYGYFT---NSTYV----S-FQRIII   EMILSPPLLQ   GPKIF   VMDFF   KK---HINEII-D-LL---   1897
Query_LOBOS  1769   TGYCFDEFI---KRTHS----DEHRRNII   HQILLPMLKA   LPTLS   VSEFY   VL---NIARIM-G-VI---   1821
Query_RACEM  1748   TETCLAYYK---DDGNR----IEVRRNVI   SLLVSLFIL-   VPDDY   VVDFY   LD---NILFIMtG-AT---   1800
Query_PVIOL  1884   FSHCFQIFL---SQDYQ----DELKITLI   DKFCVPLIHH   MKETV   LVQLF   CQ---NLTSFM-S-II---   1936
Query_BOMBU  1669   LQAAYKLFM-----NLT---NE--RFDVL   QRFLLPSFEF   CKTTE   INKFF   ET---NIKEIY-T------   1716
Query_AEDES  1788   KQI-FTKFR---EESLE---PNI-RLTML   KRYLVTSLRV   CNVET   IYTFY   KL---YIKNIN-D-MI---   1839
Query_ANT__  1692   LQLTYDAFM---ETNTT---SETERLDIL   REFLLPAFKF   CNAIA   IERFF   KK---NILDLK-K------   1743
Query_WHITE  1722   LNVMFDCAV---GLPEL---ENY-RENIL   KTYLEPTLRY   CCIKS[4]LKVFY   EG---CIENLM-T-TV---   1778
Query_BEETL  1665   VKLYDLCLL---ED--L---DNF-QKYKL   IIDTSSVFQY   CNVTS   FEDFF   TE---TISTVL-T-VL---   1715
Query_THRIP  1686   LKLVFSMFE---NETIS---VEY-RAKIL   NEFLLKLMQF   ASLSS   QVSFS   VS---TISKLV-K-RI---   1738
Query_TERMI  1755   IEAVFKVFV---DDQYD---PSV-RLGCA   QRFVRPLLQG   CRRSA   AREFY   RH---CIQDLQ-G-ML---   1807
Query_FLEA_  1685   LRIPYFQIK---SREYT---PSI-RVSIL   KRVLGTILTS   CDQQI   IELFY   KN---FFEEDF-NnLI---   1738
Query_MIDGE  1748   NQT-FDEFS---TNKCE---PEI-RLKFV   SRFLLTIMKN   SHVNV   MKSF-   ------IRQKM-N-LI---   1795
Query_LAODE  1722   LDELYKFRN----DKILk---AHSRVKLV   QDFMLPIIES   TSFKQ   AEEFF   LK---HIKDIV-V-SL-kd   1776
Query_LOTTI  1776   LDIPYKIFT---SSG-S--YPSAIRIAAL   SRVCLPMLRL   VHKSA   VIEFF   LT---HISFIM-S-EI---   1829
Query_APLYS  1817   LNVPYAIFS---KEQ-S--YFHEIRWAAV   ERVCLPMLHL   SPRVT   VIEFF   SS---HIHELI-N-NI---   1870
Query_SCALL  1815   VDVAYGVFS---KES----YPAEIRTATL   HKVALPMLHL   CHKSA   LIEFY   LD---HMKELR-S-IT---   1867
Query_CRASS  1785   VDVAYNIFC---QEK-L--YPGEVRRTTI   EKVALPILRL   VHKSA   LLEFF   TD---HMGDLR-T-TV---   1838
Query_OCTOP  1745   VQIPFDLFL---KTQS---CTNNFRKFVI   KKVCLPFLRN   MSKIP   FVEFL   KS---NIPAIV-N-IV---   1798
Query_POMAC  1791   LDVPYKIVF---CES-G--FQDMIRRTTL   ERVCLPILRC   VHKAA   LIEFF   LD---HILDLM-R-IV---   1844
Query_WHIPW  1741   ADKLYSTFL---DASSCp--TEDYRSSVF   SVFLLPLLQC   MDSAA   LGHFA   CS---HVSSIV-S-TL---   1795
Query_TRICH  1749   LELLDKMFF---NPSGCv--CASLRLALL   NEFFIPFLHS   LKNDL   LLTWT   CR---KIQTLV-S-CL---   1803
Query_AASTA  1642   CLEVLPLVF---SPSLTd----HIRTTLL   RQVFVPLMHR   LDEAA   IVAVY   TS------TFS-G-AP---   1691
Query_AEUTE  1600   CVELLQLLF---DFGLDd----YIRTVLH   RQVFSRLVHQ   LEEAA   IIKLF   TM------TSF-G-VP---   1649
Query_PLASM  1735   MAELLDLLL---DPNVDv----AIRKILL   EEVFTPFVEC   QTGEI   LLHFY[8]ST-------II-S-LL---   1791
Query_BREMI  1721   LAELLEVLL---DSTLDv----SIRRVLL   DEVFTPLLEC   QTGEV   MLQFY[8]LS-------VI-S-VL---   1777
Query_OLIGA  1740   --ELLEVLL---DPVGDl----TVRRMLL   ENLFVPLAEL   QKADN   LREFY[4]KTsk---VTMI-S-KL---   1794
Query_LIGNO  1850   VTWPWRLFR---GDSGGggFTGFHREQLV   SRVLAPLLTQ   CHGSV   LLDFV   ATddc-LEYCL-T-CLlna   1911
Query_CESTO  1635   LASSINYFEqacPKSNPspGFISLWLRYV   HKFLKPLLLS   VSPSA   LERLA   AQs---ITKWM-Q-TL---   1694
Query_OVI2_  1995   LDLCHEQLIpsmHGEKAh---QPVALDFV[4]EKFFEPLLLH   VSHTS   -LEIF[1]VN---HVNDWL-G-VL--r   2056
Query_PRIAP  1804   ADSGYSILA---EEG-N--FPGEIRRSCL   DNVVLLLLRQ   CCHHA   VVEFF   RC---HLVEII-G-II---   1857
Query_ROTIF  1761   VKQNWSNFQ---SSD------DERKFLLF   EKVIFIFLKN   CDKPV   FIDFM   SA---NIVSLI-E-VL---   1811
Query_SPONG  1804   LKFVYEMFK---NDK-D--FPVAAVKGIG   ERLCVPLFTE   CSSAV   LLEFF   VL---KISDVM-C-IM---   1857
Query_TARDI  1809   FEIPWKIFT---TPQ----YGPDHRRTAV   EEILLPMLQA   ATSTA   RMDFF   SR---TMKDMI-T-IL---   1861
Query_TRICH  1781   LDVPFKMFV---D-DKS--YRPDIRRAAI   ERVCLPMLQK   ASLKI   TKDFY   VS---HIKEIM-D-II---   1834
Query_OSTRE  1746   VADVYADFR---TQSLA-----------A   SIILPTLLEA   ANAQE   IVEFF   MEnh---ETIMkD--Vgg-   1793
Query_NITEN  1876   LELSLTLLH---DGEKQ----PRLKKAVM   DMVAVPALYA   ADESF   LSEWF   AD---KFEDFK-R-TL---   1928
Query_CLUBM  1698   FSYCMDVLK---DATKL----PRLRKAVL   DYFFVPLICL   SPVDT   LTGWF   TP---RVTELM-D-SL---   1750
Query_MOSS_  1795   FNLCLTMIM---DRGKS----YELRRAVV   EVIFEAFVNS   SSMEF   VTVWY   SQ---HISDLF-Y-VL---   1847
Query_LIVER  1820   LNLCIQLLG---DSTKL----PQLRRAVI   DRVFVSLMNA   ASPLF   VADWY   SK---RIKVIV-D-AL---   1872
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1848   EAKPTK--APEHAFEAQ---LTSKIGCFELLEVMYSRLPNTDVNSKDSSI   NKSFCSPK[1]PD-T   GKE-MNLTI  1912
Query_LINGU  1850   SAKLSK--SSESEIERL---LTLKLCCFELVEVMYSSFPKEEVNSKNSKI   NLKYCGGK   VE-T   GKE-MTTAV  1913
Query_LIMUL  1837   SSKVIK--GSEERIKAQ---LLIKIGAFQLTEVLFSCLPKEIISSKESLI   NKAFCDGN   VD-T   GKE-LTTEL  1900
Query_SPIDE  1838   KENLSL--ASG-VKDIH---LVKRICCFKLLELMYSTLNKDLLHSKESQI   VETYLGRK   PE-V   GNE-LSTDI  1900
Query_SCORP  1835   KSKRKS--SNE----IQ---LVSKIGAFQLIEILYTYLSKEEVS-KESQI   NCSY----   ----   GKE-LTREA  1886
Query_TICK_  1766   AAKIRH--GSE----SE---LTSKIVALSCLEVLYSCCPKDDVTGKQSSV   NAAFCRSR[3]DV-V   GNE-LTKEI  1828
Query_MONOS  1754   SSKHHS-dTAGWSREVA----------CQLLAQMFRLLPVHEVYGTRAAV   NQAFHDQA[4]AV-A   DNC-LTKAV  1815
Query_ROSET  1734   MSQDTP------LTSTP--aIVIKMCTFTLIDILFQRLNTTDLYGKSAVI   NKTYCEAA[3]-VnD[1]GSN-LTTNT  1798
Query_BELCH  1878   GAPQTK--AGEASLRTQ---LTNKLCCFQLLETMYSRLPKEEVNTPPSRI   VRAYKGEE   VK-A   GNE-MTTAI  1941
Query_CIONA  1846   QEEHIK--YPESKWSYQ---LQMKCLAYRLLELMFALLPEDQTKPS-SPV   NHSFTSSQ   LT-Q[1]GNS-LIKTF  1909
Query_CAECIL 1865   KGRFTK--SNELAFDAQ---IMMKMSYYKMLEVMYSRLTKEDIYSKGSRI   NQAFHGSS   NA-E   GNE-LTKIL  1928
Query_HUMAN  1852   KSRFTK--LNESTFDTQ---ITKKMGYYKILDVMYSRLPKDDVHAKESKI   NQVFHGSC   IT-E   GNE-LTKTL  1915
Query_MOUSE  1849   KSRFTK--LNEFTFDTQ---ITKKMCYYKMLAVMYSRLLKDDVHSKEAKI   NQAFHGSR   VA-E   GNE-LTKTL  1912
Query_XENOPUS1874   KSRFTK--VPEAAFDSQ---ITKKWGYYKMLEVQYSRLSKDEIY---STV   NNAYHVSS   KP-E   GNE-LTKAL  1934
Query_PLATY  1856   KSRFTK--SNESTFDTQ---VTKKLGYYKMLEMMYSRLPKDDVHSKESRI   NRAFHGSS   VI-E   GNE-LTKTL  1919
Query_TASDE  1858   KSRFTK--LSESTFDTQ---VTKKLGYYKMLEVLYSRLPKDDVHSKESRI   NQVFHGSS   IV-E   GNE-LTKTL  1921
Query_CANAR  1854   NSRFTK--SHEHTFDTQ---VTKKMGYYKLLEVMYVRLSKEEVYSKDSKI   NQAYRGSM   SV-E   GNE-LTKTL  1917
Query_OSTRIC 1824   NSRFTK--SNESTFDTQ---VTKKTGYYKVLEVMYVRLSKEDVYSKDSKI   NQAYCGSI   SV-E   GNE-LTKTL  1887
Query_TURTLE 1863   KTRFTK--SNESAFDTQ---VTKKMGYYKILEVMYSRLSKEDIYSKDSRI   NQAYQGST   NV-E   GNE-LTKTL  1926
Query_GECKO  1863   TSRFTK--SNETMFDTQ---LTKKMSYYKILEVMYSRLSKEDVYSKDSRI   NQAFLRST   CV-E   GNE-LTKTL  1926
Query_DANIO  1863   QTRFTK--SVESVFESQ---IMMKIGCCKLLEVLYSRLPKEEVYSKNSAI   NQAFCG-T   GCaE   GNE-LSKNL  1926
Query_SNAKE  1862   KSRFTK--SNETSFETQ---LIKKISYYKILEVMYSRLSKEDVHSKDSRI   NQVFQRST   HV-E   GNE-LTKLL  1925
Query_STERL  1875   QARFTK--SSESTFETQ---LTKKNGCYKLLEVLYSRLPKEDVYSKESLI   NQAFCRPN   KA-E   GNE-LSKSL  1938
Query_MILII  1859   QSRFTK--TSESSFDNQ---LTRKIGCYKLLEVTYIRLPKDDIYSEDSKI   NRAFHSSS   NM-E   GNKlLSKSL  1923
Query_BAMBO  1858   QSRFTK--TNESSFDNQ---LTKKIGCYKLFEVMYTRLPKNDVYTDGSKI   IQAFHSAG   NI-E   GNK-LSKTL  1921
Query_ALLIG  1821   KSRFTK--SNETVFNTQ---VIKKMGYYKILEVLYSRVSKEDVHSKDSQI   NQAYQGTT   NV-E   GNE-LTKML  1884
Query_GAR__  1863   QARFTK--LNEAMFEGQ---LIKKNGCYKLLEVLYSRLPKEEVYSKESKI   NQGFCGPS   RAaE   GNE-LSKTL  1927
Query_TETRA  1827   -TDFMK---QENMKEVL-rwVQELIQVFNIFESLYKALSLDKIKDDVHKK[2]--------[2]TT-Q   GNE-LTKSL  1886
Query_PARAM  1727   NVTWDD--NWKNMLFQL----QWRGLIFDLFGIAYKRMSTEQIK------   -KQLHQNL[2]-Q-A   DNE-ITKKL  1783
Query_SLEMN  1816   ESIFLQ---KDPLEQY--qlIKQKKWIFLVFEVMFRRLKPETIR---QQV   HKKLYGES[1]-P--   GNE-LTKSL  1875
Query_STENT  1602   EIKFTE---IYDSKEYLfrlREKTWILLI-FERFFAKLPSNRLK---ENV   HKALYGPG[1]-Q--   GNE-LTKQI  1662
Query_NEMAT  1847   EAKQLK--TYDPGFESQ---LVSKMCAFRMIEVLYSRLGKAELNSPQSSI   NLAYCPND   N--T   GKE-LTKAI  1909
Query_PISTI  1875   EAKELK--PTDPNFESQ---VTSKLCCFKLVEILYSQFSKSELNTLESTI   NKAYVGGD   VK-T   GKE-LTLAI  1938
Query_CORAL  1854   ETRFSK--ATDSVFESQ---LVSKLCCFKLLALLYARLEKTDVSSMESRI   NRAYNNDV   K--T   GKE-LTMAI  1916
Query_HYDRA  1774   DEKFNK--LLQSKLASQ---LINKICSFNVLSLMYELLGKEELTSATSTI   VQTFCPNS   ST--   GKE-LTVFL  1836
Query_LATIM  1863   QARFTK--SNESAFDTQ---VTKKTGCYKLLEVMCSRLPKDDIYSEQSRI   NQAFHGSC   SA-Q   RNG-LSKTL  1926
Query_SHRIM  1827   SAPVHG--RSDEQ-HKT---LVTKVGSFMILQIMYSKLPREKLFAPGSEI   NAAFRPG-   DQ-S   GKE-LTKEI  1888
Query_MCRAB  1822   SLTIQG--RSDEQ-QRL---LVTKIGSFKILQVIFSRLPRERIFNPGSEI   NAAYQPS-   DT-S   GKE-LTKDV  1883
Query_AMPHI  1846   SCSLPS--RPEQQ-LCV---LTGKLAALRLLQLLYAKLSRDQCFSSASAV   NAVVNPA-   DA-S   GKE-LTKEL  1907
Query_BARNA  1730   PPAVPA---DSEPTEHQ---WVSRMSGFSLLEVLYASLPKDQTHAATSPL   VTAFAGA-   TA-R   GNE-LTKAL  1791
Query_COPEP  1746   GVEVLG---SEEKRASV---LISKIGTFQLFGVLYSRLDQSTLHKECGII[4]KAVFKKHS[4]---Q[2]GKD-LAIYL  1816
Query_DAPHN  1667   --QICE----EARGGVE--aFVHKIVVCALFEVAIVRVDYQNLTGHLNAN   ---FCALR[3]QPsE[1]GRD-LLKYI  1729
Query_STARF  1884   DAKTIK--APETVFQNQ---LVSKTCCFQLVELMYSRLVKEDLFSMDSAI   NKAY-TND   VK-T   GKE-MTQAL  1946
Query_STRON  1873   EARLTK--NPVPAFENQ---LVSKICYFELIEVLYSRLSKDDLNSKDSAL   NREYRHGN   VE-K   GNE-MSTAL  1936
Query_SEACU  1848   ENKLKK--ISEAVFAHQ---LVSKICCFELIELLYTRLGKDQLNSTTSVI   NSAYCMGD   CK-T   GKE-MTQGI  1911
Query_FUNGI  1802   NTSDDE--ETKKTWMTK-----------KISYIAYRILDPQTI---REIM   NRRFSGE-[1]-A-K   GNE-LTASV  1853
Query_DPURP  1941   QPLQPK--YLSDSQERK-ssIIEKICCFHLIEALYQSLSTNII---KEKI   NPYFY-D-[2]DS-K   GTE-LTAAI  2003
Query_HALBU  1944   QPLAPK--YLATAIERK-ssVVEKICCFKLTQQFYESLPATVI---RDQV   NPYFTGD-[2]DA-K   GTE-LTAAI  2007
Query_GLOBO  1715   AAAAPI--GDESALLQY---LVTKAICFMLLEVAYKRLPPSELHSAKGRI   LQAYVGP-[1]KS-G   EKE-ITIAL  1778
Query_PALUS  1766   LSRPLS--TKEHDQESQ---LLTRTACYELMDVCYSRLDTHDVHSTAGNI   VKAYCNG-[1]PK-T   GKE-LSVDL  1829
Query_CALIF  1843   IKEQLI--NKEDELCIQ---LNIHSSIFNILQIMYEYLPNTEVNGANSPI   NKAFCDG-[1]VK-N   GKE-LLTNL  1906
Query_GLOMU  1898   ERDKYT--RNVDEKIYY---LNERAGSLTLLQILYLRASEEDVYSEKGEI   VKAYCQS-[1]PL-K   AKE-LTRKV  1961
Query_LOBOS  1822   KQETPR--STDTEMRRD---YIERECCYGLVHVLYMRLPSEMVNTKDSKI   VDAFTGH-[1]TT-T   GKE-LTVDV  1885
Query_RACEM  1801   RQLLRR--GTDEEIRVD---MEEKTCCIYLMQTFYLRVRHGKQFGMRSEV   SKKWEQET[7]DA-K   GKI-MMTKL  1871
Query_PVIOL  1937   QPLTPK--YLSDSQERK-ssIIEKICCFRLVEALYQSLSSTII---KDTV   NPYFY-D-[2]DA-K   GTE-LTAAI  1999
Query_BOMBU  1717   VICQSL-vGGTSDIKQL---IVSKIGCYDLVAIMFAKLQIWDIDNAESV-   --ITRNAI[2]VV-T   GKE-LLHSL  1780
Query_AEDES  1840   RSNYGQ-fGSGWEAEHA---LVNRLGGYQLIELYVAILQRDYLLTDDCIV[1]KALYGESG[1]-I-P   GNR-LITDL  1905
Query_ANT__  1744   MMTKNI-dCNDDAIKQI---IVSKIGCFQLIAIMFARVDENKID-ANGT-   --IAQNEL[2]IL-I   DRN-HYREL  1806
Query_WHITE  1779   NAEFTL-sAKENEINTE---IVSKIVALTLLEILFGYFDRSDFDTE----   ---WFKTA[2]VS-S   GQE-FLQKL  1838
Query_BEETL  1716   KEPVTT-eN-----------LSNKILHFLLVGVLFLRCPLQSDDSE----[1]SRITQAAL[2]KP-N   NKT-LLLQL  1771
Query_THRIP  1739   TEGFSM-kTDVGKQKHK---LMSHIGSWQLLQHIYGFSMKSLLEEKGSEL[1]QAAFSGSE[2]VT-S   GKE-LYTLI  1806
Query_TERMI  1808   NVTLDC-rGQLQTE-HQ---LVGRIGAWNILEPLFVHLESSELESK-ECV[1]TAAAFPGA[1]-K-T   GKE-LISDL  1871
Query_FLEA_  1739   ESPYTS-aSRAWDLKQA---YVDRTCCFHLITILVYLSDPS-IFDDASNP[1]LRSFPDFK[1]-E-N   PRE-VLSTL  1803
Query_MIDGE  1796   WSLIEA-p-LNVDMENG---FVNRCAGYMIIEAFAATVPCSIIEKESFT-   ---YGGKL[1]-V-K   GTE-LIKDF  1855
Query_LAODE  1777   ETNQNK-kINSTNLETL---VPPRAASMLLINALFSKFPSSELTSIMCPI   VQKSEEIG[2]---K[1]GND--TAVL  1841
Query_LOTTI  1830   EAKLVK--GSKAELENQ---LTTKIGCFQLMELLYSRLTKDEVNSKNSVI   NEKYCQTK   VE-T   GKE-MTQKA  1893
Query_APLYS  1871   DCKLNK--SSESVFAHQ---LTTKSGSFQLLEVMYSSLSKDDVNSKTSAV   NLRFCHGK   VD-T   GKE-MTQRI  1934
Query_SCALL  1868   EAKQVK--GPG--LENQ---LVSKTGCFQLIDIMYSHLAKENVSSKNSVL   NEKFCQIK[3]IK-T   GSE-MTRVI  1932
Query_CRASS  1839   EAKLVK--GSG--LENQ---LTSKLCCFQLLELMYGRLSKEEVFGKSSVI   NKKFCEVK[3]VD-T   GKE-MTQVI  1903
Query_OCTOP  1799   EAEMSQ--LPD-LMEQH---LTDKMCCFQILEVMYRSMDVKELNGPGSEI   NSVYCKAR   NI-E[5]GKE-LTKVI  1866
Query_POMAC  1845   EQNLAK--GPESILKTQ---LTSKVCCFQLLEVMYSRLSREEVNSLSSEI   NKKVCFGK[1]-E-T   GKE-MTQKI  1908
Query_WHIPW  1796   SETDLN--PIHQLYEKQ---LVTKTGCYNICSLLYDLLSLDQLK---GPV   NEAFCHGK   SV-D   KTE-LIKKM  1856
Query_TRICH  1804   TESLTL--EIQTDCEQQ---LITKIGCYKICAIIYNELDLECLK---NEV   NVSYLGSD   SQ-M   KNE-LIKKL  1864
Query_AASTA  1692   VVQYLM--TVVTSDTSP---SLAVVVTFGLLELLYTCLSGDAVR---RLV   NPIYAATA   GA-K   GNE-LTMRL  1752
Query_AEUTE  1650   FVKYLM--NIVLNDQSP---LISLVMAFGSLELLYSCLSGDAIR---HSV   NPVY-ADA   PA-K   GNE-LTMRL  1709
Query_PLASM  1792   ATLISA--SAEGAFEES---RIGTVVAFSLVEILYRLVDPDIIR---SKI   NSAFLGHK   SG-K   GRE-LTMLV  1852
Query_BREMI  1778   ATLISA--SADVQSRGC---QIGVYVAYSLVDILYRLVDGELIR---TDI   NAAFLGHT   NG-K   GRE-FTMLV  1838
Query_OLIGA  1795   ATVVSN--SSDVASQGA---ISTAAVAYTLVELLYRLVDPESIR---GEI   NRAFLGHD   NG-K   GRE-LTMLV  1855
Query_LIGNO  1912   QQSLAG--FNESETEEI---LLLARGCCRVLSCFYMACGLTELALPRLCD[2]-KRFYSRQ[1]-AaQ[3]-DRsLSNWV  1980
Query_CESTO  1695   EDNASEd--ASHVITQWlwrLLNRALTFYLLVTIYNRLPKAVLHGLTAGG[4]LRAFLGPS[7]GKiK   GTE-LTAAI  1773
Query_OVI2_  2057   QLDISSd-----GTDAL-hdLFVGTLCFNLFTVLYNRLPKVLVHDKDAKL[4]IQADHGGQ[3]EY-K   GVE-LTRSL  2126
Query_PRIAP  1858   EAKVVK--LPERQYESQ---LISKICSLKLLELMYACLPKEDVYGKDAAL   GQVFLASR   PA-G   SKE-LTPVI  1921
Query_ROTIF  1812   N---WN--LNEPAFEMM---CLNKKCVFQLIELAYKRLHKDEIFFTNAKI   CMAYETIK[3]VK-D   GKE-LTKEV  1875
Query_SPONG  1858   EAKLAR--VTDPKLECQ---LVHKSCCFQFIEILYTRLPSSLISSMESTV   NTAYCNGS   PQ-T   GKE-LTQAI  1921
Query_TARDI  1862   DTK-ER--STTAGL------LIVQTCCFSLLELLCVTLSQGDLTSPTSKV   VLAFCAIR[3]AV-T   GKE--LIQF  1923
Query_TRICH  1835   ETKLVK--T--ADLENQ---LISKLGSYKLMKLMYSKLSKQDVNTAESSI   VKAYTGRP   TE-D   GKE-ITKSI  1896
Query_OSTRE  1794   ----ANg-------------LLCTLRSLNVLLSVYQKCDKQQLIDGPRSK[2]--------   ---K[1]-NAlVSKSL  1838
Query_NITEN  1929   GKEYSY--RTRTDAETEqteLTTRTAILSLLELMYLRLSAEQL---REKC[1]VDKQK---   NV-E   MIR-LCDDV  1990
Query_CLUBM  1751   GRTVMS--PDMETEQEL---LLAKIYSYTLLELLFTQVDKERI---T---   ---FS---   YQ-Q   LMK-FASND  1802
Query_MOSS_  1848   KQSSSM--LSAEREFED---LFKKICHYELLELLYSKADNVSI---TKTI   APVFT---   SS-D   LMR-QAGDT  1905
Query_LIVER  1873   GKPVNS--IDAETEEEC---LVVKVCHYSLLELLFLKVDAPLI---KEKI   TPSAS---   NA-S   LMK-VTMSE  1930
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1913  TKVSNATK   GE   NMNGETL   LIP   LRRALHCAAYNTLIAALICTQKE   K    KFYTAFLFTENPAKNQ     1972
Query_LINGU  1914  TKAAHEAK   KE   DVRGEKR   LPE   LRRQYHCAAYNALMAIITCTQTE   M    KFYTGLLFKEDLAKGD     1973
Query_LIMUL  1901  TRQSHATK   KD   DLRGETT   LYD   IRRRLSCAAYNVIIAVISCTQSD   P    KFYNVFLFQENSAKRE     1960
Query_SPIDE  1901  TKLTLEFK   KF   SYETEPE   YKV   LLRKLNCAKYCALIAVVSCTQNE   L    KFYNAFLFKENPAKNE     1960
Query_SCORP  1887  TGFAIAAR   KE   NLEGESI   SKE   FCRKYCCAAYNAVIAIICCTQTD   S    KFYSTFLFQENFAKGE     1946
Query_TICK_  1829  TKHANSAK   KE   VVRRDVE   FAG   LWRQLRCAAYRLVVAVVSCTQTE   S    QFYDVFVFQERPEKGE     1888
Query_MONOS  1816  VDSLRGFL   QD   KTTQIPG   LTN[1]-PQTLKQKAFNTLFHVLDCTQTD[4]L    KFFTTLCFESKKGPSP     1879
Query_ROSET  1799  CRALRDTV   --   --TGEES   LVE[1]-RRQASCAAYNALTTVICKTQTQ   E    KFYSAFCFKEATGKVR     1854
Query_BELCH  1942  TKFGHAAK   SE   DLRGETG   LLD   LRREFHCAAYNALIAIISCTQTD   M    KFYTAFLFQENTAKGQ     2001
Query_CIONA  1910  MLRVNRDR   NE   NMKGEKV   LLE   ERRQFHCGVYNALIAAVISTQSS   I    KWFKVFLFDENPSKQQ     1969
Query_CAECI  1929  VKSCYDSF   SE   NMAGETQ   LFE   KRRQYHCAAYNCAIAVISRAFTD   A    KFYQGFLFTEKTEKNL     1988
Query_HUMAN  1916  IKLCYDAF   TE   NMAGENQ   LLE   RRRLYHCAAYNCAISVICCVFNE   L    KFYQGFLFSEKPEKNL     1975
Query_MOUSE  1913  LKLCHDAF   TE   NMVGESQ   LLE   KRRLYHCAAYNCAISLISCVFNE   L    KFYQGFLFNEKPEKNL     1972
Query_XENOP  1935  IKLCYDTF   TE   NMCGETQ   LLE   KRRQYHCAAYNCAISLISCVFSE   L    KFYQGFLFTEKKEKNL     1994
Query_PLATY  1920  IKSCYDAF   TE   NMAGETQ   LLE   KRRQYHCAAYNCAIAVISCVFNE   T    KFYQGFLFTEKPEKNL     1979
Query_TASDE  1922  IKLCYDAF   TE   NMAGETQ   LLE   KRRLYHCAAYNCAIAILCRVFSE   M    KFYQGFLFTEKPEKNL     1981
Query_CANAR  1918  IKSCYDAF   TE   NMAGESQ   LLE   KRRQYHCAAYNCAIAVISCVFTE   S    KFYQGFLFSEKSEKNL     1977
Query_OSTRI  1888  IKSCYDAF   TE   NMAGESQ   LLE   KRRQYHCAAYNCAIAVISCVFTE   S    KFYQGFLFTEKTEKNL     1947
Query_TURTL  1927  IKSCYDAF   TE   NMAGESQ   LLE   KRRQYHCAAYNCAIAVISCVFTE   S    KFYQGFLFTEKPEKNL     1986
Query_GECKO  1927  IKSCYDAF   TE   NMTGETQ   LLE   KRRQYHCAAYNCAIAVISCVFNE   S    KFYQGFLFTEKPDKNL     1986
Query_DANIO  1927  LKSCFEAF   TE   NMTGEMV   LLE   LRRQFHCAAYNCAIALISCSFNE   T    KFYQGFLFTEKPDKNQ     1986
Query_SNAKE  1926  IKSCYDAF   TE   NMSGETQ   LLE   NRRQFHCAAYNCAIAVISCVFNE   N    KFYHGFLFTEKPDKNL     1985
Query_STERL  1939  LKSCFDTF   TE   NMAGETQ   LLE   KRRKYHCAAYNCAITVISCSFNE   T    KFYQGFLFSEKPEKNQ     1998
Query_MILII  1924  LKLCYDAF   TE   NMTGEKQ   LLE   KRRQYHCAAYNCIMAVVTCIFTE   A    KFYQGFLFSEKAEKNI     1983
Query_BAMBO  1922  LKSCYDTV   TE   NMTGETQ   LLE   KRRQYHCAAYNCMIAILCCIFTE   A    KFYQGFLFSEKPEKNI     1981
Query_ALLIG  1885  IKSCYDAF   TE   NMAGESH   LLG   KRREYHCAAYNCAIAVISCVFTE   S    KFYQGFLFTEKPEKNL     1944
Query_GAR__  1928  LKSCFEAY   TE   NMAGETQ   LLE   ERRQYHCAAYNCAIAVISCSFNE   T    KFYQGFLFTEKPEKNQ     1987
Query_TETRA  1887  IKIAFEYR[6]AD[4]NVQNNPD[3]-ID[3]ILRLFSCAVYNCLSQIITKTQNK   E    NILNQFLFQQ---KND     1958
Query_PARAM  1784  IVDCNKAS[4]-E[2]NILNINQ[5]QVN[3]IVEDYQSSAFITYIYVLMKTQTK   E    NIFYAHLFKKERDISY     1856
Query_SLEMN  1876  IQLCTLAK[7]--[1]NFVKDSI[6]YNS[3]VQRQYACAAYTCLTAIIMGTQTK   E    NLFSQFLFATLRDKGE[1]  1951
Query_STENT  1663  IVFCSKYR[7]--[1]LIDSI--   -SA[3]YIRQFCCAAYSCLLTCIRKTQSQ   E    KVYVNYLFRDQNWK--     1726
Query_NEMAT  1910  TKSAHAAK   SE   DMRGETV   AFE   LRRQLHCAAYNALLALISCTQTE   L    KFYTAFLFNENPVKGQ     1969
Query_PISTI  1939  TKAARAAS   IE   DMRGETV   ALE   LRRQFHCAAYNALVALICCTQTD   L    KFYVAFLFSENPIKGQ     1998
Query_CORAL  1917  TKAAHAAK   SE   NMQGETV   LLP   LRREYHCSAYNALISVISRTQTL   L    KFYNGFLFSEDLSKGQ     1976
Query_HYDRA  1837  SKSASQSY   KE   SYNENN-   -ME   LQRKYHCSAYNCLISIISCTQNE   M    KFFAGLLFAENLLKGS     1894
Query_LATIM  1927  LKLCYDEF   TE   NMAGETQ   LLE   KRRQYHCAAYNCAIAVISCIFTE   S    KFYQGFLFSEKPEKSL     1986
Query_SHRIM  1889  FRYAAKIS   KG[2]--RHDQT   HAE   LRRLCACAAYNAIIAVLACVQND   L    RFYTNFLFSANPTKGE     1948
Query_MCRAB  1884  LRQSSKVA   KG[2]--RHDQT   QAE   FRRLCACAAYNTLIAVLVCVKDD   L    KFYTSLLFNANPTKGE     1943
Query_AMPHI  1908  LKRTVRLA   KG[2]--RHDTT   HAE   LRRLCASAAYSLLIAILINVQDD   L    KFYNNFLFVQNTAKGE     1967
Query_BARNA  1792  LGSGAAAF   KT[3]-RCMEPA[1]LSE[1]-RRRCMCAAYCALAAAVSATQTE   A    RFYQTFLLAEAPERNQ     1854
Query_COPEP  1817  LPYLGDSR   --[1]KLSGLEI[1]-SE[4]LYRKYHCARLNCCIAITSCTNFN   V    NAYDTFIFRENKKKNQ     1879
Query_DAPHN  1730  MQLVLNIR   KA[3]--FNENE   -LE[1]-YQDMERFSFRLVLSIASVTPTL[1]E    ESMDKLIFRPMNQKFA     1790
Query_STARF  1947  TKACHAAK   SE   DIRGETN   ALE   LRRQFHCAAFNALIAIISCTQTD   I    KFYNAFLFSENTAKGQ     2006
Query_STRON  1937  TKAALGAR   NE   DCRGETV   AVE   LRRQFHCAAFNCSAAIISCVQDN   I    KFYNAFLFSENHAKGQ     1996
Query_SEACU  1912  TKAAHFAK   SE   NVGGEKV   CID   LRRQYHCAAYNALIAIITCTQTD   V    KFYNAFLFQENVAKGQ     1971
Query_FUNGI  1854  MRLGHNAK[9]RP[4]ASNEDKQ   TYT[4]RCFEYQCSAFNALAAVVSRTQTK   E    NFYSVFFFRENRDKGE     1930
Query_DPURP  2004  MKAAHSAK   --[2]KLTSDDR   YVT[2]LCTMYHGAAFSALASVIIATQTK   E    NFLHVFFFKENKDKNE     2065
Query_HALBU  2008  MKAAHLAK   --[2]KLATDEQ   HVT[2]LATKYHGAAYSALASVVATTQTK   E    TFFHVFFFKENKEKNE     2069
Query_GLOBO  1779  IKSGVDFK   KK[4]E-R--SI   ASY[1]--LIVNQTAFNAVAACLLATQNI   S[ 2]EIFNSYLFVEKP----     1836
Query_PALUS  1830  IKLANTSK   NE[4]EPG--GV   ASF[1]--TTYHQAAYNVLASVIALTQRD   Q[ 2]KFCTTFLFKENRQKQE     1892
Query_CALIF  1907  MKPVAYIA   RS[4]NSNNEDY   LKA[1]--KYYKQSSYCLIAAMLYCTQTK   E    NFFNVFLFQENDAKNE     1969
Query_GLOMU  1962  MEKAHKIK   SR[4]DNGNKLV   SEA[1]--LHFRCEAYNALAAVILCTQNN   P    VFYKAFLFAENTSKGE     2024
Query_LOBOS  1886  FRTANAAK   FK[4]PMTPET-   AAL[1]--EDYKRAAYNALASAILCTQRK   E    DFFRQFLFRDNEAKKE     1947
Query_RACEM  1872  VDSLDQAK   SE[4]EA--EVL   VET[1]--LRYHQMAYNTMTVAILNTQTK   E    AFLNGFLLKARKGAH-     1931
Query_PVIOL  2000  MKAAHSAK   --[2]KLTSDDK   FVT[2]LCTLYHSAAYCALASVILCTQTK   E    NFFHVFFFKENKDKNE     2061
Query_BOMBU  1781  YSNALNVR   ML[1]--THEPE   LKE   TTRLLHCSAYNCSIAIVSNIKRD   -[ 1]DAYAS-IFIENRKKEQ     1838
Query_AEDES  1906  TKKAYAAR   SE[2]-LTPDSP   TAE   LFRKYQCAAYRALAAIISNTKED   -[ 1]QLYNVLLFRENADKNE     1966
Query_ANT__  1807  FNNTLNTQ   SP[3]--TRIPG   CQK   LVRLLHCCAYNCLLAILSLKKEE   -[ 1]-FYNF-AFGNDNKRGL     1865
Query_WHITE  1839  LK----VF   VE[2]NCTSNFK   SAD[3]ALRRYHCASFNATMALICNSKRD[4]S[19]DFYTKFVFSKLGTTG-     1921
Query_BEETL  1772  LKFTLNAI   KD   TIDCVVE   YEE   LFRLYKCHSYNAMISIVCNSMKD   -[ 1]NFYTKLFVRQEDNKD-     1830
Query_THRIP  1807  VKTGLAAF   QN[2]NFEGISD   AKE   LYRQYLCAIYNTIATVFCNLKTQ[3]-[ 1]KFYKV-LFAENTRK--     1868
Query_TERMI  1872  TSRAYKYR   SE[4]NLTGNVQ   LVE   LFRQSQCAAFNTLVAIISNVKC-[1]-[ 1]RFFTGLLFSENKAKE-     1934
Query_FLEA_  1804  TRKAFKAR   TF[3]-DENDTE   SAE   LFRQYQCTAYKTLCILICSTKSD   -[ 1]KFYTGLLFRENPLKKE     1865
Query_MIDGE  1856  IKKAMEVR   TV[3]-IVDDPK   RQE   LFRQFHCYCYRALASIVSNTKDN   -[ 1]LIYNLALFKESSKTG-     1916
Query_LAODE  1842  LRTSTVRV[8]--   NQAAKDN   LKD[3]LYLSYKQIAFNTLVILISNTKRD   V    -------------KSL[4]  1901
Query_LOTTI  1894  TKCAHEAR   SE   DTRGETE   LVE   LRRQYHCTAYNLMISIISCTQTE   A    KFYNCFLFSEKPSKGQ     1953
Query_APLYS  1935  TKAAQEAK   GE   DCRGESL   LVE   LRRQYHCHAYNLLMAIISCVQTD   I    KFYKGFLFQDNVPKGQ     1994
Query_SCALL  1933  MKIGHAAK   GE   DTRGETT   LLS   LRRLHHCAAYNMLISVVSCTQTE   A    KFYQVFLFKEDETKGD     1992
Query_CRASS  1904  MRTAHEAK   GE   DVRGETS   CVE   LRRQYHCAAYNLLIAVLSSTQTD   A    KFYTAFLFKEDEAKGQ     1963
Query_OCTOP  1867  TRIAHQTK   KE   DVHGETL   FLE   LRRQVHCAAYNLLIAIISCTQTD   L    KFYNVFLFAENEEKKE     1926
Query_POMAC  1909  IKAANNAK   SE   DTRGDFM   VTE   LRRQYHCYSYNLLIAIICCVQNE   H    KFYKGFLFQDNTSKGQ     1968
Query_WHIPW  1857  VSLSVASR   RQ   -LLGADT   W--[2]RSWQLRCAAHNCVASVVLKTQTQ   L    KHFEGFVFSE------     1909
Query_TRICH  1865  VSCAVGAK   KE   -FCRDSK   LDS[3]HFFEYRCAAQNCVIALVKKTQSH   L    KHYEGFIFND------     1920
Query_AASTA  1753  CKTASQKS   --   --ASA--   ---   -DRPVAIAAFRCLLSTVRTTQTQ   E    KFYTQLLFADA-----     1797
Query_AEUTE  1710  CKAASQKS   --   --ANP--   ---   -ENSCAVAAYRCLLVTVRKTQTQ   E    KFYTQLLFADQ-----     1754
Query_PLASM  1853  CKCASKVV   --   --TKIYL   EAD[1]LVRLACCAAYNCLLAAISQTQKQ   E    KLYDQILFQSA-----     1904
Query_BREMI  1839  CKCASKLV   --   --TKEYS   NVD[1]LNRMACCAAYRCLLSAVSRTQKQ   E    KLYDQILFQPA-----     1890
Query_OLIGA  1856  CKCASKFI   --   --TKTHD   GLG[1]TTRFSCSQAYNCLLTAVSRTQKQ   E    KFYDQILFQEA-----     1907
Query_LIGNO  1981  SVQCYDLV   TQ[2]-----PR   CRE   SARQMQCEAYNLLVHCIKAASDK   E    RSYLP-PFTENAAKRK     2036
Query_CESTO  1774  IKKACSTL[4]NP[4]PLTDELE   SIA[2]LKQTVWSASLACLIAAINATQTQ   E    RAYSHILLGDSLTRFL[4]  1847
Query_OVI2_  2127  IKLANDVL[4]SN[4]LETSSSR   LKK[1]-RRHFFSAAWSCLSAAVCSTQTQ   M    KFYQKLLFRDGLFQQL     2194
Query_PRIAP  1922  TRLCYDAC   RE   DMRGERA   LLE   LRRQYHCTAYNVCVASVSRTKTE   P    VYYSSLLLQENPAKGQ     1981
Query_ROTIF  1876  LRKSRKHM   NE[4]NTGQESA   KIE[4]LERILHCSAYNCLISLFIRTQTE   P    KLYLAFLFKDDPSKNE     1943
Query_SPONG  1922  TKCCHGAK   SE   DLHGDRT   HYE   TRRQYHCHAYNTLMSVISCTQSK   P    QFFVGFLFKEDVVKGQ     1981
Query_TARDI  1924  VTRCANEF   AH[1]NVASEPG   MEE   VRRLFHCAAYRTLAAAVTATPIA   L    KFYHLCLFEQNAEKGK     1984
Query_TRICH  1897  TKAAHACK   SD   DLRGETR   SVE   LHRQCCCAAYNALISVISSTQTE   M    KFYTVFLFTENVARKQ     1956
Query_OSTRE  1839  MFELSGLG   --   -VAGKAQ[3]--E[1]----IALRVYHLAFLVFSCLLAK[2]T[ 2]KFYLKLFESTKS----     1893
Query_NITEN  1991  IRGKERLK   --[1]MGIAAEQ   SPE[3]AWRGIQQAAYGCLAALLIKTQ-G   P    NFFPRFLFKEGAKEG-     2050
Query_CLUBM  1803  TRGKNDPR   --[1]Y-----P   FQE[3]FWRDLHVSAYTCLAAVVMCTQKD   E    KIF-NLLFKDAGNP--     1856
Query_MOSS_  1906  CHGKLDPE   --[1]K-----L   YPE[3]TWRELHAAAYNCLAAVILCTQID   E    KWFNAFLFSEKSRP--     1960
Query_LIVER  1931  SRAKNDPV   --[1]C-----L   HFE[3]LWRDLHISAYKALAAITLRTQEQ   E    KVF-GLLFKEGSTG--     1984
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  1973  L--HWNNIVD-QDVV   YEF---PLEM---ENPIERRKKFVSIRTSVRDV[2]SDEPTG-[4]VHYLAS-   Q   YL  2032
Query_LINGU  1974  F--LLDNIVD-QNRK   YLF---EIEI---TSLPQRKKRFVSIRSEAGAQ[2]GNSQDS-[4]VHYLAS-   Q   YL  2033
Query_LIMUL  1961  F--IWENIVD-LRKE   YYF---PVEL---EAPMERKKKFLTIR-EGNKE   SNGIENN[3]VHYFASQ[4]-   --  2019
Query_SPIDE  1961  Y--IWKRIVD-EDIK   LTF---SMEI---DDSVAKKNKFLAIRHQIRKV   FDNDESE[3]SGFQSSL[5]E   NL  2024
Query_SCORP  1947  Y--IWENIID-LNKE   YTF---PIEL---DT-IERRRKFISVRHENREL   NSNKTLN[4]LKYAPSQ[4]-   --  2006
Query_TICK_  1889  F--LWNNLVD-EKAV   YTF---PLER---EVLFQKTRHVVGVTEGQVPS   -------[3]SGRMAGS   -   -L  1938
Query_MONOS  1880  Iv-----CVA-PTAD   ETFe-fKAEL-d-QAPTLRLRQLTALRSSARTQ   -------[2]TRYLTSQ   -   LL  1929
Query_ROSET  1855  D--CVFNLID-EKRK   YTF---TTEL---PAAMSVRSKCVLALREGEAT   PSTTAAA[5]-RYLPSQ   -   IL  1913
Query_BELCH  2002  F--LLDNLVD-TNKL   YQF---DAEF---SAPVSRKKQFVSIRQELRGQ[3]EAGS---   SRYMQS-   -   YL  2055
Query_CIONA  1970  Y--FLNNVLD-CNKN   YTF---DTEI---SQPLTKRKKFVAIRKKKAEK[1]GEASI--   -QYMAS-   -   -L  2020
Query_CAECIL 1989  L--IFENLID-LQRC   YTF---PVEV---EVPLERKKKYVAIRQEARNT   ANGESDA   PHYLSSQ   S   YM  2044
Query_HUMAN  1976  L--IFENLID-LKRR   YNF---PVEV---EVPMERKKKYIEIRKEAREA   ANGDSDG   PSYMSSL   S   YL  2031
Query_MOUSE  1973  F--IFENLID-LKRC   YTF---PIEV---EVPMERKKKYIEIRKEARDA   ANGASGS   PHYMSSL   S   YL  2028
Query_XENOP  1995  L--IFENLID-LQRN   YTF---PIEV---EVPMERKKKYFAIRKEARDA   SSTESDE   PSYLSSQ   S   YM  2050
Query_PLATY  1980  F--IFENLID-LKRC   YTF---PVEI---EVPMERKKKYIAIRKEARDS   VNDNSDE   PHYLSSL   S   SM  2035
Query_TASDE  1982  F--ILENLID-LKHC   YTF---PVEV---EVPMERKKRYIAIRKEARDA   ANGDSEG   PHYLSSL   S   YM  2037
Query_CANAR  1978  L--IFENLID-LKRE   YTF---PVEI---EVPLERKKRYIAIRKEARDA   GNSSQDE   PKYLASS   S   YM  2033
Query_OSTRI  1948  L--IFENLID-LKRQ   YSF---PVEV---EVPLERKKRYIAIRKEARDS   GNGNEDE   PKYLPSA   S   YM  2003
Query_TURTL  1987  F--IFENLLD-LKHR   YSF---PVEV---EVPMERKKRYITIRKEARDA   VNRDSDE   PQYLPSA   S   YM  2042
Query_GECKO  1987  L--IFENLLD-LQHR   YSF---PIEV---EVPLERKKRYIAIRKDAREA   MSVDSAE   PHYLASA   S   YM  2042
Query_DANIO  1987  F--IFDNLID-SQRV   YNF---PIEI---DVPIERKKKYVMIRK---EV   SGENGDA   PVYLSSQ   S   YM  2039
Query_SNAKE  1986  L--IFENLLD-LQHR   YSF---PIEI---EVPLEKKKRYITIRKEARQT   MSGDSAE   PRYLASA   S   YM  2041
Query_STERL  1999  F--IFENLID-AKHT   YSF---PVEI---EVPLERKKKYVAIRKEARQG   GEGDSDE   PHYLSSQ   S   YM  2054
Query_MILII  1984  F--IFENLIE-AHRT   YNF---PVEI---EVPLEKKRRFISIRKEAREA   GDGDSGE   PQYLSSQ   S   YM  2039
Query_BAMBO  1982  F--ILENLID-SQRA   YNF---PVEI---EVPVEKKRRYFAIRKEARES   AEGDSDE   PHYLSSQ   S   YM  2037
Query_ALLIG  1945  L--IFENLID-LKHQ   YSF---PVEI---EVPMERKKRYIAIRKEARDT   VNGGPDE   PQYLPSA   S   YM  2000
Query_GAR__  1988  F--IFENIID-AQQT   YRF---PVEI---EVALERKAKYVAIRKEAREA   GDRDAEE   PHYLSSQ   S   YM  2043
Query_TETRA  1959  H--LWDLLIDgTQKE   FNF---EIETnf------EERSYARTRNIFSRI[1]-KFDTKE[3]KKYLNET[3]T[2]--  2018
Query_PARAM  1857  ----WDNFIN-KEWQ   FNF-----QV---ETNFEEKPLYTEVKKQQQDY   -------   FQEKIVK[1]-   YL  1901
Query_SLEMN  1952  C--FWNVLID-QEKA   YNF-----EV---ETNFNYIQLKQIEFKIEQAN   -GGDISQ[3]NNLINEY[5]S[1]--  2011
Query_STENT  1727  --------IL-LEEK   EDF-----G------------------------   -------   -------[2]P[1]--  1740
Query_NEMAT  1970  F--LLDNLID-TDRK   YEF---EVEL---STPLERKKQMTSIRINARVQ   SDQVEEA[6]VRYLSS-   Q   YL  2030
Query_PISTI  1999  L--LLDNLVD-CNRT   YTF---EVEL---SAPLERRKQF---RMQARIP   SPKENSL[2]SAGRTGL[7]Q   YL  2060
Query_CORAL  1977  F--LLDNLVD-CTKK   LHF---EMEL---SAPLDKRTLLTSIRCEARAT   RMNSSEN[6]LRYLSS-   Q   YL  2037
Query_HYDRA  1895  F--VLENLVD-LKKE   YNF---KSNV-----SFEKMKVMPLIRKASCEL[1]HQPNQRE[2]-EYILSN[3]R   FL  1953
Query_LATIM  1987  Y--ILENLID-MKCS   YNF---PIEI---EVPIEKKRKYVAIRKEARGS   TIGDLDE   PQYMSSQ   S   YM  2042
Query_SHRIM  1949  V--IWECLVD-CSRN   LEF---DVEI---NFKPASKKRFVAVRQKLRET[2]EEGAVVA[2]VQYMAS-   H   YL  2007
Query_MCRAB  1944  A--IWECLID-CNRN   LQF---DIEI---NLNPGSKRRFVAVRKNLKEA[2]SSGKGAA[2]VQYLAS-   H   YL  2002
Query_AMPHI  1968  A--LWASMID-VNQK   LEF---EEEV---GVASKNTSRLVAVRQVLSKS[2]VAEARYQ[2]VLMFKS-   N   YL  2026
Query_BARNA  1855  L--IWERLVD-AERR   YHL---PVEG-----GSELRRWHVSIPRPAGAG   --GGQSA   PRRTGSV[2]-   LL  1907
Query_COPEP  1880  F--IWSEIID-ISQP   IHLgl----V--lDKEFFKRKHLVTVRQSDSSS   ----SAL[1]---IGSQ[2]-   YL  1930
Query_DAPHN  1791  ---LWNAMVG-TTKR   YNLv---AEV-------ERHRKRFRLHLLTVSS   -ENRPTK[1]--YIPSQ   -   FL  1838
Query_STARF  2007  F--LLDNIID-ITKN   YSF---EIEM---NAPMERKKRFVSIRNQIQDT[4]SSQQSDS[8]GRYLST-   Q   YL  2073
Query_STRON  1997  F--LLDNLVD-PKKE   YGF---AIEM---DAPVERKKRFVAIRTQARET[3]EESGYVS[1]PAHYLSS   Q   YL  2056
Query_SEACU  1972  F--LLENIID-EERS   YEF---EVEI---SAPLERKKKFVALRLQAA--   SGGTFDT[8]GRYLSS-   Q   YL  2032
Query_FUNGI  1931  H--DYK-----LEVE   TNF---STVK---R-TVGSTHAESY---TETES   -RQKR--   --SLIN-   -   FL  1971
Query_DPURP  2066  Y--LWENIVD-CDQQ   YNF---ESET---NFTVSFSSPQSL---F---A   -NDLR--   --YLSSQ   -   YL  2109
Query_HALBU  2070  Y--LWENIID-LEQQ   YNF---EPET---NFQISSQLSSAV---ANSQS   -RDLK--   --YLSSQ   -   YL  2116
Query_GLOBO  1837  C--LWENIVD-TSRP   IEI---E--T---DLSAPLIRVGL-----KDFS   -EKVSM-[2]-KYLSTQ   -   FL  1883
Query_PALUS  1893  D--PWKNIVD-IHE-   -------------DLRDKLNDISLLQ--SRTIG   -TPLAR-[1]-AYLPSQ   -   FL  1935
Query_CALIF  1970  K--IWENIVD-THHK   IFF---KVED---GVPPKKSKRYYEN--IQYRK   -NIIRK-[2]--DNFSS   -   FL  2020
Query_GLOMU  2025  L--LWENIID-LEIH   YKI---K--I---ELNQPFFRTKVDD--FRSRS   -GKTFK-[7]LKYLASQ   -   YL  2080
Query_LOBOS  1948  F--VWENIVD-LEAK   YHF---E--Q---VLSGPLVKTRLSD--LRTKS   -LNPNK-[5]YRYMSSQ   -   YL  2001
Query_RACEM  1932  ----WENIID-MKKL   KES---PLRI---ELDQPLATTRLGE--FRERS   -GFRKQ-[5]VNYIASM   -   QL  1985
Query_PVIOL  2062  Y--LWENIVN-TELE   YKF---EPET---NFIINYNSPQSL---F---S   -NDLR--   --YLSSQ   -   YL  2105
Query_BOMBU  1839  L--IWENIVD-CQKQ   YNF---QQTV---AENPKYRKRLVNIRK--RI-[1]QRQASSH   YSYIYS-   Y   DL  1891
Query_AEDES  1967  Y--IWRKLVN-CTED   HLYdsfSQEL---EDYPKIKDKIVSIQRLSSD-[1]GKSHSNR   FKYIQM-   T   SV  2024
Query_ANT__  1866  V--ILKNIVD-CDTQ   YEL---DQTF---KEYPKTREIIVNIRTTEVD-[1]GGRQTHR   YAYVHS-   Y   DL  1920
Query_WHITE  1922  ---FMDKIIN-LKQN   YRF---HVDFns-DELPKQKKTLVSIRRAAR--[2]TGRAPIQ   GSMDIS-[1]Q   SL  1978
Query_BEETL  1831  I--LWSSLID-TKKT   YVF---QADF---DNIPKQTKILLNIRDEMRL-[4]ENKQTQS   VRYIET-   Q   RV  1888
Query_THRIP  1869  ---FWRLLVD-DDKC   YML---PFQ--q-TEHLKNVEKIVSIRRNVS--[2]KDSSQRM   NAHKSLQ[3]S   PS  1926
Query_TERMI  1935  V--PWGKLVN-CNKE   YAL---ALD--w-EEIPARRKQLVSIRCQARAQ[4]LGFASQS   VRYIPS-[2]Q   YL  1995
Query_FLEA_  1866  I--IWRNIID-CKRI   YEF---PQEF---ENYPKTKKHLVEIRK-----[1]ASVDSEE   LNYLTQ-   Q   SI  1916
Query_MIDGE  1917  Y--IWRNFID-TKNE   DLYgnfTQEF---DELPSLREYIVSIKDLQTA-[1]TTSSQRK   IS-----   S   SI  1970
Query_LAODE  1902  L--IFKNVKF-IMPR   ELI---TITNmafDEIPKRRQAQIGIRRAVRSE[2]KRRNSTS[2]ESYQLFT   Q   SL  1964
Query_LOTTI  1954  F--LLDNILD-CEKM   FEF---KAEL---DSPLERKKRFVSIRNEVKE-[1]RQGEEE-[4]SSYNIST   Q   NI  2012
Query_APLYS  1995  F--LLDNIID-MDFA   HSF---DIEL---DSPLERRKKFVSIRSEAQEQ[2]GAGGGAG[8]PSLHLAS   Q   YL  2060
Query_SCALL  1993  F--LLDNLID-KERT   YEF---PTEL---TSPMERKKQFVSLRKEVQER[2]EDGVTM-[4]SSFNLAS   Q   YL  2053
Query_CRASS  1964  F--VFDNLLD-KDKA   YVF---SSEL---KAPLSRKKQYVSIRNEVKES[2]QNGEAD-[4]SSLHLAS   Q   YL  2024
Query_OCTOP  1927  F--LLDNIID-KEKK   YSF---PIVF---EMQPKKKTIFHTLQKEIKQK   VVADASQ[2]-SYHLSS[5]-   --  1985
Query_POMAC  1969  F--LLENIVD-RDQK   YEF---PLEL---ESQFERKKKFVAIRSEMQEK[2]DK-----[4]SSYHLAS   Q   YL  2025
Query_WHIPW  1910  ---CFDYITD-SELK   YEF---GEWF---AEEESDGQKTDDMSGLNRVL   -------   YKKAPSI[2]Q   TL  1959
Query_TRICH  1921  ---CWESIID-KEVK   YEF---DREI---QEEEQSAGKMYYSRQDRPSS   -------   AEYENQV[2]S   VL  1970
Query_AASTA  1798  ---IWPNIMA-KTED[6]YSF---ETQT-a-AFPTKYLSRRQGGRSRTTHR[1]-------   NSRLMDI[2]Q   FL  1855
Query_AEUTE  1755  ---IWPKIVD-QSQD   YIF---ETET-t-SFPTKYLS---DARKRTQLR[1]-------   FQR-MDI[2]Q   FL  1802
Query_PLASM  1905  ---LWSNLVD-SALE   YEL---RAET-e-NFVTIPLSSLSTVSLKQRYE   --VDVLE[3]TKRNNSA[2]Q   FY  1963
Query_BREMI  1891  ---LWSNIVD-PARQ   YEL---AAET-e-IFTTIPLSTLSTVSLQSKLD   --TSNMA[3]QKRTEST[2]Q   FF  1949
Query_OLIGA  1908  ---LWKNIID-GARE   ISL---KAET-k-AFDRVPLSSFSARTLRNQQE[4]GGTQSTS[3]GKRNSST[2]Q   FF  1972
Query_LIGNO  2037  F--IFDNIVP-TDVD   YQL---PLQV---EQSASRKQYLVSVRPAASSA[2]SSASEAG[9]PSSSSTS[7]S[7]FL  2117
Query_CESTO  1848  F--IFPRRRS-RKYR   SAF---IELReqfWLPSEREPYFGKIDVKSQDM   DSSLQQK[1]----KSG[1]-   LL  1903
Query_OVI2_  2195  LpdSFETPLY-LPSQ[6]GRFlnvPETLws---SDELNSRAITFQPMINEK   -------[6]PPYGSPS[4]S   AF  2264
Query_PRIAP  1982  L--LWDNIID-CHRR   YGF---DREL---SAPLQRKRQFVAIRTAVREI[2]DADGQGA[5]-HYLAS-   Q   YL  2042
Query_ROTIF  1944  F--VFEPLID-KIKS   YNF---CVEM---ENFTQRKTKFISIRDEFKVI[1]ADTNSRS[9]TSSLSTQ[1]-   -L  2008
Query_SPONG  1982  L--LWDNIID-CEKK   YEF---QVDF---EAPLKRRKQITSIRDDLKQS[1]AAGNSSG[9]PRYISSQ   -   YL  2046
Query_TARDI  1985  I--LWANLVD-LKRQ   YTF---EVEF---IQAPRKKKQLAAIHKEAQQD[4]NGGVGSS[9]TSYASAR[4]S   LV  2057
Query_TRICH  1957  F--LWENLID-MERE   YAF---EIEL---SYQDKQKTTFRSLREQSRSS   VSSESAD[9]PSYHLS-[1]Q   YL  2021
Query_OSTRE  1894  ----WNNLVDrTKRI   ------ELEA-e------------AINEEPDSA   ----EDI[1]---LKSP[4]S   FT  1933
Query_NITEN  2051  F--LWSNVVD-CDKV   YDH---FDVE---TAELGSARETLRGLRAERKA   -RESRRM[9]HATLSSQ   -   YI  2113
Query_CLUBM  1857  ---LWQHLVD-CSKV   YSS---FPVQ---TSTLVESHAIVKGIRAMIKG   -HGP---[5]SMTLSSQ   -   YL  1911
Query_MOSS_  1961  ---LWQHLVD-LTKV   YDD---MSVE---TSQSSIISQTVKGMRADLKA   -RRTR--[8]ASTISSQ   -   YM  2019
Query_LIVER  1985  M--IWQHLVD-TIKV   YNE---LPVE---TNQLASSFQTVRGMRDERRS   -RKT---[8]LSTLSSQ   -   YV  2043
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2033  AGSSLSTDVSQFDF--GS   SMVSSADSS[1]DNQRSRE-   -SSIEV-PDVVL    -KEHY-I---EMEMDAL--N  2088
Query_LINGU  2034  ADSSLSEDVSAYDY---S   GSLQGERSE[1]VPPRTRQ-   RAVSEE-AAVPA    -EADY-E---EIEMDEL--N  2089
Query_LIMUL  2020  --SSLSEDVSRFDF-STS   EQVFSTPIE   EHMARKRR[2]SSNPEQeRQVIM    -VNDY-I---ELENDDL--N  2078
Query_SPIDE  2025  ELSSLAEDVSKFVF-TTN   EQIFKPEGK   NELNK---   -STDNMsSKNTY    -VADY-L---ELD-DLL--N  2078
Query_SCORP  2007  --SSLSDDVSQFDF-SSS   ELAFSQDMK   NKNRKEEI[2]NPSNEAkDDDLV    -SNIT-V---ELEIDDL--N  2065
Query_TICK_  1939  RGSSLAQEASQYSL-SAS   ----GLAFW   DVETVDSD[1]SKGTQApAKEPV    SRLKV-M---DLEDDAL--N  1995
Query_MONOS  1930  ADSSLSQSMASWAD-PNG[1]-SSLEAAPV[1]-GAEAAGA[7]TEAKNL-PEPMD[13]AGSTLtLsgtQLELDAL--N  2012
Query_ROSET  1914  ADSSLSQDAAQLEAlRQG[3]GESQAVEDA[5]SPPSSDGP[5]MDASED-SPGA-[ 7]VVEDS-----QMEMDEL--N  1991
Query_BELCH  2056  AGSSLSEDISQFDF-TSG   TQVKHSQTF[3]QERYSQAS[5]LQARQQ-QDSSV[ 2]-HGDF-L---EIEMDEL--N  2123
Query_CIONA  2021  AGSSLSEEISVFDRnRSL[4]TSSANISSM   SVTNSTDE   FENTDS--QPAI[ 3]FSGNF-V---ELEVDSL--N  2086
Query_CAECIL 2045  ADSSLREEMSQFDF-STG   VQSFSYSSQ   DPTAAQFH   SRKREH-PE-FT    VQEDV-M---EFEMDEL--N  2102
Query_HUMAN  2032  ADSTLSEEMSQFDF-STG   VQSYSYSSQ   DPRPATGR   FRRREQ-RD-PT    VHDDV-L---ELEMDEL--N  2089
Query_MOUSE  2029  TDSSLSEEMSQFDF-STG   VQSYSYSSQ   DRKPTTGH   FQRREH-QD-SM    TQDDI-M---ELEMDEL--N  2086
Query_XENOP  2051  ADSSLIEEMSQFDF-STG   VQSFSYSSQ   SKMLSQSA   SRKKEQiTE-GK    TFDDV-M---EFEMDEL--N  2109
Query_PLATY  2036  ADSSLSEEISQFDF-STG   VQSYSYSSQ   DVKSTGYH   FKRREY-RD-PV    TQSDV-L---ELEMDEL--N  2093
Query_TASDE  2038  ADSSLSEEMSQFDF-STG   VQSYSYSTQ   DSKATVVH   FRRRER-ID-SM    VQDDV-M---ELEMDEL--N  2095
Query_CANAR  2034  MDSSLSEEMSQFDF-STG   VQGFSYSSQ   DVTASSTR   FRRKEP-TE-YK    VLNDE-M---ELEMDEF--N  2091
Query_OSTRI  2004  MDSSLSEEMSQFDF-STG   VQGFSYSSQ   DVTASSPR   FRRKEP-TE-YM    ALDDD-T---ELEMDEL--N  2061
Query_TURTLE 2043  VDSSLSEEMSQFDF-STG   VQSFSYNSQ   DPTASSIH   FRRKEP-TE-SM    DPDDL-M---ELEMDEL--N  2100
Query_GECKO  2043  ADSSLSEEMSQFDF-STG   VQSFSYNSQ   DPTASASH   FRRKEP-TE-SM    GLGDI-M---ELEMDEL--N  2100
Query_DANIO  2040  ADSSLSEEMSQFDF-STG   VQSFSYNSQ   NPSGVSSS   SRMRER-KE-VL    SQDET-V---ELEMDEL--N  2097
Query_SNAKE  2042  VDSSLSEEMSQFDF-STG   IQSFSRSSE   DPTSSGSH   IRKKEI-TD-TV    VPGDM-M---ELEMDEL--N  2099
Query_STERL  2055  ADSSLSEEMSQFDF-STG   VQSFSYSSQ   DPSARAAR   GRRRER-VE-SL    VQDDV-V---ELEMDEL--N  2112
Query_MILII  2040  ADSSLSEEMSQFDF-STG   VQSFSYNSQ   DPRGHAVR   SRRKEQ-AELSV    VQSDI-M---ELEMDEL--N  2098
Query_BAMBO  2038  VDSSLSEEMSQFDF-STG   VQSFSYSSQ   DPKVSVLR   TRKTEW-TESNI    MLNDI-V---ELEMDEM--N  2096
Query_ALLIG  2001  VDSSLSEEMSQFDF-STG   VQSFSYNTQ   DPTATSAR   FRRKES-TE-SM    VLDDE-M---ELEMDEL--N  2058
Query_GAR__  2044  ADSSLSEEMSQFDF-STG   VQSFSYGSQ   DPSTRPMP   SRRRDG-AE-AA    LQDDV-V---VLEMDEL--N  2101
Query_TETRA  2019  -----KTSASQAFLpSFA[1]-ELISKNSE[5]DEEDESEN[4]TKKYMQqQQNL-    FFGDQ-Tg-eNLEMDLI--N  2084
Query_PARAM  1902  YTASLSVEQVTETTmEIE[1]-ENENLQST[1]--------   --------EVQ-    ----------KIEMDEV--N  1940
Query_SLEMN  2012  --MVFETQPDKINIkSE-   EEKLQKL--[2]NKKQNTNQ   -DQNMLdKSAA-[ 2]DSEIQ-Dv-sQLELDNV--N  2071
Query_STENT  1741  --TLFVKTEKKLEIvDE-   PKKLSYVYL[5]SQQATTIR   -AQKKIiVENN-[ 2]DEE-K-Sn-vFLEADEL--N  1804
Query_NEMAT  2031  ADSSLSEEVSQFDF-STP   VQVNIINYL[2]TGKPDQLG   EDSYEI-GE---    D----------VEMDEL--N  2082
Query_PISTI  2061  ADSSLSEDISQCDF-SSP   IQVHSGTEA[2]PLYKLNES   SSTAEQ-TD---    DWYEV-Se-dTLEMDDL--N  2120
Query_CORAL  2038  ADSSLCEEVSQFDF-STP   IQA--LSQS[2]TEQDIAGG   ENSGTP-ASPKE    SDSDI-Ss-eELEMDEL--N  2098
Query_HYDRA  1954  MESSLSQEVMEFSInKPF   --------K   PYTNNEEK   SLSNEE--ESVE[ 2]-----------LELDEL--N  1999
Query_LATIM  2043  ADSSLSEEMSQFDF-STG   VQSFSYSSQ   NRAAAVVH   SRRKER-TE-SV    VQDDV-M---EFEMDEL--N  2100
Query_SHRIM  2008  SDSSLREDMSQFDF-NNS   ---VVLAMS   QTEKSSER   VTGANE-MGGVL    -M--P-V---HMDQSDY--N  2060
Query_MCRAB  2003  DDSSLREDLSQFDF-SNS   ---VVLAMS   QREDNQYR   IAASND-ETGEL    -L--R-M---NFEQSDY--N  2055
Query_AMPHI  2027  ADSSLSQDLSQYDF-SSA   ---LVLSLG   EKTQHS--   -KPRDI-TEGTL    -L--P-M---TLETCDF--N  2076
Query_BARNA  1908  ADSSLGSQVTDYDF-TAS[1]----DLSTQ[1]-LQQQAEA   GGRERS-APAS-    ----L-----QLPADDL--N  1957
Query_COPEP  1931  SDSSLSQDISHYDY-TAS[1]---------   --TQGVTI   EVRKDT-PELV-    EPINY----iEIEEDAL--G  1978
Query_DAPHN  1839  FGSSLSEDVYRFDFnQAS[2]TKFGSYDEE[1]-NLPVSSP   -----TiSDAS-    ---SV----lYLTLDDL--N  1892
Query_STARF  2074  ANSSLSEDIQQFDF-SGG   VQQFNASID[5]SQRSSVPH   IKKRYS-NEET-[ 5]-QGEY-I---EMESDAL--N  2140
Query_STRON  2057  SDSSLSEDVNQFDF-SSS   VRHFGSSM-[3]RDGGSQGS[7]SMASRQ-DSKPV[ 2]PGGEY-L---EMEEDAL--N  2126
Query_SEACU  2033  ADSSLSEDVHKFDF-TGP   VTQTGS---[2]DQSQSQGS   --RGSQ-IRGTS[13]ELGDY-L---EMEGDEL--N  2101
Query_FUNGI  1972  ADSSLTQELGSSTTfLGK[2]RTATPTLEK[2]NPA-----   ------------[ 5]SPAK---e-eLVELDEI--N  2024
Query_DPURP  2110  VDSSLSQ--DFISKsFLE[2]-KDLNSDQP[2]TPV-----   ------------[ 5]IHQ------sDIEIDQV--N  2157
Query_HALBU  2117  QDSSLSQSLDVDTSsSTT[2]STTTTSNN-   --N-----   ------------[ 5]ASAM---n-mPIELDKI--N  2164
Query_GLOBO  1884  ADSSLSQMSSLFGQkDAR[2]-SLTQSSIP[2]SQDFSS--   -------QNSDF[ 9]SAVND-As-nALEMDLL--N  1949
Query_PALUS  1936  ADSSLSQTITI-----PL[2]-PSTLLANA[2]GPVYSA--   -------PTQSS[13]VMNAS-Pt-gDQRSTSIifD  2002
Query_CALIF  2021  SASSLSQSVDFLGGnDDM[2]------NNQ[2]ENETNN--   -------NENDL[13]ET--------FFDKDCV--N  2079
Query_GLOMU  2081  ADSSVTQGGALFKElP--   DESEGESKN[1]---HSD--   -------TIPME[12]-----------LEIDVF--N  2133
Query_LOBOS  2002  SDSSLSQSVGMIGEsEVY[3]DDGTTNQGN[2]TPELSG--   -------PNPNA[12]----------MLELDAI--N  2064
Query_RACEM  1986  AGSSLSQSSLSDMAlE--   --ASGQVSS[2]TDEEAV--   -------DDNDD[13]EVSGR-Ge-nELEMDVF--N  2050
Query_PVIOL  2106  VDSSLSQ--DFISKsFIE[2]DKVVDSNNI[2]GQK-----   ------------[ 5]VNEI---d-tDIEIDQI--N  2156
Query_BOMBU  1892  SSCSLTEDINAYDF-NET[1]I-RNKHDST[2]KESMS---   ------------    ----L-----TFENDEL--N  1933
Query_AEDES  2025  FESSLSQDVTKLDL-NYS[1]VRTTQEMAT[2]AEEFAALK   ---QHR---AT-    ----V-----SLERSKI--N  2075
Query_ANT__  1921  SACTLSEDINAYDL-NKC[1]V-LPANFHR[2]STTNSDDS   ---LRThNATS-    ----I-----ALNSNDF--N  1973
Query_WHITE  1979  FAGSLSEDLTSYDL-SNH[1]V-VEVEAVQ[2]QVAAGSDI   ---NES---AC-    ----L-----ELEMDVI--N  2028
Query_BEETL  1889  FNSSLAEDVTKFDF-TSS   ----ILRTQ[2]DDAESNEL   ---FVQ---SE-    ----I-----ELDCTDI--S  1934
Query_THRIP  1927  LAGSLETDITRFDF-TYS   R-LLKPTDN[2]ADAEK---   ---------SV-    ----I-----MLENDSL--N  1969
Query_TERMI  1996  FDSTLNEDVTLFDF-STS[1]V-STEARPG[2]EKTPKESG   ---AES---TE-    ----V-----ILESDDL--N  2045
Query_FLEA_  1917  FSSSLSQDITEFDM-THN[1]V------RR[2]VDHKTKVE   ---NFG---TQ-    ----I-----YLTEDFL--N  1961
Query_MIDGE  1971  FEHSLSQSLTKTDL-SHS[1]VFSNREALQ[2]QQRMEDEQ   ---QRT---MR-    ----L-----QLESTPI--N  2021
Query_LAODE  1965  GNSTFADEIRKYDF-PNS[1]--------L[1]EDTPTQEN   ----ED-TDQM-    ----I----vELEADCL--N  2008
Query_LOTTI  2013  EESSLSDDLNQYDF--NS   GTAQVFFSD[1]FSDLTS--[4]QPDQFN-TTV--    ---DY-I---ELELDEL--N  2068
Query_APLYS  2061  ADSSLTDDLNQYDF-NIS   SSAQGSFSG[1]DVKSGS--[4]SEESQL-SEPGV    -SGDY-M---ELEMDSL--N  2121
Query_SCALL  2054  ADSSLHDDVSQYDF-STA   GTSQFSPSP[1]DKGKTRLG[7]PDDDTE-TKVVV    -QDDY-V---ELEMDDL--N  2119
Query_CRASS  2025  ADSSLHEDLNQYDF--TA   TPQMSEGSP[1]EKVLTV--[7]EEDDTD-TKVMV    -EDDY-V---ELEMDEL--N  2087
Query_OCTOP  1986  --SSLSVDIGRYDY-LSS[4]QEIASISEK   ENSAN---[2]AISEEKiIDY--    ----------ALEMDEI--N  2038
Query_POMAC  2026  TESSLSEDLSKYEF--TV   SGSAQNTSG[1]PKTKASL-[3]SSYGGD-NKVRL    -DGDY-V---ELEMDAL--N  2085
Query_WHIPW  1960  WESSLAPSVIQYDF-TES   FEPIN---T   ETTSAAIF   --------SYKL    ----P-A----SRSDAV--D  2003
Query_TRICH  1971  LDSSLSVMATQFDF-TES   LEARTHCAE   TRTESAND   --------KDVM    ----K-Fy-sKTKLDLL--D  2020
Query_AASTA  1856  QGSSLSQSAAPI----ED[1]-------ND[1]EPMANDDE   ----------E-    -DEDV------LEMDAL--N  1893
Query_AEUTE  1803  AGSSLSQSSEGE----QL[1]-------SN   ---PTSEE   ----------T-    -EDEV------LELDIL--N  1836
Query_PLASM  1964  TNSSLSMDAVTL-TaSTV[1]-------EP[1]DINRGASL   ----------Y-    -QDIE------IELDKI--N  2004
Query_BREMI  1950  TSSSLSMDAERV-GpSTM[1]-------AS[1]DIDQAVAN   ----------F-    -QNII------VKLDDL--N  1990
Query_OLIGA  1973  TASSLSQTTMTSEGiGSL[1]-------LD[1]ELVAREAI   ----------Y-    -QGID------IELDSI--N  2014
Query_LIGNO  2118  ADSSLSSDVSQFDRlVTA[1]LSNTSDSAS[4]SLSLSTSS[3]INQRRS-RKNRF[ 7]TRRNL-Iv---LESDSV--N  2192
Query_CESTO  1904  LGSSLSLEINRFMR-ASG[3]IRGSTSSTD[5]SPGVSPPL[6]FRQLEK-VEAE-    --PNY----vNLEVDPL--N  1973
Query_OVI2_  2265  AGSTLALELNRFDR-IAG[1]-QLLELEMH[1]-NTVETEY   ADTNEPvSNIT-[ 9]TTGKC----vQLPMDSV--N  2332
Query_PRIAP  2043  ADSSLSEDLSRYDF-NVS   TQIRERGDP[2]ERKPEGEA   --EEQE-EEAAM[11]LAGDY-V---EMEMDEL--N  2112
Query_ROTIF  2009  YSSSLSEELSVFDFsVHS   SYNQAYNSQ   SSVTRAPS   FRRSYTdDELNS    -QEQS-V---EIELDEL--N  2068
Query_SPONG  2047  SDSSLSTDVSQYDF--TL   PHHHQYGSV[4]STESSSID[3]GNSQEA-SSYVV    -SEEQ-L-----EEDDL--N  2108
Query_TARDI  2058  QGSSFNDEMSRYDY-SEG   ---------[1]DRTSTPLA   --ARED-AMGTI    ----------ELEMDAL--N  2100
Query_TRICH  2022  ADSSLNEEISQFDFnSST   VQGYSENNS[3]IPTRETGD   SEKAKT-II---    --NKF-Ad-eELELDDL--N  2081
Query_OSTRE  1934  LSSTLAAMYSTLKTeSCV[1]-------SD[2]DDEWGKET   ----EIpAKGQ-    ------------PRDVL--E  1977
Query_NITEN  2114  NNTSATQGGGVISTfLGG[2]ARLGADGGA[2]AGSQGGSL[7]DLSASMlPDAPS[13]DEDSVvEt-sALDEDDF--D  2201
Query_CLUBM  1912  VGMSTSQEPTVISSfVGD[2]MASLPETDE   --------   ------------[13]VEF------vALDQDDF--D  1965
Query_MOSS_  2020  HSASLSQEPAVIRTfVGG[2]KE-ETSAAG[2]AGL-----   -----LlDE-ES[13]EII---Et-tALDRDDF--D  2085
Query_LIVER  2044  VNIGLSQEPAVISSfVSQ[2]DHSEGDSAS[2]LEF-----   -----NlGEGET[13]VSDHMvDt-iPSDQDDF--D  2114
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2089   QHPCMAALSALLQHM   KLNQI--TP   E   VEKGSKPSSMPDWMKNIHDKIK--NSD   TQL--NIRLFLAKVIL   2150
Query_LINGU  2090   QHECMTAMTALLKHM   QRNKI--TP   D   VVSNPLD--VPPWMKYLRDKMN--STT   THI--NVKLFIAKLIL   2149
Query_LIMUL  2079   RHECMASLTALLQTM   QQKGI--IA   A   TDQKS-ITPVPKFLEGLCDKLN--NSE   VHK--NVKLFLARLVV   2139
Query_SPIDE  2079   QHECMLSMLSLIRSM   KNKSF--IS   D   EKAGA-P---PAWLSAIVQSLE--DRK   SSD--NVKLFLGRFIT   2136
Query_SCORP  2066   QHECMNSLTMLVQHM   EKIQL--LS   S   SPKE-----IPKWLDNLCKKLE--DYS   VHR--NIKLFLTKLFI   2122
Query_TICK_  1996   GHECMASVCAVLCRM   ADLGI--IP   Q[1]ADEEVNIGPAPKWLTFLTNAIV--GNH   TTR--NALCFICRVVV   2058
Query_MONOS  2013   RHPNMLTMLALVEQM[9]ESGSAaaSA[3]A[2]APAIPAAETLPAWMSELLSIVN--AAH   VHM--NQKLFIAKVII   2090
Query_ROSET  1992   SHESMVSLMGLIQRI[3]-HAES--KP   Q[3]DGLILPSSTMPSWMSNLLKSFE--STS   TSH--NVRLFIAKLIT   2058
Query_BELCH  2124   QHECMSSMSAVLQHM   ARNNI--TP   Q   VQKGAAPKDMPPWMSCLNKKLG--GSS   THK--NVKLFLAKLIV   2185
Query_CIONA  2087   EHECMPALVGLITHM   HENNI--GF   P   EEEKEEESSLPSWMSSIHKKLS--SSS   TSK--NVRLFIAKLII   2148
Query_CAECIL 2103   QHECMAAMTGLIKYM   QKTQI--TP   K   VEEGTVPQDLPPWMKFLHGKLG--NTS   TPL--NIRLFIAKLIA   2164
Query_HUMAN  2090   RHECMAPLTALVKHM   HRSLG--PP   Q   GEEDSVPRDLPSWMKFLHGKLG--NPI   VPL--NIRLFLAKLVI   2151
Query_MOUSE  2087   QHECMAPMIALIKHM   QRNVI--AP   K   GEEGSIPKDLPPWMKFLHDKLG--NAS   VSL--NICLFLAKLVI   2148
Query_XENOP  2110   QHECMAAMTGLIKHM   NRSEI--TP   K   VDEDASPQELPSWMKFLHVKLG--NTS   TPL--NIRLFIAKLIV   2171
Query_PLATY  2094   QHECMATMTALIKHM   QNNQI--TP   K   VEEGSGSTNLPPWMKFLHGKLA--NPL   GAL--NIRLFLAKLIA   2155
Query_TASDE  2096   QHECMTTMTALIKHM   QRNHI--TP   H   VKEGSEPPSLPPWMKFLHNKLE--NPS   TPL--NIRLFLAKLII   2157
Query_CANAR  2092   QHECMASMTSLIKHM   ERNQI--TP   K   VEEGTVPADLPLWMKFLHGKLG--NPS   VPL--NIRLFIAKLIV   2153
Query_OSTRI  2062   QHECMASMTTLIKHM   QRNQI--TP   K   VEDGTVPVDLPLWMKFLHGKLG--NPS   VPL--NIRLFIAKLIA   2123
Query_TURTLE 2101   QHECMASMTTLIKHM   QRNQI--TP   K   VDEGIVPRDLPPWMKFIHGKLG--NPS   MPL--NIRLFLAKLIV   2162
Query_GECKO  2101   QHECMASMTTLIKHM   HRNNI--TP   T   VEKGAVLPDLPPWMKFLHSKLG--NPS   VPL--NIRLFVAKLIV   2162
Query_DANIO  2098   QHECMANMTALLRHM   QRNNI--TP   K   VEEGVRPSDLPPWMKFLQGKLD--NPS   TPL--NIRLFIAKLII   2159
Query_SNAKE  2100   QHECMTSMTTLLKHM   HRNNI--TP   V   VEKDTNSPELPPWMKFLHNKLG--NST   VPL--NIRLFVAKLIV   2161
Query_STERL  2113   QHECMASMTALLNHM   QRNNI--TP   K   VDEGTVPADLPPWMKFVHGKFG--NAS   THL--NIKLFIAKLII   2174
Query_MILII  2099   QHECMPAMIALINHM   QRNNI--TP   Q   VEEGKTPQELPPWMKFLHGKLG--NSS   IPL--NIRLFIAKLIV   2160
Query_BAMBO  2097   QHECMATMTALIKHM   QKTNI--TP   K   VEEGKTPHELPPWMKFLHGKLG--NSA   TPL--NIRLFVAKLIV   2158
Query_ALLIG  2059   QHECMTAMTTLIKHM   QRNQI--TP   Q   VDEGAVSVALPPWMKFLHSKLG--NPS   VPL--NIRLFLAKLIV   2120
Query_GAR__  2102   QHECMASMTVLINHM   QRNNI--TP   E   VEEGTVPSDLPPWMKFLQGKLG--NPS   TQL--NIRLFIAKLIV   2163
Query_TETRA  2085   RHPCMPTVLTNIDKM   ---------   -   VQQFPLTEAMPQWMTNLLEEVK--KDS   CSL--GQKIFIIKIVL   2138
Query_PARAM  1941   KHKLMIPMLSLIDFM[6]-QQQIagQQ   Q[1]QQTQQQEEEMPSWMTAIKQVFN--DDS   LYL--SLKIFMLKLII   2010
Query_SLEMN  2072   KHECMKNIKRCIDKM   -SE------   -[3]REWDQNTNEMPKFMQNLCLEIK--RPD   IHL--NFKIFILKILI   2130
Query_STENT  1805   TLTIMKPLIEVIERM   -EN------   -[2]----ADSQEMPQWMQCLHSDLS--NNS   NFL--AIRLFIIKIIL   1858
Query_NEMAT  2083   SHECLTKICSLLTHM   KESNI--SP   Q   PDKGALPKDMPSWMSALHKKLT--NFE   THL--NIQLFIAKVII   2144
Query_PISTI  2121   RHESMAKMISLLKHL   VDKKI--SP   E   QVKDVPPTDMPAWMAPLHKKIA--SSS   THL--NIRLFIAKLII   2182
Query_CORAL  2099   RHDCMTTLVGLINHL   VKNKI--SP   Q   PEQGKEANEMPPWMSYLHKKLS--HHS   SEL--NVRLFIAKLIT   2160
Query_HYDRA  2000   QHECMSKYISLINHM   -RNKN--IY   I   QPNDKEVKDMPPWMTYFHGKIT--NKD   TEF--NVVLYIARIVI   2060
Query_LATIM  2101   QHECMAAMTALINHM   QRNKI--TP   Q   VEEESGPVELPPWMKFLSHKLA--NPA   TPL--NIRLFIAKLVI   2162
Query_SHRIM  2061   GHEVMANLTAVINHM   VDAEImvLH   K   GDTVPKATDVPQWMTLVQQKLD--SAD   TPR--NVKLFLLRVMM   2124
Query_MCRAB  2056   EHEVMANLTAAIYHM   VDAKImmLH   K   EDSTPSAAEIPPWMTIIQQKLE--SGE   THR--NVKLFLLRLMM   2119
Query_AMPHI  2077   SHEVMPELTAVLHH-   -------IH   R   LTEDKSQSSPPLWMTEILKELS--NSK   APR--NVRLFLLRLLT   2132
Query_BARNA  1958   EHECMACATALVGHL   -AERV--WP   L   PEVGPPAQPAP-WVTALCARLQ--SDN   TPD--NVLLFLCRLLT   2017
Query_COPEP  1979   AHPCMAQLIAVAEHM   TACNI--YS   T   PTENFSNTTLPSWMQSIHDVLV--EGS   TPR--NVKLFMLKFIY   2040
Query_DAPHN  1893   KHSCMSLLVNFLNNF   YANQT--LL   -   ------ANSEPKWMITLREALL--NPN   AHN--NVVLFIIRLVV   1947
Query_STARF  2141   QHECMASLTAVLKHM   QANKI--TP   E   IPEGSKASEMPGWMASLQKKLT--SQS   THV--NIKLFIARLII   2202
Query_STRON  2127   QHECMAAMTGVLKHM   TNNKI--NP   D   ITKATKAEEMPQWMSGLQKKLS--SQS   TEH--NIKLFIVRLVI   2188
Query_SEACU  2102   QHECMASVTAVLRHM   MENKI--TP   E   VTQAKQSSNMPPWMKMLHGKMS--SSE   TSV--NIKLFVARLIV   2163
Query_FUNGI  2025   QHPCMSVMLDLLDHM   YKLFQ--QD   F   ------SKEVPKFMQEMHSKL--sSKD   SSP--NTLLFILKTIV   2080
Query_DPURP  2158   SNPSMIAMLKIIDFY   QNKFV--IP   T[8]TQHKLTPSDMPKWMYEIYQKL--qQG-   VHP--NIVIFMIKIII   2226
Query_HALBU  2165   SNPSMQPMLKIIDIY   QERFR--ES   I   ---DQSPNDMPKWMNELYLKVK-nDTN   LHP--NILIFILKVII   2224
Query_GLOBO  1950   ENICMRGILSVIQRL   SK------P   G   ADQE-----MSLWMKSLHGKME-kPD-   AHI--NVRLFIAKIII   2002
Query_PALUS  2003   HNICMSRIVDIVKKL   NAV-----P   N   TDGK-----TPTWMQELLKKMS-eTD-   TPL--NVRGFIAAIVM   2056
Query_CALIF  2080   DNICMPIILKIIKKI   SNNSM--KA   P   TQET----TIPAWMNAILKTFT-sYA-   THI--NVKIFIAKLIM   2137
Query_GLOMU  2134   SHPCMKMIIRVIDRL   HSQLT--KP   V   DPVE----SMPGWMNDIYKKFT-nIGN[6]THI--NIRLFIAKIIV   2198
Query_LOBOS  2065   SNPCMKMILLVINEL   HTSIT--PP   P   KEMAKEDSAMPSWMKDIHRKLV-nPN-   TNL--NVRLFLAKIVI   2126
Query_RACEM  2051   KNPCMPFLVQTIKRL   HSTIS--PP   D   PTKT-----VPAWMNTLLSTFK-kAEN[2]SGL--FLRLFLAKVII   2110
Query_PVIOL  2157   SNPSMIALLKIIDYY   QLKFV--ST   T[1]TIGEPQLSDMPKWMTEIYNKI--nDSD   LHP--NIIIFMIKIII   2219
Query_BOMBU  1934   NHECMASICGVLSHM   IEEKI--SV   -[1]PTDDTITV--PKWLKCFLSSFT--TTN   YN---NVRLFMLKIIL   1992
Query_AEDES  2076   DHEVMATICAVIQHM   HENKI--TP   -[1]VEGARATP--PPWVKFLAASLR--DQQ   QHK--NVRLFLAKVID   2135
Query_ANT__  1974   EHECMPYICVLLRHI   --RKI--FR   -[1]SDKE------PEWLRLFLEIMQ--CDN   Q----NIQLFMLKIIS   2025
Query_WHITE  2029   EHECMPMLTAVINTM   KDAQL--IQ   -[1]PPEIIGSSDLPNWIKPIHQALI--NQN   VHK--NIKIFLCKLIL   2090
Query_BEETL  1935   THECMANVCGVIEHM   FRAGI--SE   -[1]PDEE-EQYELPAWLQAIRTLLL--DEN   TAV--NVKIFWVKVID   1995
Query_THRIP  1970   DHECMPTVCSIIRHL   VDCGI--SP   -[1]PVQGAKPT-LPEWMDLLKNSLT--DLS   NPR--NSQIFILRVIL   2030
Query_TERMI  2046   MHECMATLCGVIEHM   VDQGI--SP   -[1]PDAETKDAALPPWMESMRKCLS--DYS   QPR--NIRLFLLKIIL   2107
Query_FLEA_  1962   DHECMSIICGVIGHM   VTKGI--SP[4]Q[8]TKQGSSVPVMPSWMEALVQVLA--NAE   TPV--NIRLFIMRIID   2035
Query_MIDGE  2022   NHEVMSVLVGVVNHM   HKCRI--TP   -[1]MDFEKCDEKKFEWVLSLAEPLR--SRN   VHK--NVRMFLAKLIE   2083
Query_LAODE  2009   SHPCMPSLCGTITNL[9]STSDA--TP   K   -----VDGGTPVWISCISDALK--LSS[5]EEI--NAKLLLIKLLY   2079
Query_LOTTI  2069   QHECMATMVALIKHL   H-TVI--TP   V   LKGTIPK-EMPTWMSYFSEKMT--NPA   TSF--NIKLFIARLIF   2128
Query_APLYS  2122   QHECMAPLIALLRHM   VATKI--SP   E   PTAGAPAKEMPAWMPYLHDKMS--SPS   SHV--NTKLFIAKLVI   2183
Query_SCALL  2120   KHECMATLIALLQHM   ERAGV--TP   P   VTTGTAA-TMPGWMDYLHTKIK--SAK   THV--NVKLFIAKLVS   2180
Query_CRASS  2088   QHEGMANMIALLKHM   QRAKI--IP   L   VESGVVP-ELPAWMKNLNDKLR--SPR   THN--NIKYFITKLII   2148
Query_OCTOP  2039   QSECMLPVVAVLKHM   YQQNI--TP   K   ENKE-----MPPWMNYLRKKFI--ASS   THL--NIKLFIAKLIV   2095
Query_POMAC  2086   SHECMAPLVALLKHI   HKNNI--FP   A   AQSDASSQDMPQWMVFLHDKMTknTGS   DTL--NIRLFIAKLII   2149
Query_WHIPW  2004   AHPCMFSMVSVIQCV   VEKFG--NI   G   ----AEKGSMPRWMLFMKDFMQ----F   GSI--NVKIFIARLLI   2059
Query_TRICH  2021   QHPCMQAMVNVINYM   DHMGM--IP   Y   ----EAGTPLPAWMASMRDFLN----F   GSI--NVRLFIARIIV   2076
Query_AASTA  1894   SHPCMLPLLKTIKHM   -ERLF----   -   -QSQWTDAMMPGWMDKVWFNLS---TN   SST--SVRLFLCKVVL   1949
Query_AEUTE  1837   MHPCMIPLMQTIKHM   -ETLF----   -   -HSQWSNTAMPGWIEKVWFNLS---SH   SST--ATRIFLCKLVV   1892
Query_PLASM  2005   QHPCMIPLLRVLVQM   -KTEF----   -   -GSNWNAKVMPGWMQRIFDVIV--DLS   TGL--NIRLFLAKVVL   2061
Query_BREMI  1991   QHPCMIPLLRVLVQM   -KTDF----   -   -GSSWSTNSMPGWMKKIFNVIV--EAS   TGL--NVRLFLAKVVL   2047
Query_OLIGA  2015   DHPCMIPLLRVLLQM   -KTDF----   -   -ADSWDGKTMPGWMQKIYNVTA--DST   TEV--GVRLFLVKIVL   2071
Query_LIGNO  2193   DHPCMATMVALLRHM   KAFKL--IP   Q   LPEGVKPTQMPKWMQNMLQKVS--NSS   TDShpNVRIFLMKLLI   2256
Query_CESTO  1974   MTSLMIAFVGLLKRM   SQLGL----   -   ----LKPNEMPQIIKFLYDQFN--TNQ   ENC--NVQVFVVRLVI   2028
Query_OVI2_  2333   REPLVSGMTGLLKHL   NRLPL--ST   E   DTTVDLTRGDPFWSGYLHGQMM--RSS   VPV--NVQTCIVKIID   2394
Query_PRIAP  2113   QHECMASMVALINHL   ATTVC--PI   Q   A--ASSPAHMPAWMELLHGKFS--HVE   TPL--NVRVFIARLVV   2172
Query_ROTIF  2069   QHESMDNLISLFQSL   VVNKI--TP   I   YDNGQLPSEMPPWMTFVHKKIL--DVY   THE--NIKLFLIRAIA   2130
Query_SPONG  2109   IHECMTPLMKLLDHM   IINKI--TP   A   PSGDS--PEMPPWMTPLHKRMT--DSR   TGR--NIKLFIVRLIT   2168
Query_TARDI  2101   QHEVMATLVGLVQHM   VEKKI--TP   T   VKDPLDNPSPPAWISALAGELK--SST   AHL--NVKLFLAKLIM   2162
Query_TRICH  2082   KHECMEALLTLLDHM   AVNKI--NP   A   IEKDTTPSEMPSWMVALRKKAS--DHT   TNR--NIKLFIAKLII   2143
Query_OSTRE  1978   AHPVSDGVFRILELV   TKEGL----   -   --TGGNSEELPPSIQLLEAISSllYTP[2]LPA---VKLFVVKALL   2037
Query_NITEN  2202   QNPVMAVLLRVLEHL   HQKQG--GA   -   -----EPAEMPPWMDAIHKSF--tSPA   TPI--NVKLLLAKAIV   2257
Query_CLUBM  1966   QHPCMGVILRLIDHL   HSRFG--ST   -   -----SSQQMPEWMKSIHATL--eSSS   THA--NVRLFLGKVVV   2021
Query_MOSS_  2086   QHECMPVILRLAEHV   HVKFG--AN   Y   -----QPGVMPDWMKSLHATL--dNAA   THI--NVRLFLAKVIT   2142
Query_LIVER  2115   QHPCMEVLLKLIERL   HSKFG--SG   -   ------NREMPEWMKALHSTL--qNAG   THI--NIRLFLTKVVV   2169
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT   2151 NGESV    FRPYASFWFQPLVALAT   N---    --MGG-LNYFSLDIIVTLL-SWHE--I   AIP   EE--TTV  2202
Query_LINGU   2150 NAEEV    FAPYAKFWLGPLVQLMV   R-DM    P-DEG-LNYYVVDIMVTVL-SWSS--V   AVP   QD--TAM  2204
Query_LIMUL   2140 NNAQI    FKPYAGSLLDPLIQLVV   SGTL    G-QE--INYFILDVLVVLM-TWAT--V   AIP   ED--TVI  2194
Query_SPIDE   2137 NNAEL    FQPYAEFLMGPILNLIV   LGTA    G-FS--LNYFILDVIITTL-SWAS--V   AIP   QG--NDM  2191
Query_SCORP   2123 NNAEI    FKPYAVFLLQPLINLIL   NATF    G-RE--INYFILDLMVVLL-SWSS--I   AIP   ED--SSI  2177
Query_TICK_   2059 NSAEV    LRVHSRHLVPALLNVLA   -GDK    L-GSE-LDAFVIDVAVTVL-MWASdwS   LED   RK--TVG  2115
Query_MONOS   2091 NQPQH    FQPWAKWWLAPLMRMAV[2]-GDR    Q-AEP-FNYFVTDLCVVLL-RWSEt--   AVP   -G---AS  2145
Query_ROSET   2059 NMPDA    FQPWSSRWLRPIMLLIL   KFQD[ 1]PeTPG-FNYFVTDLCLLLL-TWADt--   AVP   QD---DS  2115
Query_BELCH   2186 NNEEV    FRPYARFWLSPLMEMIV   RGEA    S-GEG-INYFIVDLLVTML-SWGP--V   AIP   ED--SAM  2241
Query_CIONA   2149 NCEQY    FRPYAAHWITPLAQLIL   REDC    G-VNG-IDTMVCDILVTVL-SWNN--I   AVP   ED--SVE  2204
Query_CAECIL  2165 NTEDV    FRPYAKQWLGPLLQLVV   SGNN    G-GDG-IHYMVVEIVVTVL-SWAN--V   ATP   AG--LAK  2220
Query_HUMAN   2152 NTEEV    FRPYAKHWLSPLLQLAA   SENN    G-GEG-IHYMVVEIVATIL-SWTG--L   ATP   TG--VPK  2207
Query_MOUSE   2149 NTEEV    FRPYAKHWLSPLLQLAV   CENN    R--EG-IHYMMVEIVATIL-SWTG--L   ATP   TG--VPK  2203
Query_XENOP   2172 NTEEV    FRPYARFWIGPILQLIV   SGNN    G-GTG-IHYMVVETLVTIL-SWSS--I   ATP   NG--QAK  2227
Query_PLATY   2156 NTEEV    FRPYARHWLGPLLQLVV   SGEN    G-GEG-IHYMVVEIVATVL-SWTS--I   ATP   TG--VTK  2211
Query_TASDE   2158 NTEEV    FRPYAKHWLNPLLQLVV   SENN    G-GEG-IHYMVVEIVVVVL-SWTS--I   ATP   KG--VPK  2213
Query_CANAR   2154 NTEDV    FRPYAKHWLGPLLQLVV   SGNN    G-GEG-IHYMVVEIVVTIL-SWTS--V   ATP   KG--NIK  2209
Query_OSTRI   2124 NTEDV    FRPYAKQWLGPLLQLVV   SGNN    G-GEG-IHYMVVEIVVTLL-SWTS--V   ATP   KG--NIK  2179
Query_TURTL   2163 NTEEV    FRPYARQWLGPLLQLVV   SRDN    G-GEG-IHYMVVEIVVTLL-SWTS--V   AAP   KG--NIK  2218
Query_GECKO   2163 NTEEV    FQPYAKQFLGPLLQLIV   SGDN    G-GEG-IHYMVVEIVVTLL-SWTS--I   ATP   KE--NIK  2218
Query_DANIO   2160 NTEEI    FRPYAKHWLGPLMQLVV   SSSN    G-GEG-IHFMVVDIVVTVL-SWAS--V   ASP   KG--NTR  2215
Query_SNAKE   2162 NTEEV    FQPYAKWWLGPLLQLVV   SGDN    G-GEG-IHYMVVEIVVTLL-SWTN--V   ATP   KG--NIK  2217
Query_STERL   2175 NTEEV    FRPYARHWLGPLLQLVV   SGNN    G-GEG-IHYMVVEIVVTVL-SWTS--L   ASP   KG--NTR  2230
Query_MILII   2161 NSEEV    FRPYAKEWLGPLLQLVV   SGNN    G-GDG-IHYMVVEIVVTVL-SWTS--V   ATP   KG--NMK  2216
Query_BAMBO   2159 NTEEI    FQPYAKEWLGPLLQLVV   SGNN    G-GDG-IHYMVVEIVVTVL-SWIS--V   ATP   KG--DLK  2214
Query_ALLIG   2121 NTEEV    FRPYARHWLGPLLQLVV   SRDN    G-GEG-IHYMVVEIVVTLL-SWTS--V   ASP   KG--NVK  2176
Query_GAR__   2164 NKEEV    FRPYAKHWLWPLLQLVV   SGNN    G-GEG-IHFMVVEIVVTVL-SWAS--V   ATP   KG--DTK  2219
Query_TETRA   2139 NKPLL    FEKYSQQWFDVLVDYAK   LNDN    G-SKF-FHYFLRDVTTTLI-DWCSk--   -KP[4]RRi--AS  2197
Query_PARAM   2011 NKSDI    FNRYASGWIGVLLDYCT   KENK[ 2]G-GKG-FHYFLRDVIVVLI-KFSQk-L[2]RQE[2]QM-hPQS  2074
Query_SLEMN   2131 NHPQL    FSAYAKEWFEPIANYLC   DKN-[ 1]G-GKG-FHYFMRDLCTLLI-SWNYi--   ---   PE-dNTK  2184
Query_STENT   1859 NRPQS    FKKWADLWLGPICT-VM   SGN-[ 1]G-GTG-LHYFLRDVCTLFLyDWDNi--   ---   SI-kNKE  1912
Query_NEMAT   2145 NQSQV    FEPYARFWLSPLMELIL[4]EEGE    N-KRG-LNYFVVDLLVTIL-SWAT--T   AIP   ED--SYG  2204
Query_PISTI   2183 NEAKI    FEAHAKFWFNPLMELIL[4]TEPA    G-TEG-INYFIVDLVVTLL-SWST--T   AIP   ED--SHN  2242
Query_CORAL   2161 NLPEV    FEPYAKFWLDSLMELIL   LPTF    K-NSE-LNYFVIDIVVTML-SWSS--V   TIP   ED--SYS  2216
Query_HYDRA   2061 NEPKL    FQPYAQFWFLPIIELLL   SQRI    L-GDY-LNYFVVGLIETML-SWSP--S[3]NTA   SG--SI-  2118
Query_LATIM   2163 NTEEV    FRPYAKHWLGPFLQLVV   SGNN    G-GEG-IHYMVVDIVVTVL-SWAN--I   AAP   KG--NTR  2218
Query_SHRIM   2125 NNARI    FEPYAAHFVSTILTCLV   DGTI    G-DK--INYFFCDVMVMVM-GWGSk--   AVP   QD--TVM  2179
Query_MCRAB   2120 NTSRV    FAPYAVHFINPVLTCLV   DGTT    G-DR--INYFFSDLIVMVM-GWGKr--   AGS   QD--STM  2174
Query_AMPHI   2133 NTAQL    FAPYARHASRAVLQCIT   DGTT    G-VG--FNYFLADLVVLLV-GWGVd--   GAP   ND--PD-  2186
Query_BARNA   2018 NCDSA    CRPYGRWLLPPLLRLLA   AERL    G-PEV--NALLLDVLDMLA-GWAQq--[1]ITP   EE----S  2071
Query_COPEP   2041 NCSEI    FLPFCINWYTPILECIV   DGTL    S-PEGkLNYLSLDLLTMTL-YWSEk-T[1]TVP   -----TQ  2096
Query_DAPHN   1948 NLKSI    FRIYANFWYKALLVALL   -RKI    G-GEP-LNSLFREALDMML-EWSNe--   APP   SA--NDV  2002
Query_STARF   2203 NNNEV    FTPYARFWLSGLLQLLI   SGNT    G-GEG-IHYMVVDVVVVLL-GWVT--V   AIP   ED-----  2255
Query_STRON   2189 NAQDV    FQPYARFWLSPLMQFLA   GGQQ    G---G-INYMVLDIMLVML-DWST--L   AVP   ED-----  2239
Query_SEACU   2164 NCSKV    FEPFAKFWLNPLIQLII   GGEI    S-SPG-IHYMLLDIMVVLF-GWSE--K   AIP   ED-----  2216
Query_FUNGI   2081 NRPLY    FKAYSSYWIGAIIQALN[4]SGTA[11]A-GKT-INYCMRDICTLII-EWNVk-L[3]D--   ------S  2149
Query_DPURP   2227 NRPSY    FERFHELWIPILMDYI-   ----[ 5]G-GQG-LHYFVRDVCFILL-KWPNi-Y[3]EGE[1]RP--PTS  2287
Query_HALBU   2225 NRQQY    FYRYHKQWTPILIEYV-   ----[ 5]G-GQG-IHYFIRDIAMLLL-SWPNi-Y[3]K-L[1]QL--DTL  2284
Query_GLOBO   2003 NVPNI    FAPFAKYWWRPIAQLV-   ----[ 5]F-GSG-INYFVQDLCLLLM-EWTKn-A[3]TKA   ------D  2058
Query_PALUS   2057 NVPEA    FSANAGEWWKCFAELI-   ----[ 5]F-SKG-VRGFVQDLCLQLL-KWGGr-T[3]ETK[5]DD--DHD  2121
Query_CALIF   2138 NYSDP    FIPYVNSWYIPLVKFI-   ----[ 5]F-NDG-FTYFIEDICNLLI-KWKN---   ---[4]EN--NSE  2193
Query_GLOMU   2199 NVPEA    FEKYASLWIRPVIKLI-   ----[ 5]Y-GEG-INYFIQDLCAIVT-VWGS---   ---[4]-E--SYQ  2253
Query_LOBOS   2127 NIPEA    FEMYAHSWIRSLMKLA-   ----[ 5]Y-GEA-MNYFVQDLCVLIV-VWGG---   ---[4]ED--SYD  2182
Query_RACEM   2111 NLPAA    FEPYADFWIRPLMRLV-   ----[ 5]Y-GDP-MNYFVQDLCYLIL-AWGK---   ---[4]ED--SYE  2166
Query_PVIOL   2220 NRPQY    FERFYKYWVPILLDYV-   ----[ 5]G-GSG-FHYFIRDVCLILL-SWPNl-Y[3]D-K[1]PV--AFS  2279
Query_BOMBU   1993 NM-Q-[ 2]FQPYAKFFLQPIMHATY   SYLK    R--NQ-LNYIIVDVIEMLI-DWQTs-S[4]EST[2]QN-kDEI  2055
Query_AEDES   2136 NC-R-[ 2]LKPYAATLIPPLMQVIV   DECI    A--PQ-LNTFVTDLIALIL-EWGDq-S   -YQ[1]AS---VD  2190
Query_ANT__   2026 NTAD-[ 2]FKPYARFALTQIIKTVT   KYLQ    R--HN-LNYIITDILEILM-DWYDv-A   -VP[2]EDgkAAA  2085
Query_WHITE   2091 NV-S-[ 2]FAPYASHFILSISEAIA   AKCF    G--SE-INYFVADLVTLML-SWRKi-A   -VP[2]KK---DE  2146
Query_BEETL   1996 NT-I-[ 2]FKHFSKYFTEPLLKFIV   DRSA    G--EV-MNYFVADVVAVLA-TWCTe-V   -TL[2]QT---EI  2051
Query_THRIP   2031 NC-E-[ 2]FSTYAMVWMEPMLDCVL   RICL    Q--NG-ISYLLGDLVCMLA-KWNAe-H   -QT[2]LN---KS  2086
Query_TERMI   2108 NV-E-[ 2]FHPYAHHWMNSILKAIT   DGVV    G--RA-PNYLLSDTVAMLA-DWCRk-T   -TS[2]AK---SD  2163
Query_FLEA_   2036 NC-Q-[ 2]FRPYAHHILTPILELLV   DECA    G--ST-INHFVADTVALML-SWHDv-A   -IP[1]ST---SI  2090
Query_MIDGE   2084 NC-R-[ 2]FVYYAKFLLGPVMSVIA   DKCC    G--LK-MNFFITDLVTMLL-SW-Sp-V   -YK[1]TT---MS  2137
Query_LAODE   2080 NCNET[ 4]FTPYAKVLLPPIIKFII   EYKV[ 3]--DNK-LNYMIGDILGMLT-SWSSv--   ATP   ET--N--  2139
Query_LOTTI   2129 NTSEV    FQPYTKFWLRPLCQLIF   SGPL    S-SDG-INYFNVDLIVMMM-SWHT--T   SIP   QD--TAE  2184
Query_APLYS   2184 NCTKV    FQPFAKFWLRPFCQLLF   SPEV    S-GSG-VTYFAVDLIVTMV-SWHT--V   AIP   ED--SAE  2239
Query_SCALL   2181 NCTQI    FQPYARYWLRPLSQLIL   SGAL    G-SQG-INYFVVDVVVTMM-SWHK--M   AIP   QD--TAE  2236
Query_CRASS   2149 NTAEI    FRPYARDWIRPLTQLII   SGLF    ---TG-LNYFVIDVIVTIM-SWHQ--V   AIP   QD--TVE  2202
Query_OCTOP   2096 NTAEI    FEKYASFWVLPLSEFII   LGNF    E-EKG-INYFVLDIIVMIL-SWGE--H   SIP   GD--TSK  2151
Query_POMAC   2150 NTSEI    FQPFARLWLRPLCQLIL   STSL    C-DNG-INYFIMDLLVTML-SWHS--T   AVP   QN--SST  2205
Query_WHIPW   2060 RLEQI    FLPYCYEWFDPVAQLVL   NDQL    A-EAG-LHYFRFDLLLLLL-QWIAe-G   ACQ   PS--KPS  2116
Query_TRICH   2077 QCFHV    FRPFYLEWISPIIKLIV   MDEF    G-SNI-LHYFRVDLIVVLF-EWTAh-L   KED   EA--IP-  2132
Query_AASTA   1950 NRPML    FAPYAAKFVGPLVE--L   --ML[ 3]K-RDE-FHYLLRDVCHVVVdTWQ----[1]-DT[1]SD-----  2001
Query_AEUTE   1893 NRPGI    FEPYAAKFVGPIVE--L   --IV[ 3]K-LDE-FHYILRDICHLVVdTWH----[1]-ET[1]DY-----  1944
Query_PLASM   2062 NVPSV    FSTYSSLWLPAVMKALL   --D-[ 5]K-VSE-FNYILRDCCNLVLnSWKDv-V[1]-SS[1]DT-----  2119
Query_BREMI   2048 NVPDV    FFLYSESWLRAVIETLL   --D-[ 5]M-APE-FNYILRDCCSLVLdSWQNv-A[1]-TP[1]ID-----  2105
Query_OLIGA   2072 NVPDV    FAKYADKWAEKLMEVII   --EA[ 7]K-KAE-FNYMIRDCCHLLAdEWKNv-S[1]-GS[1]SM-----  2132
Query_LIGNO   2257 NCQDI    FQPYAHFWLRPMCQFAL   SDRN    G-GQG-FHTLLVDIVYFFV-AWSG--T   AIP   ED--DLG  2312
Query_CESTO   2029 NCPEV    FKPFYRRWLPLLLTYAS   SSTE[ 4]S--SG-LSSLAIELCLLLA-DWSR--S[1]ILP[1]TE----M  2087
Query_OVI2_   2395 NCPEV    FQPYGRTWLQVLLQFLS   SNLT[ 7]P-ART-FTTLFTDICLMLA-RWADp--[2]SNP[2]PQ--TPT  2461
Query_PRIAP   2173 LCAHA    FRPCAASWLAPLMHLVV   NGNC    GvAAA-IHSLTADVVVVML-SWTP--V   ARP   QD--NMT  2229
Query_ROTIF   2131 NTESI    FKPYCKFWYSALIGFLV   NSSL[ 2]D-LAQ-MDCFTLDLMVMLL-SWGS--I   QKP   DQ--S--  2186
Query_SPONG   2169 NRPKI    FQPYAKHWLPVLTQFIL   SGDS[ 9]GgENG-LDYFTVDAIATML-SWSS--T   AI-   -----IE  2230
Query_TARDI   2163 NVENV[36]FYQFAELFYGPLLDLLM   EIDC    --TEG-MNYFAVDLIHLLL-KWSV---   KTP   LP--G-G  2251
Query_TRICH   2144 NRPKI    FRPYAKFWLVPLIQLIV   -EGK    E-SDG-IHYFIVDLMIILL-SWAD--V   AIP   KD---VD  2197
Query_OSTRE   2038 RLNHE[29]------SLLTATIELVK[1]SGVD[ 6]R--ET------LDVALECKdLWNSk--[3]AKP[3]DA-----  2121
Query_NITEN   2258 RARKL    FQPFKREWFAPLIEIF-[1]AD--[ 4]G-GAS-FHYFLRDVCIVLL-QWDLa-K[3]P--   -------  2312
Query_CLUBM   2022 KARKV    FARFSDTFFQPLVNVI-[1]HD--[ 4]G-GFS-FHYVLRDISITIL-QWDVk-E[3]R--   -------  2076
Query_MOSS_   2143 KTRKI    FAPHAADWIRPYVQTI-[1]ID--[ 4]G-GLK-FHYMLRDICVTIL-QWNVa-S[3]G--   -------  2197
Query_LIVER   2170 RTWRV    FAPFAQVWFIPVVQTI-[1]ID--[ 4]G-GMA-FHYMLRDICITIL-QWNLp-C[3]P--   -------  2224
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2203  D-IAMTSHLVKYIV-QN-CYHQ-NR   QIMRNNL   EVLKTLVEAWKST--IR    IPTQDIYENFI---DREDKGK  2263
Query_LINGU  2205  D-RAMASRLVEFVV-RN-VNHD-TR   QIFRRNL   ELLKTLVECWKGR--FD    MPTKIIYDNF--s-DRDPKSK  2265
Query_LIMUL  2195  G-RHAANSILEFLM-TK-CAHP-RK   DIFRNNL   EVIKTVVECWKPR--LE    VPTRIICGQMC---MTDITRK  2255
Query_SPIDE  2192  L-K--ASEVLKFVM-MN-IEHS-RR   DIFRNNL   ELVKTILECWKS---LK    IPYNVIYEKMS---DFSPESK  2249
Query_SCORP  2178  G-RHAASGLLEFLM-SK-CFHI-RK   DVFRNNL   EMIKTVVQCWKER--LD    IPTYIIFESMR---NLDTKVK  2238
Query_TICK_  2116  N-DRDLQGFSELLA-QR-CGHE-RR   EVLKNNL   ELLRAWVETWRER--LR    FPADVVLEQLE---QCSLDSK  2176
Query_MONOS  2146  D-VIVVQQFIEFLV-RH-AQHE-SG   PIFRNNL   TIIKTFVEKWGAQ---T[ 3]LTMEVLCEFAAhs---GISTK  2208
Query_ROSET  2116  E-KRLARQLVQFLM-EN-ARHE-TT   AVFKNNV   NIIRLLLERWKDA--VQ[ 2]VPFKVIVDNMSsr---QMASK  2178
Query_BELCH  2242  G-RALTSRLLHFLM-AN-AFHK-TR   AVMRNNL   ELIKTAVECWKTR--VE    VPTKVIHDQLC---GSDASSK  2302
Query_CIONA  2205  G-KSLVRGLLEQLS-RR-IHHN-NR   AIIRHNL   DLIKTTLECWGER--VR[ 2]VDPNIVYQLCE---SNDVDLK  2267
Query_CAECIL 2221  D-EILANRLLEFLV-TH-SFHE-KR   AVFRHNL   EIIKTVVECWKDC--LS    IPYRLIFERFS---NGNTNSK  2281
Query_HUMAN  2208  D-EVLANRLLNFLM-KH-VFHP-KR   AVFRHNL   EIIKTLVECWKDC--LS    IPYRLIFEKFS---GKDPNSK  2268
Query_MOUSE  2204  D-EVLANRLLRFLM-KH-VFHP-KR   AVFRHNL   EIIKTLVECWKEC--LS    IPYRLIFEKFS---HKDPNSK  2264
Query_XENOPU 2228  E-EILANRLLEFLM-KN-VFHE-KR   AVFRHNL   EIIKTVLECWKEC--LS    IPYRLIYEGFS---GTDPNTK  2288
Query_PLATY  2212  D-ELLANRLLEFLM-KH-VFHQ-KR   AVFRHNL   EIIKTVVECWKDC--LS    IPYRLIFEKFA---TGDPNTK  2272
Query_TASDE  2214  D-ELLANRLLQFLM-KH-VFHP-KR   AVFRHNL   EIIKTVVECWKDC--LS    IPYRIIFEKFS---NRDPNSK  2274
Query_CANAR  2210  D-EILANRLLEFLM-KN-AFHP-KR   AVFRHNL   EIIKTVIECWKDC--LS    IPYSLIFQRFS---SGDPDTK  2270
Query_OSTRI  2180  D-EILANRLLEFLM-KN-AFHP-KR   AVFRHNL   EIIKTVIECWKNC--LS    IPYSLIFEKFS---GGDPNTK  2240
Query_TURTLE 2219  D-EILANRLLEFLM-KN-VFHQ-KR   AVFRHNL   EIIKTVVECWKDC--LS    IPYRLIFEQFS---GGDPNTK  2279
Query_GECKO  2219  D-EILANRLLEYLM-KN-AFHQ-KR   AVFRHNL   EIIKTVVECWKDC--LS    IPYRIIFERFS---GGDPNTK  2279
Query_DANIO  2216  D-EVLVNRLLGFLF-KN-CFHS-KR   AVFRHNL   EIIRTVVECWKDC--LT    IPYDLIYERFA---GTDPNSK  2276
Query_SNAKE  2218  D-EVLANRLLEFLM-KN-VFHQ-KR   AVFRHNL   EIIKTVIECWKDC--LS    IPYRIIFQLFS----GDPNTK  2277
Query_STERL  2231  D-EILANRLLEFLM-KN-AFYD-KR   AVFRHNL   EIIKTVIECWKDC--LA    VPYSLIFDRFS---GKDPNSK  2291
Query_MILII  2217  D-EILANRLLEFLM-KN-SFHE-KT   AVFRHNL   EIIKTVVECWKDC--LS    VPYSLIFSRFS---STDPNSK  2277
Query_BAMBO  2215  N-EILANRLLEFLM-KN-CFHE-KR   VVFRHNL   EIIKTVVECWKDC--LS    IPYSLIFDRFA---GRDPNTK  2275
Query_ALLIG  2177  D-EILANRLLEFLM-KN-TFHQ-KR   AVFRHNL   EIIKTVVECWKDC--LT    VPYRLIFEQFS---GGDPDTK  2237
Query_GAR__  2220  D-EVLANRLLEFLM-RN-AFHE-KR   AVFRHNL   EIIKTVVECWKDS--LM    VPYSLIFDRFA---GSDPNSK  2280
Query_TETRA  2198  DkKQNLCQVINKIC--RvLAHK-EK   EIFKQNI   QILKNLLDLWKEC--IY    VDENQLTKMLLle---DKEDK  2259
Query_PARAM  2075  R-NKICSVINNLIM---lSADK-NR   KIFDQNI   EIIGMFVQLYRDSmyIC[12]-------DQKNqq--LDPQAE  2142
Query_SLEMN  2185  N-RLLCTNVVNTLI-KY-SADK-TK   MIFNTNI   EIIAGLLHRWKKLvaI-[ 3]VISMMIGM---------DDKK  2243
Query_STENT  1913  K-E--AGKFLNYLV-KV-GGDL-NK   EIFESNI   KLITELIAKWK--vpL-[ 3]LIDGMLKK------QLKDDDN  1970
Query_NEMAT  2205  D-RHLASRLLQTVM-RR-THHE-NR   AVFKNNL   EIVKTMVEVWRER--VT    VPTKVIYEALS---NPHPDKK  2265
Query_PISTI  2243  D-RHLASRLLGFLM-SR-TRHR-HR   SLFRNNL   EIVKSVVEIWKSI--LH    VPAEVIFNSFS---DPDPVKK  2303
Query_CORAL  2217  S-RRLTSRLLEYLM-ET-CNHP-TR   SIFKNNL   EVIKTFVEVWKER--ID    VPYKLIYERFS---NASVNTK  2277
Query_HYDRA  2119  ----LASSLLDYLV-KN-SHNL-DS   PVLQTNL   YLINALSSSWRDR--VT    CPYKLIYSLFS---FSDQNSK  2176
Query_LATIM  2219  D-EILANRLLEFLM-KN-TYHE-KR   AVFRHNL   EIIKTVVECWKDC--LA    VPYSLLYQRFS---GKDPNTK  2279
Query_SHRIM  2180  G-RNIVGRVLRYLC-MN-TPQT-RA   DIFKYNL   DVVRSVVECWKSL---L[ 1]VPYTVIYSLME---VKDGAER  2240
Query_MCRAB  2175  G-RNMVSRVLRFLC-SN-TPHN-RA   DVFRYNI   DAVKSVVECWRAV---L[ 1]VPYAIIYSLME---VSDGRNK  2235
Query_AMPHI  2187  D-SPLLAKLMCHLI-TS-LPHL-RA   DVFRYNV   DLLRSLVSVWCSA---W[ 2]VDGSAIIELLE---TREGHKR  2248
Query_BARNA  2072  S-RADAGRLLEVVLrLTpSAHR---   GVVRIML   DVLEKLVSGWGDR----[ 2]FPGRLAVQLLTgd-------R  2128
Query_COPEP  2097  N--LVVSRMLTLIM-KN-VYHEeKR   -IFKANI   EMIRLVIEKWKDCcvVS[ 2]--FSVIEDLLEan------QG  2155
Query_DAPHN  2003  EfAQAASSLLQFIL-RN-IENK-NA   EILRDNI   EMVRAIFLVWNKI-sLN    IPYELVEELLS------PDKS  2062
Query_STARF  2256  --KRLANRLLEFLM-HF-CHHS-NR   QVLRNNL   EMIKTLVECWKEI--IS    VPTSVIYRILA---TPDPDVK  2315
Query_STRON  2240  --NRLASGVLRFLM-KN-SYHE-RK   SVMKNNL   EVIKTLVECWKDV--ID    VPTRVIYESLQ---DRSPDKK  2299
Query_SEACU  2217  --KYLASQLLRFLM-LN-VHHA-KR   AILRNNL   EVVKTLVELWKPV--LD    IPTSVIYNYLS---NKDPASN  2276
Query_FUNGI  2150  F-QANI--LLNHLA-ER-CSYP-DR   NISRYNV   SIVKSMIEQWKDY--FI[15]VVQQRIVEMLAda--------  2218
Query_DPURP  2288  L-ENHLSKFMNFLM-KN-TYCP-DR   RILKSNL   NIVKSFTERWRSL--F-[ 1]INKNIIIELLLqk--GDFKSK  2349
Query_HALBU  2285  S-KSNISKFINFLI-QN-SYCA-DR   QILRNNL   NILKLFVERWKGL--F-[ 1]VEKKSIVEALAtp-VTNFKSK  2347
Query_GLOBO  2059  D-SVTILKTMRWLI-RH-CNHE-SK   ATIRNHL   RIVRTLIENWKDF--MI    VPSDVIFELLHakpDGNNKAE  2122
Query_PALUS  2122  T-RTILFKMMESLV-ASvCVAD-SM   PEVRENL   ELIASFIELWPSL--MV    APTRIIYEYIT-----NQKRE  2181
Query_CALIF  2194  N-RFNFSNMMNYII-KY-SNIQ-DG   LSIKVVI   DIINEFLNKFKDF--ID    TPVEILYKFLSnknNSDEQYS  2257
Query_GLOMU  2254  D-KVLIYELMEFLM-KN-SYNE-SR   IVIRNNI   LIIKAIFEIWSEL--IV    VPTHVIHELFSds-VGNVDK-  2315
Query_LOBOS  2183  D-RVLLLSFLSYLM-KN-CYHE-QR   QYLLNNI   DLIKGTFENWSNI--AI    VPTTIIYNNFKn--IKDDKR-  2243
Query_RACEM  2167  D-KTLLFYFTNYLM-EN-AYHE-TA   RIAKNHV   QIVKNAIENWEDR--II    PPTKIMYNQFNpr-DKNAGQ-  2228
Query_PVIOL  2280  T-KQLISKFINYLF-KN-CYCS-DR   RILKSNL   TIIKQFTENWKGL--F-[ 1]IDQSTIVEYLVqk--GDYKSK  2341
Query_BOMBU  2056  Q-QLWETLIQKVIV-TK-TEKI-SK   SIYNYNM   NMVKTVLEVWHPY--LK    LPKN-LSDKMSta-PG-----  2112
Query_AEDES  2191  E-ISLASGLVKFVM-RN-CYHA-RR   EVFRLNL   ELVKRLIEQWKA--vIS    VPVQLLYDMIGpt-QDPESMQ  2253
Query_ANT__  2086  Q-RLFEVLINKVLMkSD-NHRV-DK   ----YNL   GLMKTMVEKWRNY--LR    VPSDFLNQKIIsa-DT-----  2140
Query_WHITE  2147  N-RTTVSELLSFLM-AN-TESF-NK   RVFRRNL   EIVKTVVEVWKSS--LD    IPYQLLLDLLKpe-DNDPKSE  2209
Query_BEETL  2052  Q-LA--SKLVAFLI-NN-LKND-RQ   DIFKYHL   QLIRLLVEAWKTS--IE    VPYDLIYAKLKve-PT---CK  2109
Query_THRIP  2087  Q----LNILFGYLV-KH-CPHS-DK   AVLRYHL   ELIKTMLEVWKNY--ID    PPHDSLLSML----DN-PDAR  2142
Query_TERMI  2164  H-HL-VSQLLYLLI-SK-VHHD-RR   DVFKHNL   ELIRTVVELWKPN--FD    IPVQLLYDMIMkh-DR-KDPT  2224
Query_FLEA_  2091  Q-RNIASQFMYFLM-KN-AYHE-RR   AIYKLNM   EIIKSVVEIWKTC--LD    FPLQYLYLALEnp-VK-DSKT  2152
Query_MIDGE  2138  E-KEDACFIMKFLM-EN-AYNE-RS   EIFRLNL   ELIKKLVETWKEIlvVK    IPSQTLLDLLQkp-LEEEMNL  2202
Query_LAODE  2140  E----IYQLLEFLI-EK-TPND-RK   DVMKLNI   EIMKTVISSWKDN--IS[ 1]IPHEKILDLLSdt--------  2193
Query_LOTTI  2185  E-RAMSSRLVQFLM-AN-IYHK-NK   QIYRNNL   EMLKTVLECWHDR--IE    IPYKIIYNQL-------KSEE  2241
Query_APLYS  2240  G-RAMASRVVSFLI-AN-VFHT-NR   QVFRNNL   ELLRTLLECWHDR--VD    IPYGTIYGLF--k-STDPKSG  2300
Query_SCALL  2237  E-RAMATRVVDYLT-AN-VFSE-TR   PVFRNNL   ELLKTLLQCWRDR--LE    VPYNVIYKLMVes-RSDPKSP  2299
Query_CRASS  2203  D-RAMSSRLVEFVM-GY-VYND-TR   PVYRNNL   ELLKTLLQVWKDR--VE    IPYNIVYKHL--k-QPDETKK  2263
Query_OCTOP  2152  D-RYVASSLVQFII-RN-IYNA-NR   TIFRNNL   KLLEALVGLWKAR--ID    IPYKLLYEMMT---SSDPNTG  2212
Query_POMAC  2206  E-RAMASRLVEFLV-RN-VHHE-TR   AIFRNNL   EMLKTLLECWKDC--IE    IPYQEVYNQL--k-MPDPSSK  2266
Query_WHIPW  2117  S-KALCGRLLEYTI-RY-SGSK-NY   HVQKRNI   QLVETMVGLLKDI--VA    VDFKILLDHHC---VA-DTDG  2176
Query_TRICH  2133  --KDSLNDLLKFCV-KH-CWSK-SG   KVFKRNL   RIIETMLVLWKSF--VT    LDPSLILEKIC---FPLGNEN  2192
Query_AASTA  2002  ----DLSRFVNHLM--AvSPHP-ST   FILRDNL   YLIEAFVTKFPTQcrFL[ 2]---DILLDMIN-a---D--DA  2057
Query_AEUTE  1945  ----DLSQFVKHLI--AvSPNT-SA   AILRDNL   YLIESFLFRYPSQcrFL[ 2]---DAIMDMIG-g---DGNDK  2002
Query_PLASM  2120  ----TTSRFVNTLV--KlCPDR-NN   LVRDSNV   LLVTELIALWKDAtyIN[ 2]---ELITFMMS-g---DEDTK  2177
Query_BREMI  2106  ----TASQFVNELI--HlCPNR-NK   LVCDNNV   LLVTKLIALWKNLiyID[ 2]---ELIIHIIS-e---DGDIK  2163
Query_OLIGA  2133  ----MASRFVNHLI--NlSPNEeSN   FIMRDNI   RAVSELIALWKDHvrID[ 2]---EVVKLLFSse---DNNPK  2192
Query_LIGNO  2313  D-RDLVERLLDKLT-GL-CYDK-ER   RIFKNNW   RLLDALIRSWRSR--VH[ 1]VPYKRIFDLLQ----QTEEYT  2373
Query_CESTO  2088  E-KMAAQGVLVFLVkYAwIRASeEG[9]PALKHNL   ELIQLLAESWLPA-gVQ    IPYRDLLAELQlr-----ADN  2159
Query_OVI2_  2462  E-RDLATQLLGMLF-EHlCGNR--E[9]QIPRSAE[6]ELFKLLIECWEET--VQ[ 1]-PCQAVHDALCrt---DVSTK  2537
Query_PRIAP  2230  E-RSLASRLLEFLM-RN-ARHA-TM   SVFRNNL   EIIKTLVECWRDR--VD    VPSQVIYENMC---AGDITMR  2290
Query_ROTIF  2187  E-AKLINRLFTLLI-KR-CYHE-NR   AVLKNNI   ELIKTMTECWRDM--IQ    VPVELINNFLT-----SNDHR  2245
Query_SPONG  2231  D-RFLGSQLLEYIM-LR-SNHK-NR   SILRNNL   EIVRTMVELWKDK--ID    IPTRVLYDQFS---NSDKSTV  2291
Query_TARDI  2252  D-QFRAQRVFLFLI-QH-VKHS-NS   QVQRNNL   TLLKVLTEAWKAQ--LE    IPPRGTIPDLL-i-KDKTKAK  2313
Query_TRICH  2198  E-ILQANRLMEFLM-RN-AFYP-NR   SVLRNNI   EIIKTAVELWKNH--IE    IPYRIIYKNFS---NPDKDKK  2258
Query_OSTRE  2122  ---------IEIVL--K-CGHSkYV   GVMKANL   ALLTRMFARLRDA-vLS[14]-RANEILELVPks--KSQEEK  2189
Query_NITEN  2313  ---SKATDLIGHLM-RV-SAHP-TR   AVLSANI   DVIRLFLAEWKSD--V-[ 1]VDKRVILRYLAad-GSDA---  2370
Query_CLUBM  2077  ---HTASRFIDHLM-AV-AAHE-SS   QVLRANV   EIIKAFIQLWKDR--V-[ 1]IDGRLILKYLQag-GKSTNTS  2137
Query_MOSS_  2198  ---RLATNFVHHLI-SV-ADHG-SN   QVLRANI   DIIRLFLENWKGS--I-[ 1]LERGQLLQYLIsg-GRNPDMN  2258
Query_LIVER  2225  ---GMASEFITHLM-RV-AAHP-SS   PVLRANV   DIIRMFIDTWKKV--L-[ 1]FDKNIILEKLAfg-GKGASGT  2285
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2264  RNVV   GLQIVGVFLANGLSPWQPDAE   VS--KD---RFY    QYLAMNM---N-QPDKQVYSASAEIVGM--ALS  2322
Query_LINGU  2266  TNTA   GVQMVGVFLANKLPAYSYTSE   VD--RE---KFY    KALASNM---G-SYWKTSSAAAAEVVGM--CLQ  2324
Query_LIMUL  2256  ENLV   GIQLSGIVLANGLSPYSDTGG[1]LE--KA---RFL    NLLASNL---T-FKYKEVYASASEVIGM--VLK  2315
Query_SPIDE  2250  QNHF   GIQLLGVVLSCDFPPYDLQD-[1]AE--GY---SFL    TTLLKNL---S-MKYKEVYGAAAEVIAL--TLQ  2308
Query_SCORP  2239  DNIV   GIQLFGIVVSNGLDPYQNTS-[1]ID--KI---SYL    KTLTVNL---K-NKYKEIYESASEVCGM--SMK  2297
Query_TICK_  2177  DHHA   -VLLFGIFVSNNLVEFGQSHA   -S--TA---RIL    EVLPRLL---A-SRYRGAYTSVADVVGL--TLR  2233
Query_MONOS  2209  RLRS   SLHLFSVALANGIYPTAQNLP[4]LD----------    -GLCNQLagyLTHKSLTMCRTAASVIGM--YLQ  2269
Query_ROSET  2179  DNLS   GLHLLSLILANDMMPSESDVE[3]LS-mRS----FF    QILVANLmf----KSRSVYRTAAFVTGM--LLS  2240
Query_BELCH  2303  ANSV   GIQLLGVVLANQLPPFNPSTS[2]ID--RE---QFF    SAVAKNL---QLTRYKDVYAAAAEVIGM--IMK  2364
Query_CIONA  2268  ENFA   GLQLLGAFLVNKCSPFAGEYT[2]VD--RE---QYY    QRIFRNM---RF-KRKEIYDAASEVAGL--TLN  2328
Query_CAECI  2282  DNSV   GIQLLGLIILANNLPPYDPKCG   ID--VE---RYF    QALANNI---ALTKYKDIYLAAAEVLGL--ALH  2341
Query_HUMAN  2269  DNSV   GIQLLGIVMANDLPPYDPQCG   IQ--SS---EYF    QALVNNM---SFVRYKEVYAAAAEVLGL--ILR  2328
Query_MOUSE  2265  DNSV   GIQLLGIVIANNLPPYDPNCD   IT--SA---MYF    EALVNNM---SFVKYKEVYAAAAEVLGL--ILQ  2324
Query_XENOP  2289  DNSV   GIQLLGLALANNFSPLDPKCG   ID--PE---RYF    QSLANNL---GLTRFKEVYIAAAEVIGL--VLR  2348
Query_PLATY  2273  DNSV   GIQLLGIVLANNLPPYDPKCG   ID--SA---RYF    QALANNL---SLVKYKEVYAAAAEVLGL--TLQ  2332
Query_TASDE  2275  DNSV   GIQLLGIIMANNLPPYDSKSG   IE--ST---RYF    QALVNNM---SFIKYKEVYAAAAEVLGL--ILQ  2334
Query_CANAR  2271  DNSV   GIQLLGIVLANNLPPFDPKCE   ID--RV---RYF    QALTSNM---GLLKYREVYAAAAEVLGL--ALQ  2330
Query_OSTRI  2241  DNSV   GIQLLGIVLANDLPPFDPKCE   ID--RV---RYF    QALISNM---GLLKYKEVYAAAAEVLGL--ALR  2300
Query_TURTL  2280  DNSV   GIQLLGIVLANNLSPYDPKCE   ID--SV---RYF    QDLANNM---SLIRYKEVYAAAAEVLGL--ILQ  2339
Query_GECKO  2280  DNSV   GIQLLGIVLANNLPPYDPKCE   ID--NV---RYF    QALANNM---ALVRYKEVYAAAAEVLGL--VLQ  2339
Query_DANIO  2277  DNSV   GIQLLGIVMANNLPPYDAACG   IE--HD---RYF    QSLANNL---SFIRYKEVYSAAAEVIGL--ILN  2336
Query_SNAKE  2278  DNSV   GIQLLGIVLANNLPPYDPKCE   ID--SI---RYF    QALSNNI---ALIRYKEVYVAAAEVLGL--ILQ  2337
Query_STERL  2292  NNSV   GIQLLGIVLANNLPPYDPKCG   ID--SE---SYF    QALATNI---SFGRFKDIYAAAAEVIGL--IMQ  2351
Query_MILII  2278  DNSV   GIQLLGIVLANNLSPYVPKCE   IG--YE---RYF    QRLVYNI---SLTRYKEVYAAAAEVLGL--VLH  2337
Query_BAMBO  2276  DNSV   GIQLLGIVLANNLPPYDCKCE   ID--SE---RFF    QRLVHNI---SLTRYKEVYAAAAEVLGL--VLQ  2335
Query_ALLIG  2238  DNSV   GIQLLGIVLANNLPPYDPKCE   ID--ST---RYF    QALVNNM---ALLKYKEVYAAAAEVLGL--TLR  2297
Query_GAR__  2281  DNSV   GIQLLGIVLANNLPPYDPSCG   IE--HE---RYF    KSLANNI---SFVRYKDIHVAAAEVIGL--VMQ  2340
Query_TETRA  2260  EKSQ[5]SILILDACIQLRIPILSNYEA[4]-QtvEE-kfAFF[ 6]DRLKRLLlf-QQHSKRQYCKAASAVCGG--LLS  2339
Query_PARAM  2143  DTSV[5]GLAILHIALLNNVVILQDYPE[5]NSslYEkyrFIL[10]TEQINRLac-TYTTKKHVLKGIANVYGS--VLQ  2229
Query_SLEMN  2244  PESN[5]GIQILALAVSFDILVIEKPEN[5]----VN-qrEFR[10]DKLMNSLlkiLESKKKQIIQAASETLGK--ILN  2326
Query_STENT  1971  KISY[5]GIIIYGSAIDNKIKILSD---[4]----Y-------    ESLDDSLlksLEVPRRQIILSASEILGK--RLE  2033
Query_NEMAT  2266  DNAT   GIQLLGVVLANKIHPFSSDSS   VD--EN---TFY    TALSDNL---TF-KYKDVHAPAAEVSGM--LMK  2324
Query_PISTI  2304  DNAV   GIQLLGVVVANGLFPVTSDST   ID--ED---RFY    GTLVNNL---TF-KYKDVHAAAAEVIGL--LMM  2362
Query_CORAL  2278  ENSV   GIQLMGVIVANKITPFKPGCG   VD--EK---RFY    QDLVKNL---SF-KYKDVHGAAAEVIGM--ILK  2336
Query_HYDRA  2177  KNLR   GLQILQFYLANGFAPFATQIS[1]LT--EY---DYF    LALSNNL---LF-TYKEVYATCAEVCGI--SLK  2236
Query_LATIM  2280  DNTV   GIQLLGIILANNLPPHDPTCS   ID--ME---SYV    KALANNL---SFTRYKEVYAAAAEVLGL--VLQ  2339
Query_SHRIM  2241  EVEA   GLQLLGIILANGLPPYTVDSE   TE--KK---RCE    SLVIRLL---G-AHYASVYGAAAEVMGL--ILK  2299
Query_MCRAB  2236  EQEA   GLNLLGIMLANGLAPYTEDSV   AE--KT---KCE    KILLKLL---E-AHYASIYGAAAEVVGL--VLK  2294
Query_AMPHI  2249  EVDA   GLHAAACLLVAGVSPYNEQDR   MQ--ER---KMH    GVLLKCL---A-NPRAAVYNAASEVVGL--MLA  2307
Query_BARNA  2129  QLHA[3]GVCALRAALLADRPLCLDREE[4]LD----------    -RLLTLMvd--RSSTAELYRPASTVVAL-----  2187
Query_COPEP  2156  DLSH[2]ALKVLAVFLGQQM--------   FD--RE---RLE[ 3]--LCQKIcqtFTSKRKEVYQAAGVATGL--YLT  2213
Query_DAPHN  2063  KRKL   GLLILDAIVVHDFLPKNHEFN   -----S---RFK    EQLVKWL---TLDEKRDLTRSCAALCGS-iLIR  2120
Query_STARF  2316  DSST   GLQLLGIVLANGLHPIDKAS-[1]ID--EE---RFY    AALSYNL---K-NRYKMVHAAAAEVIGM--AMK  2374
Query_STRON  2300  DSAV   GIHLLGVILANNLHPLSPTC-[1]VD--AD---RFY    SALASNI---D-NRYKAVHGAAAEVIGM--AMK  2358
Query_SEACU  2277  TNTV   GVQLLGVVLANNIHPFSSTS-[1]VD--SE---RFY    AALATNL---K-SRFKQVHAPAAEVVGM--TFR  2335
Query_FUNGI  2219  NTHT[3]ALQLLAVFVLNGFPIYHDEQD[2]IS--KV---KYM    DLIMSNLfseEGRGAKELYTTTAEVVAMqlALQ  2288
Query_DPURP  2350  SRQM[3]GLVLLSILLSNGFPAYD-KDY[4]IS--EF---KFY    QILLDHL-----QDFKELYDAASEVCGM--ILL  2413
Query_HALBU  2348  ARQV[3]ALIQLSILLSNGYAAYDKEND[2]IS--EF---KFY    QILIDCL-----TEYKELYDAASEICGM--ILQ  2410
Query_GLOBO  2123  KYLN[2]GVYVLQALVANKVSAY-DARG[4]IF-sES---DLL    DSLSSLL---TLSKLKEIYSPLAECIGT--VLL  2188
Query_PALUS  2182  NLL-[1]GIYLILVYVTHGVPPF-DMRG[4]SL-sED---QFC    DAFMNLL---VNGR-KEIYSFSAEVLGR--WI-  2243
Query_CALIF  2258  VSVI[2]AIQILASFVSNDIPEYNELNM[4]SYpsET---DFY    SVLANYI---QYNHKN-VYEPASELCGL--ILK  2324
Query_GLOMU  2316  ---V[2]GLQLLGVVLAHDKPLFQDEVG[4]GL-tSD---KFY    MDLTNNL---LNHVNSNVRALAAEVCGL--AVK  2379
Query_LOBOS  2244  ---N[2]GIQLLGIVLTHDNPPFYRGPE[4]TL-rES---DYY    EALVRNM---GSSQKA-VYSSAAEVCGL--VLA  2306
Query_RACEM  2229  ---I[2]GLQLVGAILVHGINPFYNGEG[4]GL-tEH---QFY    AAICENM---DNSSRE-VAATAAELATL--AFK  2291
Query_PVIOL  2342  SRQI[3]GLVLLSIMLCNGIPAFDASNH[4]QS--EF---KFY    QILLDNL-----QDFKELYDATSEVCGM--ILL  2406
Query_BOMBU  2113  ---A   TVYLILICFINGMG--LDIIQ   ----RN---DIL    EFLERSLkkwQ--DDEETILQCCECYGL--ILK  2166
Query_AEDES  2254  NIG-   GLQLNAIVMANGLVPWTETS-   ----KF---DFI    RAIFRCLd----SEKAAVYRPASELLGM--CLN  2308
Query_ANT__  2141  ---T   VMYLILVLLDNDMA--EEIIV   ----RN---DIV    DFLLKPLsvwDTSNKDELPLQCCECLGL--YLR  2196
Query_WHITE  2210  KAQT   GIMLAGILVVNNIAPYASAS-   ----KM---KYL    SALVRKMh----SQKTTIYEPTAEVIGR--ILL  2265
Query_BEETL  2110  AVEI   GIHVSSIFLANKLLPWKGDS-   ----KD---FCK    SIFNVILk----AKYKTIYRPCAETLGL--ILR  2165
Query_THRIP  2143  V---   GINLTAIMLFHKLTPWTTKS-   ----KQ---SFI    SFLLKNLn----NNTKEIYKASAETLGM--CLN  2195
Query_TERMI  2225  VSEP   GIQVAAIILANRLEPWSSTG-   ----KL---QFL    RALLRNId----TEKRSIFQSSAEVAGM--ALE  2280
Query_FLEA_  2153  NMG-   -IELAGILLANKIIPWPDSK-   ----KD---DYY    KALCMQFn----SMESITYKPCSEVIGM--VLN  2206
Query_MIDGE  2203  NLRC   GIQLNAVLLINDIVPWKDSEQ   ----RK---QFI    AAIVNCFr----HSNSQIYRAAAQLLGM--CLH  2259
Query_LAODE  2194  DLSK[3]SIEIICCLLNNNIKPYAEESL   -D-------RFY    RKLEQKVlf----KDRSIRQKSAEILGM--LYR  2252
Query_LOTTI  2242  NKSV   GIQLLGVVITCKFPPYGPDCP   VD--RH---KFF    TTVAFLM---C-DKSKPVYTATGEVVGM--ILK  2300
Query_APLYS  2301  RSST   GVQVLAVVLSADFPPYGPSAP   VD--KD---KYL    SSLSNLM---N-HHLKIVYAATAEVIGL--VLL  2359
Query_SCALL  2300  RHST   GIQILGVVLACKFPPYGPVAP   VD--RD---KFF    STVSNYL---T-HHLRSVYAATAEVVGM--IFQ  2358
Query_CRASS  2264  ESCV   GIQLMGVVLSAKFPPYAPTAP   VD--QE---KFF    TTVAQYM---R-NRYKAVYAATAEVVGM--IFK  2322
Query_OCTOP  2213  GLMV   SVQLLGVLLKCNFSPYVNMT-[1]VP--RE---TYF    SHLSQGL---H-HSHKIVYASTAEVLGL--SLK  2271
Query_POMAC  2267  SSLI   GIQVLGVIMTCNYAPYGPSAS   VD--RN---RYF    TTLASSM---T-NSSKMIFAAAAEVVGM--VFA  2325
Query_WHIPW  2177  ENAF   GVQLLQSVLKNMKKPCELCGD[3]----DQ---RIF    EYLINRL---S-SKHAKVYKPAAYVCGL--TLS  2236
Query_TRICH  2193  ENVC   GISILCLCLNTVPETFSFLTD[3]----ES---GIF    SHLVRLL---S-LKNVKLFKLAAFVCGL--SLK  2252
Query_AASTA  2058  ENAS[4]GMQLAAVLINLAFDASSPVLW[4]LRacSS---RLE    TALLARLas-TTHKTTPQ--LAADVSGL--VLQ  2127
Query_AEUTE  2003  LDTA[4]GMQIVAMLLSLAFDATSPATS[4]LRnaRA---KLE    NNLLTRMvs--MHKTTPI--LAADIAGL--VLK  2071
Query_PLASM  2178  IKYA[4]ALQIVSAMLVAGLKS------[3]AEva---nqTIE    DGILLCMtt----KVTSLYTLAAEVGGL--YLL  2239
Query_BREMI  2164  LNNA[4]ALQIVSAMLTAGHIS------[3]VQvs---dhTIE    DGILLVMts----KTTPLYTLAAEVGGL--YLQ  2225
Query_OLIGA  2193  REAA[4]ALQIVSTMLTAGLVE------[3]PDfdMSdgkTIE    DGVLESLyh----KRASIYTVAAEISGL--LLK  2257
Query_LIGNO  2374  AAAA   GLQLYNAYLSNALPYFNEQQG[5]VD--SP---QIL    AALVRLL---N-HSRKEVYSPCAGVTGL--CLR  2437
Query_CESTO  2160  KRAI[2]-LHLFKTILIS-CQRFAPESE[4]-----S---QFV[ 1]-ALLPYIrq----TPKTLLTAAWECSAL--LAS  2219
Query_OVI2_  2538  QLLV[1]-LDLTTSMLRSKMSFDQPQMN[2]LD--KMag---L    --LIPLLd----HENHSVRNSALLAAAL---LL  2595
Query_PRIAP  2291  DNSV   GVQLLGVVVANNLPPYSHTTA   VD--SE---KFF    QAIVRCL---A-NKYKEVYAAAAEVLGM--ILH  2349
Query_ROTIF  2246  KLST   GIQLFGVVLSNGIENYEYPMD   IS--RL---DFF    KSLINCI---K-EPSKLIHAPAAEVVGM--LLK  2304
Query_SPONG  2292  SNLS   GIQLVGIAVANELNPYDTNTA[2]LD--QG---LYY    KIFVNNL---TF-KYKLIYAAAAEVLGM--IMK  2352
Query_TARDI  2314  LNIP   VINIIAIILNAGLPLTAGSFE   ---------EFL    KILMLHL---Q-ESSKPLYSVSGLTLGL--LFS  2368
Query_TRICH  2259  DNST   GIQCLGLVVANGLSPYRYDVE   VD--E------M    LALAENL---QF-KYKEVYEAVAEVLGM--IMK  2314
Query_OSTRE  2190  RFVT[5]VMQLIATLIHEETKVIQVNDE[5]NSrgEQ-----I[ 9]QAVMSCLmssTESTGGLLNSVASALLGK--IME  2271
Query_NITEN  2371  ----[6]GLQLFGVLIAVGMKIYDSIED[2]IP--EK---EVY    GRLLECL----GFKAKEVYEAAAEVIGA--ALF  2433
Query_CLUBM  2138  DRQL[6]GLQLFATVVSSSGSVNQGLER[2]VG--RE---DVY    ETLLSNL----KLKAKEVFAASADVLGM--VLR  2204
Query_MOSS_  2259  DRNI[6]GLQLLGAIVGSGIPIYDPSCD[2]TS--EE---HIC    EALLNNL----NVKTKEVYEAAAELVGI--TMI  2325
Query_LIVER  2286  DRSV[6]GLQLFGAIAASGLPLYDPACD[2]I---EK---DLY    QVLLQNL----NFKGKEIYEASGELIGI--VLK  2351
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2323  SL--Q    NteSDKEWCRA-YGDNIQKMLLGIH-    GNKP-DQFLTCVHCI---QKHFSAFVE-R    FIHKVLFM  2381
Query_LINGU  2325  DL--D    --qHDKDFKEE-YVNHLHSVLSSLQ-    QSNP-DTYMLCVYKM---HLHHKPIAE-R    FMNKLLFM  2381
Query_LIMUL  2316  YM--E    E-kEGSP-DEA-FQNSIEQRVTELSN    -TKQ-DQFITCVNKV---HLHFPPIAD-R    FLQKLLFL  2372
Query_SPIDE  2309  YI--S    K-sETDILKNK-FFLKVSQVLDTMKH    -KEE-SKFLYCLNKI---HSSCPEIAN-K    FIPHLLFL  2366
Query_SCORP  2298  YY--F    D-kDDDV-YSE-FCKIVKTEISSLN-    --AK-DQYLTCLYKI---NTHFPIFID-S[ 1]FLNKILYM  2353
Query_TICK_  2234  KM--S    E--AGIATEDSaFKGSVASRLDGLLA    --AHRDQFVLCVRNV---CRFYQPLAL-R    FLDKLMLL  2291
Query_MONOS  2270  CV---[ 5]DgaSDTAGVGQmLTQATQRCLMDIKR[ 1]SAEGRERFVTALHAIcnaTLGYPAFVD--[ 1]LVQTLLYV  2339
Query_ROSET  2241  HF-kS[ 7]----------DeLMNEVVARLDQLKK[ 2]DATSKERLITCLHELcnvNTGHPPVVD--[ 1]FVQRLLFL  2305
Query_BELCH  2365  YM--A    D--VEKVTEGS-FHTQVERYLLPLQ-    VNQQ-DVFLTCVHKI---HLHFPPFAD-R    FVSKLLFM  2421
Query_CIONA  2329  YI--T    N--VEKEENGN-VHEMVRNNLSALFK    NPKTQDIFITCLHKV---SMHFPPMVD-N    FVSAVTGL  2387
Query_CAECI  2342  YV--A    E--KEKTLEGP-LFDCVIKELKQH-Q    NAKE-DKFIMCLNKV---VKNFSPLAD-R    FMTVVLFL  2398
Query_HUMAN  2329  YV--M    E--RKNILEES-LCELVAKQLKQH-Q    NTME-DKFIVCLNKV---TKSFPPLAD-R    FMNAVFFL  2385
Query_MOUSE  2325  YI--T    E--RKHVIAEL-VCELVIKQLKQH-Q    NTME-DKFIVCLNKI---AKGFPPLAD-R    FLNALFFL  2381
Query_XENOP  2349  YI--V    Q--NEKRTEAP-VFDYVVKELKRH-Q[ 1]NNKE-DKFIMCLNKV---VKNFPPFAD-R    FMTIVLFL  2406
Query_PLATY  2333  YV--S    E--KETMLEDS-ICELIVKQLKQH-Q    NSRE-DKFIVCLNKA---VKNFPPLAD-R    FLNAVFFL  2389
Query_TASDE  2335  YI--A    K--KENVLEDS-VFELVVKQLKQH-Q    NTKE-DKFIVCLNKA---VKTFPPLAD-R    FMNAVFFL  2391
Query_CANAR  2331  YI--A    E--RENILEDP-VYDCVIKQLKHH-Q    NTQH-DKFIQCLNKV---VKNFPPLAD-R    FMNAVFFL  2387
Query_OSTRI  2301  YV--A    E--RENILEDP-AYDCVIKQLKHH-Q    NTQQ-DKFIQCLSKV---VKSFPPLAD-R    FMSAVFFL  2357
Query_TURTL  2340  YM--A    E--KENIFEGP-VYDCVIKRIEQH-Q    STKE-DKFIVCLNKV---AKNFPPLAD-R    FMNAVFFL  2396
Query_GECKO  2340  YI--A    E--KENIYESP-VFDIIIKQLKQH-Q    NTRE-DKFIICLNKL---AKNFPPLAD-R    FMNTVFFL  2396
Query_DANIO  2337  YM--T    E--RENQIEGT-LFNITVTKLMDLRK    KEVD-DKFIVCLSKV---SKHFPPLVD-R    FINPVFYL  2394
Query_SNAKE  2338  YI--A    E--KENTFEGP-IYDMVIKQLKLH-Q    NTRD-DKFIVCLNQI---TKNFPPLAD-R    FMSFVFFL  2394
Query_STERL  2352  YL--A    K--NEKQLEGP-IFDFTVKELKHL-Q    NTKE-DKFVECLSKV---VKNFPPLVD-R    FLNAVLFL  2408
Query_MILII  2338  YF--A    E--NEKQVDGL-VHDLIAKELKQL-Q    QSKE-DKFIVCLSKI---VKNYPPFADrR    FLNQVFYL  2395
Query_BAMBO  2336  YL--A    ----GKPMDCA-VHDLVVKELQQL-Q    NTKE-DKFIVCLSKI---VKHFPPFAD-R    FMNAVFFL  2390
Query_ALLIG  2298  YI--A    E--KENITEGP-IYDCVVKQLKQH-Q    NTKE-DKFILCLNKV---VKTFPPVAD-R    FINAVFFL  2354
Query_GAR__  2341  YM--A    E--KEKQSEGP-IFDFTAHHLKKL-Q    NEKE-DKFVVCLSKV---VKHFPPFMD-R    FVNAVFFL  2397
Query_TETRA  2340  MM-hE[ 5]-egKRKIYEQE----IVNNSVNIINK[ 4]SKERFPIIVACITQY------YPDFLN-N[ 5]YCQDYIHQ  2408
Query_PARAM  2230  CC---[ 3]------KVSGNdYEQEIINIVTGIKA[ 6]GKEKFTTFMKILTFV------YPKVLSnN[ 1]LFQETANL  2292
Query_SLEMN  2327  KQ---[ 1]--------------SLVAQSIHPLMN[ 2]ITERHDVFVHSIEKI---TREYPELLIdR[ 1]IFMKLLSF  2378
Query_STENT  2034  YS---[ 1]--------------NLQT-SLFALLS[ 2]ESKEIGVYVNILEKV---SICYPQILNnK[ 1]ICFKLSGF  2084
Query_NEMAT  2325  YL--I    E--ERKVTECS-LIALVEAKLSPMAT[ 1]GGKDIGKFITCVHRL---QLNYPRIVD-R    FMKHIMFH  2384
Query_PISTI  2363  QL--A    K--VKKIFDGP-LHDMVGQKLLSLVN[ 1]AKPELDKFITCVHKL---KLNYPPIVD-R    FVNQILFT  2422
Query_CORAL  2337  YL--S    E--ENDKESCDgLSGLLNEKLSSMT-    --KDLDKFILCVHQL---QIHFPPIVD-G    FINQILFV  2393
Query_HYDRA  2237  YY--F    S---KCESPKV-LLEMVQGSLCELSK    NNKEK--FIVCLHSI---QINYPMVVD-R    FISQLLYC  2292
Query_LATIM  2340  YR--A    E--KEKELGGP-IIDSVVRELKQL-Q    SIKE-DKFIVCLSKI---VKNFALLAD-R    FMSVVFFL  2396
Query_SHRIM  2300  YT---[ 3]DgsVEA------LEDAVVKRVTELTV    --KMQGQGLTCIYNI---HKHYPSFVK-R    LINKLLNL  2355
Query_MCRAB  2295  FQ---[ 5]DcnETS------FEENVARKVMEMMH    --SKQGQGITVIYKI---HLHFPAVLK-R    VMNRLLSV  2352
Query_AMPHI  2308  RS--M[16]SltSEATVGCDeFFGDVVACVMDLIS    --KHQAQGLSVLYHI---HKHFNTIDN-R    FTSKLLVL  2383
Query_BARNA  2188  -L---[ 3]------RTDIEaVTPVVAAKLAAVEK    --DDRVRWVTCLARL---SAHCRPLAAsQ[ 2]---RLLFQ  2242
Query_COPEP  2214  TL---    -rsNQNVFDKV--LNIFLKALSDMSL[ 2]SADTRSRFIIILHGI---QQSYPQILI--[ 1]LASKFQYS  2273
Query_DAPHN  2121  ET---    ---------DHeFEKLVCDALQMWYN    -RKKVDVFVDLLYEV---AIRHPAVLS--[ 1]FASSNLSL  2171
Query_STARF  2375  QK--A    D--VDKVTDGF-LHSIVTSELTKI--[ 1]EIKP-DVFVTCIHKL---HFHYPEFAD-H    FTNKLLFM  2431
Query_STRON  2359  QM--A    E--NDKITDGR-LHEVTYKELSQLHL[ 1]SGKE-VIFITCLHKI---HLHYPVFTD-R    FMNKLLFM  2417
Query_SEACU  2336  QM--A    D--NDQITDGY-LHQITTKELTAL--[ 1]NMSP-EIFINCINKI---TLHYPEFAN-R    FLNKLLFV  2392
Query_FUNGI  2289  QN-dK[16]AsqRSQETED-vCEQMFRDKLNTIFS[ 1]GDYDRY--INCMERI---DLQAPFILI-H    HYPRIFQI  2365
Query_DPURP  2414  YN-sK    -----NNIQS-vFSQLLKDKINSILG[ 1]NETQRA--FSCLYFI---GLHYPLFLQ-N    FYGKIFNL  2469
Query_HALBU  2411  FN-hK    -----HNIQS-vFEKLLKDKISSFLS[ 1]SDYQRA--LNCLYFI---GLHHPSFIL-Q    FQLKLLNI  2466
Query_GLOBO  2189  FMetT[ 5]--------AS--FLSKIEDVCKKLQA[ 1]---EPKSFIVIINRI---SVHYPRILQ-N    QYKVLFNL  2245
Query_PALUS  2244  ---sS[ 5]--------AG--IAAQLEEKLRNLLR[ 7]SGAEQDRFVLILNRL---SVGYPDAVK-D    YAKDLLFI  2306
Query_CALIF  2325  RL--K[ 5]--------EA--ILNIITNKLELLNP[ 3]KNRTVDKFIICLHQI---GINYEEILD-K    YLVLILFL  2384
Query_GLOMU  2380  HL-rK[ 5]--------A---LSSLLNPMLQKI--[ 9]NESDMKNFLTVIHHI---SMHDVDVAV-R    FLKYVFGI  2443
Query_LOBOS  2307  YM-kR[ 5]--------ND--FQDLVTKKLSELLV[10]GKS--VSFINCLHKV---QLHFPEITD-A    FGPRLLFL  2372
Query_RACEM  2292  LM-iE[ 5]--------ID--MESMIVEKLKRM--[ 8]NTG---KFLSCIYRV---QTHQPEIAN-K    FANKVIYL  2352
Query_PVIOL  2407  YN-sK    -----HQIQSNtFQQMLKDKINTLLA[ 1]NDYQRV--FSCLYFI---ALHYPGFVA-S    FYSKIFNL  2463
Query_BOMBU  2167  SL--N[ 1]DntLANKRCAI--IDTIRSILRQMQT[ 1]CENRQ---MKCIRAL---CKNYPIAAM-D    YFDFVKRS  2225
Query_AEDES  2309  NL--Y[ 5]Qe-DERYQ------NEFVAKLVTMRK[ 1]QEKKF---MDIIYGI---HKAFPIIVD-S    FLAVITNA  2366
Query_ANT__  2197  SL--D[ 5]EteRENKKYEV--KQKIFTILGFVKP[ 1]HTIKQ---IKRIAVL---CRTYPEVAL-D    YINVT---  2256
Query_WHITE  2266  FL--Q[ 6]QltNDGSTL----LNSVKTLMQRIYK[ 1]TEPKF---YRCLLKI---SLSFPPVVK-E    FLRQMIVN  2327
Query_BEETL  2166  ET--S[ 1]---DLPKLCGN--VHILLKKIN----    EPDKY---AHILEGL---VLHYPPIVN-E[ 1]YLNSLLSR  2217
Query_THRIP  2196  QM--C[ 6]LsiEDQKFVDE--LENILKNVNLAK-[ 1]DCEKF---LVLLDNV---SENYIQIID-N    FVSNILVY  2258
Query_TERMI  2281  LL--A[ 6]LegDDAEF-----LSGIVDKLNLTKD[ 1]SKDKF---LTCLFAV---QEYYPRIVD-A    FMNHVLFH  2341
Query_FLEA_  2207  LS--S[ 5]Ni------L----LSSVNRRLDYLRR[ 1]NTDIF---IICLYGI---AKHYPDIAY-P    YITYVIFN  2261
Query_MIDGE  2260  QI--V[ 5]VegDENFQI----VSDITERLEKIRR[ 1]AGNDFNAFLQLLYGI---QKSFPRILD-P    FMTLIKFS  2323
Query_LAODE  2253  FL---[ 3]DgnSPSLIEER-LENLTKQILRHALD[ 3]NPDGLDTMILFTSLMkn---HYPSIVT--[ 1]FIKDIVPL  2318
Query_LOTTI  2301  YL--E    --vKEQETEGE-FHTYIINNISSLQ-    LSKP-DNFIYCIHKI---HQQYPPTAD-R    LLNKLLFM  2357
Query_APLYS  2360  KM--E    --rSDGEVEGS-FHDHLVSQLSNLS-    HAKT-DNFIVCLHRI---HKTFPKFTV-R    FLSKLLFL  2416
Query_SCALL  2359  YL--A    --eKEKETDGA-FYNHVTGLMNSIH-    LSKP-DNFVIGVHRL---QRHYPPVAA-K    FINKVLFM  2415
Query_CRASS  2323  HL--A    --eKDRQTEGT-FHDYVVNLMNSLQ-    QSNP-GNFLLCIHRM---HLHYPPIAD-R    FMNRLLFV  2379
Query_OCTOP  2272  YF--D    Q-hKDSE-FEK-FMKQIERVMKTLE-    -PKL-GNFIVCLHKM---QLHYQPVAE-S    YLPKLLYH  2327
Query_POMAC  2326  YL--A    --dKENETDGQ-FHSHVESVLGKVA-    ---P-HNFLVCVHCI---HRHYPPMAD-R    FLNKLLFM  2379
Query_WHIPW  2237  AL--S    -------VNGGaFKEALRSAMVKMTG    TTLDRSKFLLCLSEI---HKSETSFAS-S    FLAELLSW  2291
Query_TRICH  2253  VL--N    ----ELQTDCEwLLKNMKAVMSRMTG    TMLGRSKCLFCLHEI---LPNYPEFAL-Q    FCNDLLDW  2310
Query_AASTA  2128  HM---    -----PASTA--FESGIDNCLQSFFQ    -LDKPDRFLTCLKQL---ATHFPAIVS-G[ 1]VLNRLASI  2181
Query_AEUTE  2072  HM---    ------HSTT--FQTSVETQLALFFQ    -EEKPDRFLTCVKQV---SFHYPPIVD-E[ 1]TLNRLMSI  2124
Query_PLASM  2240  TM---    -----DQATSKcFLCNLKKLIADSYT    -SEEIGYFLAILRNV---SIHQATIVD-S[ 1]MLQRLSSV  2295
Query_BREMI  2226  TT---    -----RQIDTDaFMCKLKSLIIKAYD    -EEDFGRFLAILRYA---SLDQANIID-S[ 1]MLQRVSFV  2281
Query_OLIGA  2258  AK---    -----TGSKAQaFRDKLAEHITKYYS    -EEEFGRFLALLRSV---SLHEPNVLD-S[ 1]MLHRLSFV  2313
Query_LIGNO  2438  QM--V    --tAQSVVPAF--YKQVAEQLGKMLT    -ARP-DAGLHALAQV---SRCCPRLACiE    FRQRLLFQ  2494
Query_CESTO  2220  KC-fT[11]RsrGSDILPLSaFPKDLHPLINELWS[10]CKRPSDTPIMSVIEAacaSASHWSVLAlH[ 3]--SLLGTG  2308
Query_OVI2_  2596  HKgkT[13]---SQSLPTLDsMERNIHDRITCLGQ[ 8]AHQSLSLELACRAAL-----SYIALQT-H[ 3]-VASVAFQ  2677
Query_PRIAP  2350  RM--V    E-rEDEEAEGA-LRDAVVTAMSTLH-    ADAK-QNFIVCLQKM---TLHYPAIAD-R[18]LMSKLLFM  2425
Query_ROTIF  2305  RF--D    S-hDKEYFENT------VSHLFEILR    -EADTSLFITSIHKI---QLNYAAISE-R    FMTKLVFF  2358
Query_SPONG  2353  RM--A    E-sYDPFLDRL--EEMVGEQLFKLSQ[ 1]LEKSEDKFITCLHKI---TLCYNKIVP-R    FVEKLLFL  2412
Query_TARDI  2369  RA--E    GagIAGETIEA-GVEALRQFLSGIVA    --TDRDKFVAVLHQA---SKSYPAIVA--[ 1]FQARILFL  2427
Query_TRICH  2315  HL--A    S--VKKETEGP-IFDYVTQHLSSMVS[ 1]GRMDVARFILCLYRV---QLHYPIIAD-R    FVNKVLFL  2374
Query_OSTRE  2272  ID-gN[ 5]-------------HDYLRKRCQSLYD    -AQAWDPFLDTLDRV---TKRCPNFPHdT[ 1]LAPLVHSV  2327
Query_NITEN  2434  ER-dK[ 5]--------QS-lLERPLHQRLLSLYR[ 1]NSATKATFYNIAAKI---CGRFPKFVE-K    WASKVAGD  2493
Query_CLUBM  2205  DK-sE[ 4]---------S-aVENKFKQFFVSTFK[ 1]RNYSK--CLVILDRL---TSKYPPYVG-A    YIHVLIDL  2260
Query_MOSS_  2326  YR-kE[ 5]--------EE-iLEAKLHKILKQLYR[ 1]AEYER--FLNILNKI---TIRHPGFLE-G    YAEMVVDQ  2383
Query_LIVER  2352  QR-hQ[10]--------EQ-mLEAPLKQCLLKFCR[ 1]SNLDY--FLNILYKI---TFRFPQYGE-A    FASLLVDL  2414
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2382  LPGLYGVF-RTYALEIIAAGIENI   EG   AFVEIKNKNIIGIL    RERDELSQLMSLK   IIRRLLEK--LEAS  2445
Query_LINGU  2382  LPRLHGEY-KRHCLEVISSRVQNM   EN   AFIELKNKGLLTML    THRDDSTQAVCLK   IVNELLRR--LQPD  2445
Query_LIMUL  2373  LPKVHGNF-RTLIMEVVASRVEYI   QG   GFTELKTKGLLMMI    THRDESVKKFSLQ   AIKGLLPQ--LTSD  2436
Query_SPIDE  2367  YSRVHGIY-KNYILEIFYSRVLEF   EN   SYQELKHIGVFENL    EKKNDIIQHLCLQ   TLEKLVPN--LSCD  2430
Query_SCORP  2354  LPNVHGIF-RKQILEIINSHVENI   EN   AFAELKNQGLLNML    EHRDESIQFSSLN   IIQKLLKE--LDID  2417
Query_TICK_  2292  LPRALADF-KVAVLEVLSLCCHQM   EQ   PVGRLQGLGFLELL    AHGDEKMQTVLLE   ICTKLVGA--ATEE  2355
Query_MONOS  2340  LPACAGNA-KTQALEVLMWRATAI   PD   LFRELLGKDLPGLL    QRHAPSTQCAVLT   ILKAILPS--LEPS  2403
Query_ROSET  2306  APSLPGQF-KTMALETLLWRAEDI   PD   LFRELQGKGLLSLI    KHHHTETQAVVLS   ILKVLLHH--LNAS  2369
Query_BELCH  2422  LPKLHGIF-RTHCLEVIRGRVNVI   DN   LFLELKTRNFLGML    THRDDGLQAIVLQ   IVYALLQK--LKVE  2485
Query_CIONA  2388  LPKMHGDL-RTFCYEIINQRASHI   ED   LYQQLDMKNLITNF    SVGGDPLQLASLR   ICSKLVMA--TIIN  2451
Query_CAECI  2399  LPKLHGVL-KTYCLEVVMCRAEEI   PD   LYIQLKSKDFNQIM    SHRDDDRQKVCLD   IIYKMLPK--LKPM  2462
Query_HUMAN  2386  LPKFHGVL-KTLCLEVVLCRVEGM   TE   LYFQLKSKDFVQVM    RHRDDERQKVCLD   IIYKMMPK--LKPV  2449
Query_MOUSE  2382  LPKFHGVM-KTLCLEVVLCRAEEI   TG   LYLQLKSKDFLQVM    RHRDDERQKVCLD   IVYKMVAK--LKPI  2445
Query_XENOP  2407  LPKLHGVL-KTQCLEIIMHRAEDI   PD   LFIELKNKDFCQIM    NNRDDERQRVCLD   IIYKILSK--LTPA  2470
Query_PLATY  2390  LPKLHGVM-KTFCLEVVLCRAEEI   TD   LYLQLKSKDFAQVM    RHRDDERQRVCLD   IIYKMLAK--LKPG  2453
Query_TASDE  2392  LPKFHGVM-KTLCLEVVLCRAEEI   HD   LYLQLKSKDFIQIM    RHRDDERQKVCLD   IIYTMLAK--LKPV  2455
Query_CANAR  2388  IPKLHGVM-KTYCLEVIMCRAEEV   PD   LHLQLKSKDFIQIM    NHRDDERQRVCLD   IVHKMLSK--LKPL  2451
Query_OSTRI  2358  IPKLHGVM-KTHCLEVIMCRAEEI   TD   LHLQLKSKDFLQIM    NHRDDERQRVCLD   IVYKMLPK--LKPL  2421
Query_TURTL  2397  IPKLHGVL-KTYCLEVVMYRAEEI   TD   LYLQLKSKDFIQVM    SHRDDERQRVCLD   IVYKMLPK--LKPL  2460
Query_GECKO  2397  IPKLHGVM-KTYCLEVVMGRAHGI   PS   LYLELKSKDFTQIM    SHRDDERQTVCLD   IVYKMLAK--LTPA  2460
Query_DANIO  2395  LPKLHGML-KTHCLECVLSRADVI   PE   IFLHLKTKGLSQIM    SHKDEGRQRVCLD   IIHKILAC--LKPE  2458
Query_SNAKE  2395  LPKLHGVM-KTHCLEVVMCRAQEI   PD   MYLELKSKEFIQVM    CHRDDERQRICLN   IVYKMLPK--LKPA  2458
Query_STERL  2409  LPKLHGVL-KTHCMEILMSRAEEV   TD   IYLQLKSKDFAQIM    SHRDEMRQRVCLD   IVHKILAK--LKPV  2472
Query_MILII  2396  LPKLHGVL-KTHCLEVVMHRAETI   ID   IFQELKNKNFVQIM    SYSDEERQRICLD   IVYKIIAK--LKAE  2459
Query_BAMBO  2391  MPKLYGIL-KTHCLEVVMYRAEEI   PD   IHIQLKSKDFTQIM    SHRDEDRQRVCLD   IIYRIKAK--LKSA  2454
Query_ALLIG  2355  LPKLHGVL-KTSCLEVVMCRAEEI   TD   LYLQLKSKDFMQVM    NRRDDERQRVCLD   IIYKMLSK--LKPL  2418
Query_GAR__  2398  LPKMHGVL-KTHCLECILSRAEEI   PD   VYLQLKSKDFSQVM    SHRDEARQRVCLD   IIHKIMAK--LKPV  2461
Query_TETRA  2409  KSKMDRAY----ILQSILIFLKKK[1]-D[2]---ENMKEDIHKIC[11]EDENVENQEKFLS   LFVFLSSPqiLNSF  2481
Query_PARAM  2293  IVQASNKE-RAEILRIFSEFLEQI[4]QD[4]---------INEIV[11]EDQEESVHYALLN   LLKQLVQLkcLNAC  2368
Query_SLEMN  2379  VKTLTGSM-RASILKSFERFINVC[2]RK   -----KIEDINEIA[11]ADISDENQQSFLI   LITQIAKLnlDTG-  2451
Query_STENT  2085  INPLSGST-RASLLLLILHYAEYI[2]CS   -----TFANLPDLA[11]NDNEDSHRIALLK   LLLSLMDLlaNPSI  2158
Query_NEMAT  2385  LPTIHGVF-KAQCLEIVLSRVSDI   PD   LFTELRSKGLNVMF    EHRDEGVQVALLK   IIQGLLEK--LKMS  2448
Query_PISTI  2423  LPSVHGDF-KSYCLEILCSRVEHI   PN   LFTELKGRGLKEML    SQRDDGIQIGVLR   IIKHLMKK--LTIP  2486
Query_CORAL  2394  FQHVHAQF-KTLCLEIILSRVGVI   PN   LFLELKTKGFLNTL    THRDEATQVACLR   IILGRLSQ--MSVG  2457
Query_HYDRA  2293  TSNMFGTL-KTMCIQCLASRTEQL[4]QT   IFSQVKGKEIDYIL    NTNDEETQIGGLA   LIDGLLKG--LSL-  2359
Query_LATIM  2397  LPKLHGVL-KTHCLEVVMSRAEEI   PD   IFLQLKSKDFIQLM    NHRDDARQRVCLD   IVYKIMFK--LKPV  2460
Query_SHRIM  2356  LPKLYGIH-RKNVLECLIAHSATM   ED   IYLHLKEQNLLDIL    AKKEASTQLVALQ   LVNAVMP--rLTPS  2419
Query_MCRAB  2353  LPNLYGIF-CTRVLECLASYSSSM   ED   IYSHLKEHNVLGLL    SKKEESTQLVSLE   LVNAVLP--rLTPS  2416
Query_AMPHI  2384  LPKLYGLQ-KMRAMECLIAPLSTH   PD   CFQHLQQQRITDAL    QIREPATQLVLLQ   LLNSSLT--qLTAS  2447
Query_BARNA  2243  LPSLHGEQ-KALGLQVLHDIIGDV   PQ   LSSELAALRLDRLT    EARDPQLRRAALQ   LLLAAAEGdrLSAG  2308
Query_COPEP  2274  LKTLVGED-LSLCLEMMVSHASEL[3]PQ[1]LPTELNMIGLKNFL    QRNDPSSQSLALQ   IFKKCHSAldVGVR  2343
Query_DAPHN  2172  LTGLYGGL-KSRCLEVVARSWYYM   ED[1]-WKEIGFKTIESTL    KGSNPADLCPMLA   LVKDLIPR--VEVE  2235
Query_STARF  2432  LPSLHGQP-KTLCLEVVASRVSHI   PN   IFLEMKSKGIVELL    THRDEGTQLVSLK   LVNAMLPN--LKPD  2495
Query_STRON  2418  LPSIHGQP-KTECLQIITARIEHI   DN   AFIEMKNKNILALL    THKDEADQLTSLK   LVDGMLPK--LKPS  2481
Query_SEACU  2393  LPSIYGAP-RVQCLEVISSRVEKI   EN   VFQEVQSKGLKELL    THRDESTQVVCLK   LINGMLPR--LSPA  2456
Query_FUNGI  2366  LPRVNGDF-KTVALQLILDNVEHI   TD   VVDKLKP-ELIRLM    SQKNDAAHDVSFK   ILIRMLKAgiLGDN  2430
Query_DPURP  2470  LPQISNET-RLIVLNIVLWCVDDI   QD   LYLKLKSNNIENLI    RIREGEIQVVILK   ILYKLVQKngTPVS  2535
Query_HALBU  2467  LPQLSNEI-RLVAFNVIYWIADDI   PE   LFIKLKSANLDSLI    RIREEQTQIIILK   ILFKLIKNnyTTDQ  2532
Query_GLOBO  2246  YPKLDLDL-KGKSLEAILSWADAI   PD   LFTEILGLDLEDIF    LSQDEECQTYVLG   ILSILAP--tLSHS  2309
Query_PALUS  2307  LPRLYGRV-RAKCLYALSSSAPDI   EG   LFADLRAKNLQQLI    QHRDDESQAHVLV   ILLNLAP--sLRNE  2370
Query_CALIF  2385  FPNLSIDR-KIQALEIIRLRCQTR   KD[1]LFDELKGKELESLL    KNSSENIQYLTLD   IIYNIKD--vLSIN  2449
Query_GLOMU  2444  LPRLLNN--KILALEIISFCAGHD   PD   LFTNLQRDRLLGFL    KHRDENEQLLALR   ILAGILQ--eIEVS  2506
Query_LOBOS  2373  FPQLLGEE-RRLALEIFSGRAAHR   PD   LFQSLQGLNFLGCL    KHRDEGTQFAALS   IIYALWE--sLKVD  2436
Query_RACEM  2353  LPSLVGDL-KVLALQIVGRAVTEL   DD   VFKILESKGLLAML    GMRNVDAQIALLS   MLNQIAP--nLSSD  2416
Query_PVIOL  2464  LPQLTNDI-RSMALNIISWCSNDI   PD   LFQKLKSCNLDNLI    HIRDGDTQLVLLQ   ILNKILTNptVDSS  2529
Query_BOMBU  2226  VFRVDFQG-KSNGLEIFLLCIPNF[2]DQ   IVKELAHMKLHDL-[ 1]TNKILPCEKITLK   IIHSLVY--vLPPP  2291
Query_AEDES  2367  VPNTNGAP-KRILLEMLLSRI-EV[2]EQ   VHRELVSLDLKGM-    -LKDHQYQLLALH   LINKSLSLmsGE--  2429
Query_ANT__  2257  -ISANDLA-RSYCLEIFMLAISRF[2]EE   LVNNLRRMELQRI-[ 1]TNRESSCEKMALR   IVRDIVT--tVVPT  2321
Query_WHITE  2328  FTGTKNNNlQEKCLDLLQLSI-DS[2]SD   LCELLQEVQIIDLW[ 4]STSGSQCHLAALK   LFAKLLL-nnLLPT  2398
Query_BEETL  2218  IHGVPK------PLKIIYLRI-LL[2]CD   DLHKLSDFKTERW-[ 4]ENDHLEVQLLALE   IVHKSLNVlrFAPT  2282
Query_THRIP  2259  LNQIAG-H-RALCLNLLSKRL-EV[2]GN   PYQELKGCGLLTM-[ 1]GDQESLVQVKVLQ   ILRELALApsLQDS  2324
Query_TERMI  2342  LPKTYGIF-RTMCLKIVYSRL-EI[2]DD   TYAEMKSKNLFKT-[ 1]SDRDPDQQNLALS   IVHKMV--phLKPE  2406
Query_FLEA_  2262  ISRAIDLQ-KTMLLQILLSVL-EL[2]EE   FYSDIKTLDFDLL-    -LKDKDTVILALK   IIHKLLPFlqNDLN  2326
Query_MIDGE  2324  IPKAIGKI-KTIYLEMFLARL-EI[2]EF   VYREIISMGIKNL-    -LKQNEFQLISLH   IANKSLEYlkND--  2386
Query_LAODE  2319  LSEVNGYL-QQHVLDIVNTALELD[3]EN[5]LYLGINSQLINVFR[ 2]---NVDTQEAALK   VVTALRDF--LDYD  2389
Query_LOTTI  2358  LPNLNGES-RRLCLEVVHGRIGQI   DN   VFMELKTKGLLSFL    TQRNEKTQLVALN   IITSIREK--LKSK  2421
Query_APLYS  2417  LPNLHGEF-RPACLEVILGHVDQM   ED   AFLEINSKGLPSFL    THRDEDTQLVSLK   IVKALALK--MSPA  2480
Query_SCALL  2416  LPKMHGEF-RSLSLEIIHGCIESI   DN   VYMELKSKGLISYL    THRDEQTQLIALK   IVKAMMVK--LEAK  2479
Query_CRASS  2380  LPNLHAEF-RNRCLEIVLTRIDTL   EN   VYVELKSKGLLAFL    THRDDETQVIALK   IVKSIQSK--LQPS  2443
Query_OCTOP  2328  LTKLHGDF-QLYCLEILLGNINNV   DG   AFRELKSKNFFSFF    AHRNKNIRLISLN   IVKALLAK--LKPN  2391
Query_POMAC  2380  LPGLHGEF-RVQCLEVIMGCIEQL   DD   AFRELQSKGLLNFL    TQRDEDTQIVALG   IVNAMVPK--LSAT  2443
Query_WHIPW  2292  LPTLHGRL-KALCLEMFVSFVDQR   ES   LFIDLKGHGFLDIF    SSHDHVCQLASLK   LLQSLLPQ--LSTA  2355
Query_TRICH  2311  FPSLHGDL-KTACLEGFFAVSKYK   GD   FFEELKSRGFLKVF    SLRYERDQMLGMY   IVKNLHAN--LTVE  2374
Query_AASTA  2182  LPRILIQD--TLSLHLLDIVH-AC[1]VD[3]LFRVVRPH-LPRLL    SHHNPAVTLRLLD   VLQKWWPH--LTAD  2246
Query_AEUTE  2125  FPRIWVQD--TLLLCTLDIVY-SS[1]FQ[3]IFRFLRPH-LPRLL    RHNNPDVPLRLID   VMTKLWPN--LSED  2189
Query_PLASM  2296  LSKAVSVD--AWALLAADTVYNAC[1]ND[5]IFTHVQSV-LGRFV    AHRHSGVQISTLR   AITRLLDH--LTMS  2363
Query_BREMI  2282  LPRAILVD--AWALFAIDIINHAC[1]ND[5]IFTHVQSV-LGRFV    AHRHSEVQLGTLR   VIKCILNN--LTVS  2349
Query_OLIGA  2314  LPKVIAIE--AWSQSAVHCLFLAL[1]NE[5]IFTHCQQV-LGRFV    LHRSAVVQLTALQ   LIERLLDH--LDTS  2381
Query_LIGNO  2495  LPHLTGDS-LADCLVCLAAAATAA[4]DN[5]AAVELKERGFFDSL[17]SRVPQNSLALCER   LLLLLQQQerLADA  2586
Query_CESTO  2309  IPESIEPNsLVHCLRVFTTLLEDF[1]EN[1]LEFSLDQRTKIEVL[12]----KSENAVVLF[5]ILRMVQLLknIPQN  2390
Query_OVI2_  2678  LNQLSGST-----LNAYLNLLKQF[1]AD[3]VSSTMVASLTQCLF[20]SLKSSDTVKLGLE[5]-IRTVVSYskGLPK  2767
Query_PRIAP  2426  LPSLQGSF-RSQCLEVILTRVAHT   QN   LFLELKSKGFLGMI    THRDEATQLASLH   ITNALLSS--LSPG  2489
Query_ROTIF  2359  LPQLYGDF-KKMCAESILSSIRVL   QD   PL--FKSKNFIDML    THRDTPIQLICLK   MIHETLDK--QTDE  2420
Query_SPONG  2413  LPSLQGVF-RSSCLEILDTQADVQ   PQ   FYLKLKDKGFHTMI    SHRDEATQTASLS   IVQKLLVQ--LTND  2476
Query_TARDI  2428  MPTLPGVI-KAHGLDVLLAGADAM[1]-D   PYRDFKSCGIFACL    KDGNETVQGLTVE   VLRKVAPR--ISAS  2491
Query_TRICH  2375  MPSLYGEF-KTKCLEIILARADHI   KE   LYTELRNKGLKDML    LHRDESKQQVSLD   LVRAMLKN--LDIS  2438
Query_OSTRE  2328  IDKVYGEK---KCLALMIVTRSED   -E   ---EVVRSTYERLL[ 7]RQKDEQLLEVLFA   ALTRGVKVatIESS  2394
Query_NITEN  2494  VMSLPGDF-KILVLNMLLGYAANL   LE   LFEYLTP-ALRSLL    EHRDEQAQLKTLQ   ILEAILP-siSRER  2557
Query_CLUBM  2261  LPRLHGPS-KTVALDILLRDPATA   NS   YFNMIVS-FLSKLL    RHRDPKGQLRSLQ   LFAEVVK-geADPN  2324
Query_MOSS_  2384  LSRLFGAY-KVNALDTLLRFPKAD   AN   LFNQISR-SLPKLL    AHRDEIAQLKILQ   LLSQLLK-nvNQDT  2447
Query_LIVER  2415  LPKLPGVL-KVITLDIFLRFPSCL   RS   LLSMLLP-TLTKIL    THRDEFAQLKCLQ   LLLELMK-dvPDQS  2478
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2446  Q    LLYLLP-AICE   ISDLSSVLCRREMYQILMWCYDNY-KDEAQ   SES--[2]EIITLTKDVLLRGLS-DDDL  2509
Query_LINGU  2446  E    VLYFLP-AVTA   FSSSPSMPCRGLMYDILMWIYDNY-RDE-N   NAT--[2]QIMSRTKDALLNGMS-DKDL  2508
Query_LIMUL  2437  N    LLEVLP-MIIK   LQSQQNPLIREVMYQILMDVYANF--SMKS   GAKEK   EITRLAKETLLLGLA-DQNI  2499
Query_SPIDE  2431  D    ILHILP-SIVG   IHTQPNSHSRLALFKILFLLFEKF--SDSS   GAKEK   EITRYSRETLLNGLA-DPDP  2493
Query_SCORP  2418  K    MLNVFP-IISK   VHKSAETICRKVMYEIMMHAYEKF--NTKN   GIKEK   EISRYAKEMLLHGLA-DSDL  2480
Query_TICK_  2356  D    WGAILP-VVCQ   LQGHTSGACRKLVYDVCMRLYAKY----RS   KTSAA[1]-LAKQSKHVLLVGLG-DPED  2416
Query_MONOS  2404  V    VQRYFPaLIAH[3]---HDDAQVRLLHYQLLMVLYDLTlTAELK[3]-GESN[1]-VQRETLRLLVMAFS-DESA  2471
Query_ROSET  2370  S    IKQFLP-SVTA[3]--SHDDYTVRAIAYDILMAVYDQK-LSHER[4]LADDE[5]EIAQLARTHLLLGLQ-DSSP  2443
Query_BELCH  2486  E    LLYLLP-SVAA   FASHPAPECRRTMYDIMMWIYDNY-REEAS   QQESA[1]RVLGTAKDCLLQGLS-DEES  2550
Query_CIONA  2452  T[11]VLQFLPeIFKV   FSKHSNPICRKEAFELAMMLYDHQ----KS   LSSTD[3]EIMKLSKIFLLKCLS-DPEE  2527
Query_CAECI  2463  E    LGELLN-TVVT   FSSHPSPMCRERMYDILMWIYDNY-RDPES   QANDD[2]EVFSRANDVLLQGLI-DENA  2528
Query_HUMAN  2450  E    LRELLN-PVVE   FVSHPSTTCREQMYNILMWIHDNY-RDPES   ETDND[2]EIFKLAKDVLIQGLI-DENP  2515
Query_MOUSE  2446  E    LRELLN-PVVE   FVSHPSPTCREQMYNILMWIHDNY-RDQES   QNDED[2]EIFKLAKDVLIQGLI-DENV  2511
Query_XENOP  2471  E    LHEFLI-PVTA   FSSHSFPVCRERMYDIFMWIYDNY-RDHES   QNDSK[2]EVFNMAKEGLLQGLV-DENT  2536
Query_PLATY  2454  E    LRELLN-VVVE   FISHPSPVCREQMYNILMWIHDNY-RDPES   QADEE[2]EVFKLAKEILLQGLV-DENH  2519
Query_TASDE  2456  E    LQELLN-CVVE   FISHPSPVCREKMYNILMWIHDNY-RDLES   QTDEV[2]EVLKLAKEVLLQGLM-DDNL  2521
Query_CANAR  2452  E    LKELLP-GVTG   FISHPSVMCRQRMYDILMWIYDNY-SDPES   QADDD[2]EILKVAKEVLLQGLI-DENA  2517
Query_OSTRI  2422  E    LKELLP-GVTG   FISHPTVICRQRMYDILMWIYDNY-SDPES   QADGD[2]EVLKLAKEILLQGLT-DENA  2487
Query_TURTL  2461  E    LRELLP-GVTG   FISHPSPVCRERMYDILMWIYDNY-RDPES   QAEEE[2]EVFKLAKEMLLQGLM-DENS  2526
Query_GECKO  2461  E    VQELLP-TVTG   FITHPSPLCRGRMYDILMWIYDNY-RDPES   QADEK[2]EVFKLAKEVLLQGLI-DENS  2526
Query_DANIO  2459  E    LKEILG-AVTA   FASHPSPVCRERMYDILMWIQDNY-SDSES   REDST[2]EVLSVARETLLQGLT-DENY  2524
Query_SNAKE  2459  E    VLELLP-SVTG   FISHPSPLCRGQMYDVLMWIYDNY-SDSES   QADEE[2]QLLKVAKESLLQGLI-DENS  2524
Query_STERL  2473  E    LKELLI-SVTA   FTTHPSPVCRERMYDILMWIHDNY-RDPES   QSDSV[2]EVFNLAKETLLQGLT-DENQ  2538
Query_MILII  2460  E    VKVFLV-PVIA   FISHPSPVCRERMYDILMWIYDNY-RDAES   QSDPD[2]AVFVTTKDTLLQGLT-DDNP  2525
Query_BAMBO  2455  E    LKELLR-SVVA   FISHPSPICRERMYDILMWIYDNY-RDPES   QSDPD[2]EIYTETSDALLQGLT-DENP  2520
Query_ALLIG  2419  E    LRELLP-GVTG   FISHPSPVCRERMYDIMMWIYDNY-RNSES   QADED[2]EVFKVAKEMLIQGLT-DEIS  2484
Query_GAR__  2462  E    MREMLG-AVTS   FVTHPSPVCRERMYDILMWIHDNY-RDSES   QSDSD[2]EMLNVAKETLLQGLV-DENQ  2527
Query_TETRA  2482  D[ 3]--QLMRiSISI[4]YERHYQEKFATAYYTIVYNLYDYLiNQESD[4]NQINE[5]QFKSQVKSCILCGLT-SQHN  2561
Query_PARAM  2369  T[ 6]FNQILD-KIDQ[1]YGVDRDIGIVKIYQNILKLLHEYL------[1]NYDHK   ELQGKVTTYILSLFT-SQHI  2435
Query_SLEMN  2452  D[ 3]-QRVLP-RLKQ[1]FQHNKNDYSRGLFYDLMVYLYDQFeNFKPL[2]-----   -------SALIRGLS-DKTK  2509
Query_STENT  2159  K[ 3]-SGIIP-NLSS[1]ASASSKDI-RKLLYEVMQKAYDESkDMDES[2]-----   -IRSKAREVLLKGLS-DRE-  2220
Query_NEMAT  2449  E    IEYVVP-SIAT   MATSSSTACRELIYDIFMWVYDTY-RSDET   FRNE-[4]EILATAKDYLLRGLA-DDNK  2515
Query_PISTI  2487  E    LKFIFP-LLTS   FVSHPNLLCRDLMYDIFIWIYDSY-RADEA   FQAE-[4]DILADAKDHLLQGLA-DDHR  2553
Query_CORAL  2458  D    IQTLLP-SVTD   FLSSKSVACRGLMFDILIWIYDTY-RFDET   FKNED[5]DVLCRTKNVLLQGLA-DENK  2526
Query_HYDRA  2360  -[ 6]LKKIEPlQKSI   FVK-----VREIVYKIFMKCFDHF----KE   KREEE[3]LILQICKETLLVGIA-DKNT  2424
Query_LATIM  2461  E    MKEFLV-AVTG   FISHPSTVCRERMYEILMWIHDNY-RDQES   QANNE[2]EVLKLASEVLLQGLT-DENP  2526
Query_SHRIM  2420  Q    LLYFMP-GVVA   FASHPAPSCRETMYDVLFWVYDNF--SNND   TEDGH   QLESQARGVLLRAVK-EEDA  2482
Query_MCRAB  2417  Q    VLYFLP-GITG   FASHPGAKCREMMYDILFWIYDNF--RESS   DSEGS   QIEAKTRVSLLQAVK-EKDA  2479
Query_AMPHI  2448  Q    VLTFLP-FLVS   FLKHPVSSCREVVFTSLTFLHDRY--QCES   SEDQK   EVYCTVHDALLQAVA-DEDT  2510
Query_BARNA  2309  Q[ 1]-ARFLP-AIDF[1]-VNHGGAPGRRLAAELCIELRRRFdRDDSE   --EAA[1]-VRARATDALLVLLT-DADL  2371
Query_COPEP  2344  K[ 3]--ELLK-----[1]FTSSKYTEQRKHLYEIFK---DGY-LDPTN   -----[1]DLKSYCLEQLMFGLT-DTDE  2398
Query_DAPHN  2236  E[ 1]-KPLWPtLIRL[1]--RHQNVECRSRLYEVFKKAFDLS----GS[1]--DSE[2]-------QAVLKACS-DQDE  2290
Query_STARF  2496  E    ILYLLP-PVRT   VFNLPSPACREVMYDILMWIYDNF-RDEES   RTEDG[2]DILCMAKDTLLQGLS-DEEA  2561
Query_STRON  2482  E    LRYLLP-AVKD   FVSHPSPSCRDTMYSILMWIYDNY-RDEDK   REGQD[2]EALSIAKDTLLQGLS-DQDN  2547
Query_SEACU  2457  E    LTTLLP-DVEA   FTSVSSIGCREVMYDIFIWIYDHY-RDEVR   MEEEG[2]KVLGMAKDALLLGLS-DKDP  2522
Query_FUNGI  2431  D    IKPLLEaIVKS[1]VTRTTSQVKNSM-YELMELLYEKD--P--L   YRSGE   ---QDLTSQLIMGLS-NE-E  2488
Query_DPURP  2536  T    IGQILQlLLLN[3]TSNITSEYSRSLFYDIVIYIYN-N--F--S   EFQDD   ---KQLILSLLIGLS-DEGE  2596
Query_HALBU  2533  T    LLQIINlLLD-[2]FNQSPNETCRLLYYDIFMWIHDYL--E--K   DVIDQ   ---NQIRLTLLIGLS-DDSE  2592
Query_GLOBO  2310  Q    TAKFLGlAVDL   FSIHGSDRVRAAFFSMIFKLKESP--AVEH   DAVLK   ---RLVKYSLLKGLS-DSDN  2370
Query_PALUS  2371  D    VSYFLAtAIET   FSGHPSAKCRSAFYTLMMRLHERT--DLE-   -AKIN   ---RVLKLAILQGLT-DPTE  2429
Query_CALIF  2450  Q    IKFFLNnS-NH   FIKHSSLRCRNIYYAILKYIYQV-----NM   DEEIK   ---NSVKVSLLNGLA-DSDE  2506
Query_GLOMU  2507  T    FEYYLDgLIES   FQDHPNNDCRRTYYTILVKMYNK----ING   TVEIK   ---HKLKINLLRGLI-DTDE  2565
Query_LOBOS  2437  Q    ILYFLDtIVLE   FSHHPSVECRKLYYSILMTLFDK----HST   TPEVA   ---NVLRTQLLRGLG-DTNE  2495
Query_RACEM  2417  N    ISTMLDtLVDI   FAKHVNVDCRALYFSILRKCYDT----LMD   DSPLK   ---IRINSEILKGLL-DGDK  2475
Query_PVIOL  2530  Q    KQQLVNlLKEQ[4]TSSTTNEHCRLLYYEMIMYIYHQN--F--N   DADQD   ---RDIILCLLIGLS-DESE  2592
Query_BOMBU  2292  N    ILSFITlVP--[1]YTKHKLSEYREIAYSIFINIYKKY-TGDSS[4]I--KE   -LMQISKEILLNGIL-DPTE  2356
Query_AEDES  2430  N    LENLLEnIV--[1]MTGNDRADVRDIVYEILI--FLQ-----EN[3]VLSSE   -SQQKLKRALLAGLS-DPEQ  2489
Query_ANT__  2322  D    LLPYVNlVI--[1]YVKDSITEYRELVYDILMKIYKRY-STDIT[4]VIVQN   -LLSISVQNLLVGLL-DPSH  2388
Query_WHITE  2399  D    LQKKVIaVFLN[1]VSSSKRDESRKIIFETVSILYSRYcNNPA-   -IEDN   -IKNRCKEILIRGLV-DRNA  2462
Query_BEETL  2283  -    FPEIVQsV--C[1]NVSNTNVLCRTKVFDVVMAICET-------   ----G   -ILPEVKPVLVQGLC-DEDT  2334
Query_THRIP  2325  E    LDEVIHhI--S[1]INCYVKGEARNLVHEISLIAFDRYlPMHES[4]EIEKK   -IFDESKALLLLGLV-DSNA  2392
Query_TERMI  2407  Q    LEELVHeI--C[1]FETQRSVKCRSVMYEILKWIYDH-------[2]DIGGV   -LIEEAKGILLKGLR-DDDP  2465
Query_FLEA_  2327  R    PSTLIQyAK--[1]LISSEDTNCRSVVYDIAMTSYKI-----IQ[4]Y---D   -ILQNVQYILLLGLL-DSDA  2386
Query_MIDGE  2387  E    IESLMNdLM--[1]LVNSPRDDVRRILYEMMI--FIV-----ER[4]EMSSL   -IRKKAIEIILKGFI-DNNV  2447
Query_LAODE  2390  N    LAGLMLeVVKH   FSNHSRLACRNAYYDVLMNVYEYY-QTTPC[4]EVSWS[5]SAKDLIKNALVSGLE-DANE  2463
Query_LOTTI  2422  E    MSSLMD-SITA   FSHHPSTLCRKVMFELLMWIYDNY-RED-E   CTE--[2]EIMLKTKECLLKGLG-DEDM  2484
Query_APLYS  2481  E    LLSLLP-SVKS   FSSHSSSACRTVMLEVFMWAYDNY-RAG-E   DEE--[2]RVMSLTKEALLRGLC-DENL  2543
Query_SCALL  2480  E    LMSLLP-LVTA   FSSHPSTSCRTIMYDVLMWVYDNY-RGD-D   NED--[2]IIIQQTKEGLLRGLG-DEDL  2542
Query_CRASS  2444  E    VSELLP-LICG   FCVSPSIPCRNVMYDILMWVYDNY-RED-E   SET--[2]TIMQQTKESLLKGLG-DEDM  2506
Query_OCTOP  2392  E    LAEIFS-SINH   QKEEPWVACRKVLYDILMWVYDNY--GEET   SEESS   QIMSNTRELLLKGLE-DKDY  2454
Query_POMAC  2444  E    LQTFLP-AITG   FSNHPSVTCRQIMMKILMWTYDNY-RDN-E   SKE--[2]EIMKVTQESLLQGLC-DDNL  2506
Query_WHIPW  2356  Q    IEELVR-VLSS   FVSHPLNSCRRCFAELCCSMMDKL-RSSFI   LSADS[5]SAMESLTEHSLSLLA-DKEE  2424
Query_TRICH  2375  N    ISDLVS-SFSS   VIGNFSPSCRRAFAEMCCSLYTKL-SSL--   --NDS[5]AALATTKMHLLTMLS-DSNK  2439
Query_AASTA  2247  M    HQVLL----DG[7]FESSEPCRMKVLDFLLA--------TTKGG   --VPN   -----VQRILLKALT-DANA  2300
Query_AEUTE  2190  Y    QMVLL----SG[7]FE-NESCQMKVLQFVIS--------IDKLS   --VPA   -----VQLISLKALS-SRIA  2242
Query_PLASM  2364  E    LNRLVAsTSEG[7]YQGHALSVCRGLLFSIAKNLYHKD-ISDPG   --KLK   -----VRISLLHGLC-DPDE  2428
Query_BREMI  2350  E    IDRLVAnALEG[7]YEGHT-SESRALLFTLAKELYHIG-ISDQS   --KSK   -----IKVSLLHGLC-DPDA  2413
Query_OLIGA  2382  V    ACRFIMrEDQG[7]TVRERDTPVRRIFYDIAKKLYQKN-LDEPL   --KDR   -----LRLALLQGLA-DPDI  2446
Query_LIGNO  2587  E[11]LAAVGEmCLAS[7]VAALTLPQARRQAYQILATWYPTH------   -----[1]---NLCAPALLAGLT-DSAE  2657
Query_CESTO  2391  N[11]VLRLMQfLLSS[4]---NSTPSSRRIAYRIFIVAKTNE-KSTEN   -----[1]EIMEICQLGLSYGLCaEDDN  2464
Query_OVI2_  2768  Q[10]YRYQIQsVLSS[4]FSLNAPRDVREAFYQICITVWEFR------[1]ELEPA[2]--MHLAKFGLLYGLTvETDL  2843
Query_PRIAP  2490  E    LLFLLP-AVTS   IGTHPRAPCRQLMYDILMWVYDSY-RDSEG   EEE--[1]KLASLARDSLLQGLN-DSDD  2552
Query_ROTIF  2421  D    LDRLMP-HLEK   FIHHPHLSCRYQFVLILISLYEIY-QFKVN   QIGST[1]QILNFSKENLLKALI-DEDT  2485
Query_SPONG  2477  Q    LSEIVPiLIAS   FSKHGSIQCRIVLYNILMRVYNKLcTDDQS   ESDDS[2]MIRSISRDQLLLGLG-DESN  2544
Query_TARDI  2492  D    LAEVLL-EMVD   LGSHGSPLMRQKYYEFCMYLYERF--ENST   VADEQ[1]-LLQIIRTRLFTGLK-DRDK  2554
Query_TRICH  2439  Q    ILYLAE-LVAT   FKTNPSAACRQKMYGIFMWIYDTF-RQEID   KNDE-[2]QLINFCKDQLLQGLN-DSSE  2503
Query_OSTRE  2395  H    LKQWHSaLIEA[3]VDTESVVNVREAHTKMCIAVAEAF---PEL   -VNEA[1]-----LRGPLLQALA-DPVS  2455
Query_NITEN  2558  V    ETVLLPiLCPV   FSRHDNPECRREFYRILEFLFRPA--PDGK   GMPDH   ---PQVQRHLLTGLC-DPAD  2618
Query_CLUBM  2325  S    TQKVVLeTCDA   FVHYETEECRLQFYQMLMFLYET------K   KLPHD   ---SRIARFLVLGLC-DVAD  2381
Query_MOSS_  2448  I    RTRVLPvLCDN   FNGHENVDCRREFYEILIYLFKN------K   GMDKE   ---RLVIQSLLMGLC-DSSD  2504
Query_LIVER  2479  I    HEKVLHvVCET   FPTHESVDCRRLYYDLLICLHAI------E   GLKSD   ---NLLIRSLLAGLT-DSAE  2535
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2510  TNRLTVQ    NFWS-HETRLPG    ETVERLVAMLEAM--Y-TPSTELH    YLSYATNLLLEMTTKSSDF-NRVLF  2572
Query_LINGU  2509  GLRLRVM    NFWS-HESRLPA    GTLERMVAMLEAM--Y-SPSVESS    YLSYATNFILEMTSKSPDY-NREIF  2571
Query_LIMUL  2500  ALRLLVN[57]NFWS-NEKRLPL    GTVERLIAL---------------    ----------ELTSLSPDY-NRKIF  2597
Query_SPIDE  2494  SIRL---[ 3]NFWS-NEKRLPL    GTVERLIALMKNM--S-SPVTEES    FLQYSTFLLLERTSYSPDY-KRKIF  2556
Query_SCORP  2481  TLRLTVH[48]NFWS-NEKRLPL    GTVERLIALMKNM-----------    -LQYSTFLLMEMTSKSPDY-NLKIF  2582
Query_TICK_  2417  SLRLSVA    NFWS-QETQLPL    KARERLVALLAGM--Y-SPEIEDH    FLSSVTFLLLELASRSPDY-GRKLF  2479
Query_MONOS  2472  IIRRSAE    EFFD-QETRLSI[ 1]-TDLRLCEVLDKL--Y-VPASEGS    FLSTCSSLLLRLTARSPDY-ERQLF  2534
Query_ROSET  2444  KLRQSIV[ 1]-YWD-QEKRLAR    NPVQRFYDGLTSL--Y-VPEAEEA    YLQFVTNVMLQLTKHSPDY-DRELF  2506
Query_BELCH  2551  YLRLLVQ    NFWS-HETRLPT    GTLERLVAMLECM--H-SPRTEPQ    YLSYCTNLLLEMTSRSPDF-SREIF  2613
Query_CIONA  2528  LLRNRAM    NFWS-DESRLPI    TPIHKVSGIFQDL--YdSSLEESR    FLCNAVGLILQSTSLSPDF-DRAIF  2591
Query_CAECI  2529  GLQLTVR    NFWS-DETRLPV    VTTDRMLALLNSF--Y-SPKIESH    YLSLATNFLLEMTSKSPDY-TREMF  2591
Query_HUMAN  2516  GLQLIIR    NFWS-HETRLPS    NTLDRLLA-LNSL--Y-SPKIEVH    FLSLATNFLLEMTSMSPDY-PNPMF  2577
Query_MOUSE  2512  GLQLIIR    NFWS-HETRLPS    NTLDRLLA-LNSL--Y-SPKIEVH    FLSLATNFLLEMTRMSPDY-LNPIF  2573
Query_XENOP  2537  ELQLIVR    NFWS-DETRLPS    NTTERMLAILSSL--Y-SPKIEKH    YLSLATNLLLEMTSKSPDY-IRKMF  2599
Query_PLATY  2520  GLQLIVR    NFWS-NETRLPS    NILDRMLA-LNSF--Y-STKIEMH    YLSLVTNFLLEMTSKSPDY-SRHIF  2581
Query_TASDE  2522  GLQLIIR    NFWS-HETRLPS    NILDRLLA-LRSL--Y-STKIEMH    FLSLATNFLLEMTSKSPDY-SRHIF  2583
Query_CANAR  2518  ELQLIVR    NFWS-HETRLPT    NTLDRMLSLLSSF--Y-STKIETH    YLSLITNFLLEMTSKSPDY-SRKIF  2580
Query_OSTRI  2488  ELQLIVR    NFWS-DETRLPA    NILDRMLALLNSF--Y-STKIETQ    YLSLITNFLLEMTSKSPDY-SRKMF  2550
Query_TURTL  2527  GLQLIVR    NFWS-DETRLPT    NTLDRMLALLNSL--Y-STKIETQ    YLSLATNFLLEMTSKSPDY-SRKMF  2589
Query_GECKO  2527  GLQLIVR    NFWG-DETRLPG    NTLDRMLTLLSSL--Y-STKTETQ    YLSLVTNFLLEMTSKSPDY-SRKMF  2589
Query_DANIO  2525  GLQLYVR    NFWS-HESRLPP    ETLERMLVVLKSL--Y-SSRIEEQ    FLSLATDLLLEMTSHSPDF-TRNMF  2587
Query_SNAKE  2525  GLQLIVR    NFWS-DETRLPT    NTLDRMLSLLRGL--Y-SDKIETQ    YLSLATNFLLEMTSKSPDY-SRKMF  2587
Query_STERL  2539  GLQLYVR    NFWS-HESRLPT    STLERMLVVLRSL--Y-SSKIEAH    YLSLATNLLLEMTSRSPDY-TRDMF  2601
Query_MILII  2526  ALQLYVR    NFWG-HENRLPT    GTHDRLMAILRSM--Y-SNKIETQ    YLSIATNLLLEMTSRSPDY-TREMF  2588
Query_BAMBO  2521  ALQLYVR    NFWS-HENRLPT    GTLDRMLASLGSL--Y-CNKIETQ    YLSIATNVLLEMTSHSPDY-SREMF  2583
Query_ALLIG  2485  ELQLVVR    NFWS-DETRLPT    NTLDRMLALLNSF--Y-STKTETQ    YLSLATNFLLEMTSKSPDY-SRKIF  2547
Query_GAR__  2528  GLQLYVR    NFWS-HETRLPT    TTLERMLVILRSL--Y-SSKIEAQ    YLSLATNLLLEMTSRSPDY-SRAIF  2590
Query_TETRA  2562  TIRENTF[ 5]---EtNDSKTPY    ---ERLLYLIQEI--Y-QPEQERL    WLLSAVPLIMCLSKRTSDY-DKNLF  2624
Query_PARAM  2436  EIRNLCF    KYLQdDLERKAI[10]TQIDRVISLMKDI--Y-QPSKETL    WLKGAVPLLLSQSTKTQEY-DAFMF  2509
Query_SLEMN  2510  VVRNKLI    SFWSdS-SRLGL[ 1]-PNQRLLQIMNDL--Y-TQEEENI    WLNNSVFLLMQICSQSSDF-NRKIF  2572
Query_STENT  2221  -FSDGIM    KFWDhQ-DHLSL[ 1]-AQMRMLQCLSEI--Y-TPECEEL    WLVASSHLLVRLTELSTDY-NKVLF  2282
Query_NEMAT  2516  ILKLKLQ    NFWS-DETRLPN    TTLERLVDVLRAM--Y-STNTEQQ    YLSYTTNLVLHLTSFSPDF-ERLMF  2578
Query_PISTI  2554  DLQLKLR    NFWS-DEARLPG    TTLERLVEVLRAL--Y-STTTEIQ    YLSYTTNLILHLTSLSPDY-ERLMF  2616
Query_CORAL  2527  EMRLKLR    NFWS-DETRLPN    STLERLVQVLRAM--Y-SPITESE    YLSYTTNLLLHLTSQSPDF-GRLIF  2589
Query_HYDRA  2425  DLNLVLL    NFWR-KE--LSE    NTIERFTEILTTF--Y-SSKTEKE    YISYSTNLILQLTSLSPDY-SLPLF  2485
Query_LATIM  2527  GLQLYVR    NFWS-DENRLPV    GTLDRMLVILKAF--Y-CNKIETQ    YLSLATNLMLEMTSRSPDY-TREIF  2589
Query_SHRIM  2483  ALRQIVL    NFW--LQGARSM    TLTQCLLHILRKM--Y-SPESEDS    FLQMAIYVLLEATRHSADY-QRNIF  2544
Query_MCRAB  2480  ALRQTVL    NFW--LERFQDM    VLHDRVLHILKMI--Y-SPETEDS    FLELAMYIILEATRYSADY-QRDIF  2541
Query_AMPHI  2511  ELRNFIL    DYW--TRRAQSL    SPHELLVFVLRDL--Y-SVSSEDT    YLHTVCHVLLQITHSSPSY-ARPLF  2572
Query_BARNA  2372  TVRQTAA    NYWA--DAGADT    -TPGRLSRVLTDM--F-RPGLEAA    YLSYATYLVLGLTAQSPDF-DRKMF  2432
Query_COPEP  2399  GLSEFCS    NFFN--LEFLSS[ 1]DSLDRMRNILHQM--Y---FHEHE[ 2]LLAFLPASILSLAALSPNY-SKKVF  2461
Query_DAPHN  2291  ELAVKMQ    NFLA---EKLPN    GTMDRLLAYVTDF--Y-SPGAEDY    FLPYCLHFLLERTTQSHLY-REPIF  2351
Query_STARF  2562  SHRLRVS    NFWS-HETRLPT    GTLERLVAMLECM--Y-SPGTEHQ    YLSYASSLLLEMTSKSPDY-KREIF  2624
Query_STRON  2548  YLRLRVS    NFWS-HETRLPK    ETLERLISMLGCM--Y-SPGTENR    YLGYATNLLLEMTSKSPDF-KREIF  2610
Query_SEACU  2523  YLRLRIN    NFWS-HENRLPV    DTLERMVALLECM--Y-SPSTEDQ    FLSYSTSLLLEMTSKSPDY-KREMF  2585
Query_FUNGI  2489  SISNRLL    LFWD---RLLPE    KSSERLYALFQAM--Y-TPDTETH    WIHYSTLLLMKLTARSPDY-DRLLF  2549
Query_DPURP  2597  SINKKLL    EFWDnNTKTLSS    NSNQRLQQFFTIM--Y-SPETESK    WVGNNCCLLLQLCNRSSDF-CKVLF  2660
Query_HALBU  2593  AVKRKLL    EFWD-TSKELSL    STVQRLMQLFEKM--Y-TKETESQ    WLGYCCSLLLQLCTRSSDF-TKLLF  2655
Query_GLOBO  2371  GIRSSLL    EYFN-SKVFVNQ    TVFDRAMDLVCDW--Y-IPEVEDS    FLKYSVLFLLDATRLTPMF-DQKVY  2433
Query_PALUS  2430  EIRGAVT    SYLR-QGGRMGQ[ 1]DIFSRTEFLLGSL--Y-EPEVESS    YLNFSTFFLIDVAKEAPQY-GTNMF  2493
Query_CALIF  2507  EIRNNIL    SIWH-DSSLVDS    -TFDLFDRVLNSL--Y-SSETESI    YLSYATPILLLSLS---DL-ENSLF  2565
Query_GLOMU  2566  SIQQALI    EFWH-AQEELSH    NTFTKLKELIGNOL--Y-STESENL    FLNYACYLLLEGSKKSIEY-NKPIF  2628
Query_LOBOS  2496  SIRHTVV    EFWY-SKDRLPS    NTFERLGEIIRNM--Y-DVQAEDI    FLNYATYMLLDRTKKSPDY-TEPIF  2558
Query_RACEM  2476  TIQQTNV    EFVE-ARHGLKE    EIFDVIQGPLSTA--Y-TPSVENM    YILYGTYILLESTKRHQDY-GSLLF  2538
Query_PVIOL  2593  MIGKKLF    EFWD-NVNSSNP    STIARIIQLFTTM--Y-APETEGK    WLSNSCCLLLSICNKSPDF-SKLLF  2655
Query_BOMBU  2357  TLQEMIL    NFWTqDAKL-TN    TCKERLLEAVDIY----TPNADQN    FLPFIFLLITDLIKKSNNY-TQKIF  2418
Query_AEDES  2490  SIQNRMF    DFWTnE-SRFPT    DIDNRFRRVLSDL--Y-EPSKEAD    YLGYATQILLDSAVKNPES-KRRIF  2552
Query_ANT__  2389  DLQDKIL    KFWTeETSLSTE    KSKDRLIALLAMHsqV-TTEVDA-    FATFVALMMLQLTTKSIDY-TSKMF  2453
Query_WHITE  2463  EIQKSLY    EFWMkD-NVLPE    KSIDRLRFVISNM--Y-IRDIEYA    FLGYCLRFVLDCAEQTLSC-RENIF  2525
Query_BEETL  2335  EIRERVY    NFWRnN-SSVPH    PVADRFLYLLKNF--Y-EESTEND    FLGNCCYFLLERLKFADDY-SKLLF  2397
Query_THRIP  2393  DLQKKSF    SKW-tDEERLPQ    NVAERLSTMFSLL--Y-SPLTEQY    FLSYCVPLLLAGTRSTPDF-DSPLF  2455
Query_TERMI  2466  HLQDSIF    EYWNsDAQQRDA    STSDRLLFILTDL--Y-CPATEHI    FLGYSTQYLLEATAGTPDY-NQKIF  2529
Query_FLEA_  2387  QMNDKMQ    TFWQsE-NALPK    TYDAIFKYLMSNL--Y-DEHLEKN    FIQYGVTILLQIALQVSDI-QRNIF  2449
Query_MIDGE  2448  EIRNRVA    NFLSlD-NELPK    DFIGRFRVLLEKF--Y-DPNIGRE    FLHYSTQLLLDISIRNPQS-KERLL  2510
Query_LAODE  2464  DLRNKMF    QFWN-NQGRVRT    DLVDRLKDLFELL--Y-SPDLESS    FLSLFLQLLLNVCQLNPAY-YSCVF  2526
Query_LOTTI  2485  MCRLLVQ    NFWS-SDTRLPT    GTLDRMVAMLEAM--Y-SPSTEQQ    YLSYATNLLLEMTTKSPDY-QREIF  2547
Query_APLYS  2544  ACRLMCQ    NFWS-SNTRLPE    STLERTVSMLEAM--Y-SPSTEPH    YLAYATNLLLQATSKSPDY-TRKVF  2606
Query_SCALL  2543  HSRLLIQ    NFWS-SETRLPP    DTLDRLVAMLQAM--Y-APVTEQQ    YLSYSTNLLLEMTSKSPDY-QRQIF  2605
Query_CRASS  2507  QCRLLVQ    NFWS-SETRVPG    NTLDRLVAMLEAM--Y-SPVTEKQ    FLSYATNLVLEMTSKSPDY-QREMF  2569
Query_OCTOP  2455  SCRILVH    NFWS-NRTRLPE    STFDRMVAMFQMM--Y-SPGTEHQ    FLYYSSSLLLEMTSKSPEY-KREIF  2517
Query_POMAC  2507  HCRLTMQ    NFWS-SESRLPD    GTVDRMISILEAM--Y-TPGSECN    YLSYATNLLLEMTSKSPDY-DRQIF  2569
Query_WHIPW  2425  PVRQIVF    SWWN-DEKRLSH    RSAARLQNVLG-L--Y-SRSPGKD    FLGQATSFMFYLSTLSPDF-TRLLY  2486
Query_TRICH  2440  SIRQLVY    NFWC-DASRLEI    STSARLLNCLSQL--Y-TCATEST    FLSHASYVMLELTCHSPDY-TRLIY  2502
Query_AASTA  2301  DVRGKVA    SFWE---AHLPS    SCDDRLLALFGQL--Y-HQETTDE    WVRYAPTLWLSLSTSTPDN-ANMLF  2361
Query_AEUTE  2243  DVRERAY    EFWE---KQLPS    SCHDRLLALFGPL--Y-HDQTTED    WVKYAPKLWLSLCQASSDN-KQFLF  2303
Query_PLASM  2429  ILRNEAY    NFWS-NFASIKK    SCFDRLLSIFGSL--H-SSELDEN    WILYATNILLGLSKECNEF-KQPLF  2491
Query_BREMI  2414  SRRKEAF    DFWN-TSAIAMN    SCSDRLLAIFGSL--Y-SPDFDDR    WVLYATNILIGMSENSKEF-KQPLF  2476
Query_OLIGA  2447  SIRTEVL    TFCN-GSELLPG    SCRDRVLELYQSM--Y-APGVEDV    WLLYATNTLMGLTRGSGSF-DSLLF  2509
Query_LIGNO  2658  SNRAWLL    EFWS-RNSGLPS    EASARWFALLSRC--HgNNTADED[14]LPAFCSAVQLDACSSTPDF-RRPMF  2735
Query_CESTO  2465  QLRRQLR[ 5]HYFN--SSPIPG[ 2]SSLSRCLQTFQLLsaShCSYTDNP[ 9]LISTSLQIVLEPAIKSAEF-RHPFR  2545
Query_OVI2_  2844  ELASNIR[27]NLWD----GLPE[ 5]-TQSLVLQMFPQLl-Y----TINA    -------LTLERATLTAEY-QQVFF  2925
Query_PRIAP  2553  SLRMIVT    NFWS-HATRLPT    QTLPRLLAIAQTL--Y-SPATEPH    FLSYATYLLLQLTSLSPDY-KELVF  2615
Query_ROTIF  2486  SIRVMAQ    NFWT-EKADMPS    NTIDRMLLIVEKL--Y-SPKTERE    FLSYSTNLLLEKTSKSPDY-NRLIY  2548
Query_SPONG  2545  EVKLKMY    EYWN-QESNLSS    NTLDRLTQLLVTL--Y-SKKTETQ    FLYHATNLLLELTSRSPDY-NRVLF  2607
Query_TARDI  2555  TIKDLLT[ 1]-FWA-QEKRMSG    STFMRLMTIFTSF--I-TPELETA    FLPTACHLLLGLCQKSTVFrDTKLF  2618
Query_TRICH  2504  ENRLQLI    NFWS-NNNRLPS    GTLERLSSVLYAL--Y-SPNAESE    FLSYSTNLILELTSRSPDY-KLSIF  2566
Query_OSTRE  2456  TTCRDSA[ 2]-FWN---TRLQY[ 1]SLFARLQDVLNLIp---NKKVESL    WLAAACRLMFAVQQMNLAS-GASLI  2518
Query_NITEN  2619  ELREEAQ    FFWG---LQLSQ    EMPARLVELLGRL--Y-EPELEDK    WAQYSDKMLLNLCEASTDY-SRPLF  2679
Query_CLUBM  2382  SVRQSLQ    SFWD---AQLPE    NMMERFSHVLNNM--H-LPDIESH    WTSFTNVLLLKLCQRSVDY-RRPIF  2442
Query_MOSS_  2505  IVRASLQ    NFWH---LQLSQ    DVRHRFLELVNRI--Y-DPCLEEN    WAQISNVLILKLCEDSVDY-KRPIF  2565
Query_LIVER  2536  IVRSTVQ    KFWS---SILSQ    NLRERLSQMINLM--Y-DPSVEEN    WAQFTNVLLLKLCEDSVDY-KRPIF  2596
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2573  DTPLSE-CKFQDF--AVSSTWYQRH-AAM-TPMFLDTQSS---Q--SNDSPDA    GLST---------GQLRAT---Q  2626
Query_LINGU  2572  EHPLDD-CPFEDF--VVRSSWKQRH-AAM-TPLFAETQFS---L--ADGTQSG[ 2]DDSF---------VGVRAT---Q  2627
Query_LIMUL  2598  PNPLSD-CQFQVY--NVSSSWRQRH-SAM-TPMFAETLLPDGTL--GSLQTET[ 2]NTHGTSGG-Q----HLRAT---Q  2660
Query_SPIDE  2557  EYPLSQ-CNFQAY--KIQTSARKRH-MAM-SPLFTETATSSIYS---SLMSLD[ 2]DLTSFRHD-L----KLKAT---Q  2618
Query_SCORP  2583  QNPLST-CAFQNY--AISSSWKKRH-ATL-TPWFVDSLSSAASF---------[ 1]DTSVKNID-E----QIRST---V  2637
Query_TICK_  2480  PRPLSD-CRFVDY--AVTGSWRQRH-AAM-TPLFAETLSWSQSQ--SSTTTSQ    -STTTSATtDp-gPMLRAT---Q  2543
Query_MONOS  2535  EQALKD-CERHPL--RVDVTWQRRN-LHL-TPLFSQSSLDASLEa--TLRQRG    ---------------IRAT---Q  2585
Query_ROSET  2507  ARPLSS-STWRNY--RVDASWRQRN-AHM-TPLFAATQQSQASAgeGEGEDED[ 1]-------------SGVRAT---Q  2562
Query_BELCH  2614  QHPLSE-CRFQDV--QIDHSWKQRH-VAM-TPMFVETQSS---H--PASSQSQ[ 2]SLSQGGQDtL--gGQLRAT---Q  2676
Query_CIONA  2592  DLPLSE-CVFQEY--EIDNSWRHQY-ASL-VPMFVETLGNSSGD--VTSRKQN[ 2]RIL--------------AT---Q  2645
Query_CAECI  2592  EHALSE-CKFQDY--TIDSNWRYHS-TVL-TPMFVETQASLSAS--RYRSQE-    GSFPAQSS-L--gGQVRAT---Q  2653
Query_HUMAN  2578  EHPLSE-CEFQEY--TIDSDWRFRS-TVL-TPMFVETQASQGTL--QTRTQE-    GSLSARWP-V--aGQIRAT---Q  2639
Query_MOUSE  2574  EHPLSE-CEFQEY--TIDPDWRFRS-TVL-TPMFIETQASPSIL--HTQTQE-    GPLSDQRQ-K--pGQVRAT---Q  2635
Query_XENOP  2600  EHPLSE-CKFQDY--TVDSSWRFRS-SVL-TPMFVETQLSQSMQ--RSRAQ--    GTIEADEP-I--gGQLRAT---Q  2660
Query_PLATY  2582  EHPLSE-CKFQEY--TIDSNWRFRS-TVL-TPMFVETQASQSAL--RSRSQE-    GLLPGQGL-L--lGQLRAT----  2642
Query_TASDE  2584  EHPLSD-CKFQEY--SIDSDWRYRS-TVL-TPMFIETQASQNST--RNRSQE-    GSLSLPRL-L--tKQIRAT---Q  2645
Query_CANAR  2581  EHPLSE-CKFQDF--VVDSSWRYRS-TVL-TPMFVETQASPSTH--RNLSQE-    RSLSPSGS-V--rGQVRAT---Q  2642
Query_OSTRI  2551  EHPLSE-CKFQDF--VIDSSWRYRS-TVL-TPLFVETQASQSTN--RSLLQE-    RSPSTSGS-G--gGRVRVT---Q  2612
Query_TURTL  2590  EHPLSE-CRFQDF--TIDSNWRYRS-TVL-TPMFVETQATQSAK--RNRSQE-    GSLSTQGS-L--gGQVRAT---Q  2651
Query_GECKO  2590  EYPLSE-CKFQEF--AVDSSWRFRS-TVL-TPMFVETQASQNPS--RHRPPG-    GSLLTHGS-M--gGQLRAT---Q  2651
Query_DANIO  2588  EFPLSE-CKFQDY--TIDSNWRMRS-TVL-TPMFMETQATQGAE--------A[ 2]SQ---AAT-V--rGQIRAT---Q  2643
Query_SNAKE  2588  EYPLSE-CKFQDF--VIDSNWRYRS-TIL-TPMFVETRASQS-----------    -AITQEGY-L--gGQLRAT---Q  2640
Query_STERL  2602  EYPLSD-CKFQDY--IIDSNWRHRS-TVL-TPMFVETQVTQGGT--QYHTQD-    SP---QGT-I--gGQVRAT---Q  2660
Query_MILII  2589  EYPLSE-CKFQDY--NIDSSWQQRS-TIL-TPMFVETQATQNSS--RSQSQE-    GSLTAQGT-I--gGLVRGT---Q  2650
Query_BAMBO  2584  EYPLSE-CKFQDY--SIDSSWHQRS-RIL-TPLFVETQASAGSI--KRSSQE-    ASTVDQNT-I--gGQIRAT---L  2645
Query_ALLIG  2548  EHPLSE-CRFQDF--IIDSNWRYRS-TIL-TPMFVETQASQSAS--RKRSQE-    GSLSTPGS-Q--gAQVRAT---Q  2609
Query_GAR__  2591  EYPLSE-CKFHDY--IIDSSWHHRN-TVL-TPMFVETQASQGGS--RYQSRGG[ 2]SA---QGS-L--gGQLRAT---Q  2652
Query_TETRA  2625  DEQLSKlAQFYDWdfRDSHNMRLYNsSQPlTPLYTLSQVNYKDSq-GFAIPAP[ 6]----TGVSqMpiqEGVRATv---  2698
Query_PARAM  2510  DKNLSDeAHYYDQefNNKKYIRQNNmSQMaGTQFGSQFTSSSQLl----SQAI[ 3]QNNPGNKNpVnqkQLIRATq-eV  2583
Query_SLEMN  2573  DAPLQD-CKFQPLeiNQFGH--QLNrSQPlTPLLS----QMDQIfqASQMQAK[10]---QNENPdEniaGDVTMIdttV  2648
Query_STENT  2283  DNPLAD-CIYQDAqmNSDDF-----rSFPmTPMFSPSYYHQESIkkASQKSSQ    ----------------------  2329
Query_NEMAT  2579  EQPLSE-CKFEEY--PIDYSWQQRH-VVM-TPLFAATQGS----------TQ-    -SHSQGGSlE--pGALRAT---Q  2632
Query_PISTI  2617  EQPLSE-CRFEEY--QIDYSWQKRH-LAM-TPLFAATQGSHGSG--MTPTQE-    -SMSVDGS----pG-LRAT---Q  2675
Query_CORAL  2590  EQPLSE-CKFQDY--EINFAWQQRN-STM-TPLFVSTLMPSGSS--QS-SFG-    -SHSMDIT-D--gGQLRAT---Q  2649
Query_HYDRA  2486  SMGLDEeKTYSSF--DVKFSWQHQK-ASL-TPLFVSSQHSNTSG--TYSKTQG[ 2]EVVPTKLNqV-----VRAT---Q  2549
Query_LATIM  2590  EHPLSE-CKFQDY--IIDSNWRHRS-TVL-TPMFVETQASQSTA--RFRSQE-    GPEQVQGS-L--gGQVRAT---Q  2651
Query_SHRIM  2545  DQPLTE-CKFREM--KVSTAWRARH-ASM-IPMFAETQASNDT---SLNSLLS[ 2]QGSDTNME-TdgpAGVQAT---Q  2610
Query_MCRAB  2542  QHPLTL-CKFREM--KVSTAWRARH-ASM-IPLFMDTQVSSDS---SLRSL--[ 1]QSNEAEDD-VdgpRGVAAT---Q  2604
Query_AMPHI  2573  EHPLTD-CKFVEM--KVCPAWRARH-ASM-VPLFSNTQSSADSNgpSLSQLVA[ 2]QGGEVGHT-L---AAIAAT---Q  2638
Query_BARNA  2433  DRALPE-SEFEEV--AIVSDWRVQSaAAV--PLFSQPAAGGGPLp--------    ------------ALLRAT---Q  2479
Query_COPEP  2462  SDPLDE-CHFIEY--HVDTNWRMQHsASL-MPMYADTLSSQSQG--SSQTQSS[ 1]-------------GMLRATl---  2516
Query_DAPHN  2352  DQPLDD-VPFQSFd-LLSLSSQTVS----qRPSWSNFGRGFSYG-----TIQP    -------------GYLRAT--qS  2401
Query_STARF  2625  EHPLSE-CRFQDV--TIDHSWKQRH-LAIsTPMFVETQSSQASS--GSQSQSP[ 2]SSLS--------rGEIRAT---Q  2685
Query_STRON  2611  EHPLSA-CKFEGV--QIDHSWRQRH-AAIsTPKFVETQVG-QSQ--GSIRGSP[ 2]QSMTSSGMsMgggGQLRAT---Q  2678
Query_SEACU  2586  EHPLSD-CRFQDI--SIDSSYKQRH-QAInTPLFVETQQG-SGA--GLGSQSP[ 2]TL----------gGQVRAT---Q  2643
Query_FUNGI  2550  DRPLSD-APYREI--NIET---NQH-SSM-TVGLSQSPF--DPR----NVLAP    EFV-------------RATvltD  2598
Query_DPURP  2661  DKPLSE-CSFKEV--QLDSSWQHRT-LNM-NPLFSSSQYGVDGE----E---A    SSNQMDIDeI------RAT----  2714
Query_HALBU  2656  DKPLSE-CTFREY--NLDASWQNRT-SNM-NPLFSSSQ--LDEE----SSQLP    T---FDGDnF----GPRAT----  2709
Query_GLOBO  2434  DRGLPD-ARFNDKdiVVDLLWKN-S-TSM-LPLFAATQD---------KASF-    ---SFDEReG----GIRAT----  2484
Query_PALUS  2494  DSGLPD-ATFADAslDTDVSWLG-T-GSM-QPLFASTQA---------RVQA-    ---RGTQAvG----FIRAT----  2544
Query_CALIF  2566  KNSLPN-AIFGKN-kKFNMSFI--S-SSM-DPLFTSSQN---------NKNY-    --------nN----MIYSS----  2609
Query_GLOMU  2629  DQQLPQ-SRFDENydDIDTSWRV-N-NTM-TPLFINTQT---------SKQK-    --GNLKRQeG----LVRAT----  2680
Query_LOBOS  2559  QEPLPN-AKFNERyaSIDTSWRH-T-VAM-TPLFVNTQQ---------SQSQL    ISDSWTLGeD----ELRAT----  2613
Query_RACEM  2539  DEPLPK-ARFDEHyqNIDTSWQG-S-TSM-TPLFVATQE---------SSKV-    ---DPEDLeL----QLRRT----  2589
Query_PVIOL  2656  EKPLSE-CTFKEV--VLDSSWQNRT-SNM-NPLFSSSQYGLDGE----SSQIS    SSMMMDQDeI------RAT----  2712
Query_BOMBU  2419  E-PLHD-CSYRDY--KIALLWRTKN-FGSkAPLFAPSLASQM-----NQTFTQ[ 1]NTTLSYMSfD------STY---T  2476
Query_AEDES  2553  QHEYEVdVKLREY--SIDTRWRQRN-SFAsAPLFVES------Q---QRLFTG[ 1]NGSQ-MERlI------KAT---Q  2607
Query_ANT__  2454  DAPLQDsCIFEKY--KIPVSWRRRN-LSCiTPMFVDSLASQL-----S--YSS[ 1]SQSVDYNDlY------GTC---L  2511
Query_WHITE  2526  DFPLA-nCNLFEL--PINAAWKRQH-ASF-VPLFADTTQRSS-Q---SSSFSD[ 1]AQ-------V------RAT---Q  2577
Query_BEETL  2398  EHPLE-dCNFEDY--YLETNWRARY-LSV-VPLFAETVRSFS------MSLQT[ 1]DSK------L------RQT---Q  2448
Query_THRIP  2456  SHALD-dCEFGEM--HLLTSWRVKN-AHM-APLYAQTLSSQI-------SLQD[ 1]SSSLDTSYpL------LAT---Q  2511
Query_TERMI  2530  KEPLQ-kCQFEDF--RMLLSWRAQN-ATV-APMFADTLASQLIQ---TQSLSG[ 1]GEHTGAGPlL------RAT---Q  2589
Query_FLEA_  2450  EHPLS-eCTFKSY--SVEDNWHIKN-SAT-VPLFAESFIPKMKS---RSTATR[ 1]TQVY-IGSqT------IAA---Y  2508
Query_MIDGE  2511  DYDSTKdKDFFEY--PIETKMNSKR-SLP--PLFVQS------Q---QKTLIA[ 1]DGSY-YDQmV------RAT---Q  2563
Query_LAODE  2527  PDRNGRsTNQSEI---LDLKQSRVNsASV-APTLVRTERKSNYNsfGSSMGSV[ 3]----------------RST---Q  2582
Query_LOTTI  2548  EHPLSE-CRFQEF--SVKSSWKQRH-AAM-TPLFAATIST---V--SESMEGV[ 1]DSLD---------GGVRAT---Q  2602
Query_APLYS  2607  EHALSE-CTFQDY--SVQSSWRQRH-AVM-TPLFASTLVS---Q--SMDTDES    DQLD---------GGLRAT---Q  2660
Query_SCALL  2606  EHPLSE-CKFQDY--NVKSSWRQRH-VAM-TPLFANTLAT---Q--AMDSDSP[ 2]DSLD---------GQIQAT---A  2661
Query_CRASS  2570  ESPLED-CKFQDY--NVKSSWKQRH-AAM-TPLFANTMAT---Q--SMMDD--[ 1]ESYS---------GEIRAT---Q  2622
Query_OCTOP  2518  EYPLAE-CQFKEY--RVKSHWQHRH-STM-TPLFFDALSSSYME---------[ 2]DIKTENMGaD----ELRAT---Y  2574
Query_POMAC  2570  EHPLSE-CTFRDY--SVQSSWRQRH-AVM-TPLFAVTQGS---E--TMEDV--[ 1]DSLR---------GGVRAT---V  2622
Query_WHIPW  2487  DRPLED-CNFEDM--HVPTHWHSRY-AQFeVPTFMSLSKPFEKIs-SMLTPQS[ 2]ETALTGLSrLe-pPSTSAW---P  2555
Query_TRICH  2503  DHPLED-CTFENL--PISTHWQSRY-ASLeVPAYMQTDMHSE-Vs-KFIS---[ 1]KRKMPYGRrTy-sTESTAW---P  2566
Query_AASTA  2362  PTPLSStCQFEPL--EIDTAWARRTqSIL-TPLFSQDAIALQSQ---QPLDQP    ---------LeraGLVKAT----  2418
Query_AEUTE  2304  SEPLSSsANFEQL--AIDTSWEHRTqNIL-TPLFSQDAQALQSQ---SAIFE-    ------------gGFVKAT----  2356
Query_PLASM  2492  KISLGGg-EYTTT--QIDASWEAKAqS-M-APLFSVEADKFWKR---LQSTQQ    ---------PscrNNTALS----  2546
Query_BREMI  2477  ASSLGSg-EYSDT--QIDASWESKSqS-M-LPLFSVEADALSAH---SESRQQ    -----------lnSSTKYS----  2529
Query_OLIGA  2510  TNPLSDg-EYRDA--EIDAAWETKNeT-M-TPLFSVEADKFS-F---SVATQM    ---------GsltEASIQS----  2563
Query_LIGNO  2736  DRPLAD-CPFREC--RIDADWRRRHgGLM-TPLFLETQQQQQQ--KASDVEP[15]VAMGTDASmLp----LRAT---A  2813
Query_CESTO  2546  LTPLDPhYSFADA--SFSRTSISISqSLLdPVVFSGTQQLRRSQ-----TQLL    ETASASTTlDlqgGDSTQT----  2610
Query_OVI2_  2926  DRPLDPsISFQDLn---LSAYNSVN--LM----------TSDPTi---ASLAP[ 1]-----------------AT---Q  2964
Query_PRIAP  2616  TQPLSD-CKFQDY--NIATSWQSRH-AAL-TPLFTETQAS---Q--ASSTQQT[ 2]AAMQT--------GEIRAT---Q  2672
Query_ROTIF  2549  ENPLSE-CIFREY--NLSADWRRRH-EIM-TPLFAETVTNTFTE--SVDLRHN    ---------------LRAT---Q  2599
Query_SPONG  2608  DTPLSE-CRFEDY-kEIDLSWQQRH-LQM-TPLFAATQLSQTQS--DTGSIGT    -------------GMLRAT---V  2661
Query_TARDI  2619  RTALTE-CDFFKQ--EIATSWKQRS-TMA--PITASTFSSGTFT--SASYGS-    GSGSRNGA---------AP---I  2673
Query_TRICH  2567  DHPLSE-CKFHDY--NVNYSWQYRH-AMM-TPLFAATQLGSQGS--MA--SF-    -SLGSSStD--eHLLRAT---Q  2626
Query_OSTRE  2519  DSDLSD-CVFTLV--KVDTLRRRGAsQTM-NPLFSSESTQLQKEgvDALKQSV[ 3]HSQQLDADpLf-----AAT----  2584
Query_NITEN  2680  DQPLAE-CHFSEY--NIDTAFTGRS-LPM-TPLFSQSQT------------A    TLQ-----eL----GSQAPapsQ  2728
Query_CLUBM  2443  NAPLSD-CEFQER--QIDSSAVGGS-LPM-TPLFSEAGS-----------LR    SMK-----eS----Q-----laE  2487
Query_MOSS_  2566  SAPLSE-CVFQDY--TINTSWVGAS-LPM-TPLFSSSQM-------------    -SM-----dI------QGMhgeD  2610
Query_LIVER  2597  NAPLSD-CEFHEY--TINTSYLGGS-LPM-TPLFSQSQA-----------FP    SSQ-----dS----RSQAVlasQ  2646
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2627  -DAAQFTQTQA-[2]SS--KGPYNWLT-QSSLDTFSDS[3]TLAT-QSES----SLMF[ 1]VGSS------[4]SKK--  2685
Query_LINGU  2628  -EIQQFTPTQG-   AV--KAPYNWLT-QSSLDTFAEY[1]TQSM-ETQS----SLLF[ 1]VGTANENQTP[4]YKK--  2688
Query_LIMUL  2661  -ANLQFTATMD-[2]AK--PKTYNWLT-QSSIDTMEQW[4]AMDASMGSSGS--SL--    LFTIGRKTKA[1]-----  2720
Query_SPIDE  2619  -VNIDFPETLDV[4]KK--RDTYNWLT-EKSLDTFAPQ[2]SGDSGYSSFTN--SPVK[ 2]LFALKNENKS   -----  2682
Query_SCORP  2638  -DVREFDPTQD-[2]SK---KAFNWLT-QTSIDTLGDW[2]------SSTSS--SL--    LFTIGSDTKK[1]-----  2688
Query_TICK_  2544  -AELQFTPTQRG[4]AAttAGSFNWLT-QSTLESNLSG   ---AERRRS---lSILS    ----------   --FAP  2594
Query_MONOS  2586  aSYYSFSQTQSF   -RpgAAPADSLR--SAQNTQTQF   ------TDSQRv--LLV[ 3]------RPRA[3]-----  2636
Query_ROSET  2563  -ASLAFTPTMAS[1]TGtaATAATFEPgGISGQTLTHT   ----QYTAAQQv--LLF    --------ER[4]GANVS  2619
Query_BELCH  2677  -DVLQFTATQ--[3]AP--KNAYNWLT-QSSLDTFADY[2]PLGT-ETQS----SLLF[ 1]TDASKKVA--[2]RRLKP  2738
Query_CIONA  2646  -QDLQFTATVE-   GG--KSSYNWMT-GSSLDTFSSL   ------AVGDHqkSSLL[ 4]--TSPASYRR[4]LVTDK  2707
Query_CAECI  2654  -QQYEFTPTQN-[1]GG--RNSFNWLT-GNSIDTLAEY[3]S-SSESLSS----SLLF    ASKKSEKSRR   TSLKP  2714
Query_HUMAN  2640  -QQHDFTLTQT-[1]DG--RSSFDWLT-GSSTDPLVDH[3]S------SD----SLLF    AHKRSERLQR   APLKS  2695
Query_MOUSE  2636  -QQYDFTPTQA-[1]VE--RSSFDWLT-GSSIDLLADH[3]S------SETLssSLLF    SHKRTEKSQR   MSCKS  2695
Query_XENOP  2661  -QHYQFTPTQN-[1]GG--RSSFNWLT-GSSMDTLADY[2]E-SPESLPS----ALLF    VNKRNENVRR   VPLKP  2720
Query_PLATY  2643  -QQYEFTPTQN-[1]DG--RNSFNWLT-GSSVDTFTDY[3]T-SSESL-S----TLLF    THKKSSRTQK   AGFKP  2702
Query_TASDE  2646  -QQYEFTPTQN-[1]AG--RNSFNWLT-GSSIDTLADY[3]S-LSESLSS----SLLF    VHKRTDKPQR   AAWKS  2706
Query_CANAR  2643  -RQYEFTPTQH-[1]SG--RSSFNWLT-GSSIDTLAEY[3]S-SSESLSS----SMLL    VSKRSEKFKQ   ATFKP  2703
Query_OSTRI  2613  -QQYEFTPTQN-[1]SG--RSSFNWLT-GNSIDTLAEY[3]S-FSESLSS----SMLL    VSKRSEKFKQ   SAFKP  2673
Query_TURTL  2652  -QQYEFTPTQN-[1]NG--RSSFNWLT-GSSIDTLAEY[3]S-SSESLSS----SMLF    VSKRSEKSQR   ATLKP  2712
Query_GECKO  2652  -QQYEFTPTQT-[1]GG--RSSFNWLL-GNSVDTFAEY[3]S-SSESLSS----SMLL    THKHADRSQR   AAFKP  2712
Query_DANIO  2644  -TSLEFSQTLA-[3]GR--RSAYNWLT-GSSVDTLADY[2]--SSDSL-S----SLLV    F----DKKRS[4]AAWRA  2703
Query_SNAKE  2641  -QQFEFTPTQN-[1]SG--RSSFNWLI-GNSVDTFTNS[3]S-SSESV------SMLL    IKKRTDRSYK   TAFKP  2699
Query_STERL  2661  -AQLEFTQTQA-[1]GR--RSSFNWLT-GGSIDTLSEY[3]S-SADSVSS----SLLV    FGKRGDKSQQ   AARKP  2721
Query_MILII  2651  -THYEFTPTQN-[1]DA--R-NFNWLT-GNSLDTFAEY[3]SLASEPGSS----SLLV    FSRWSERLHR   PAKKP  2711
Query_BAMBO  2646  -THYEFTPTQN-[1]GA--R-NFNWLT-GSSLDTFAEY[3]--FSESASS----SLLV    YSKRSERSQK   P-MKP  2703
Query_ALLIG  2610  -QQYEFTPTQN-[1]SG--RKSFNWLT-GSSIDTLTEY[3]S-SSESLSS----SMLF    VNKRSEKSQR   GTLKP  2670
Query_GAR__  2653  -STLEFTPTQD-[1]GR--RASFNWLT-GSSVDTLSEY[3]A-TSDSLSS----SLLV    FGKWGEKQQQ[4]GPKKP  2717
Query_TETRA  2699  -----YSQSQDN[2]GEmvIPSQDQAS---QMNQLEEK[5]DQMKRLVNPSQqiSSSF[18]-NKMSSISEE[4]ENEEG  2785
Query_PARAM  2584  iQQSQYFDEEGF   ------RIPMKSnRSHLSQFLQA[1]EISFQQIQLTSnnQIRY[11]----------   -----  2641
Query_SLEMN  2649  sNSNRNQRRPQI[4]IGs-QSVVEISDtQAQDGNFVRA[4]VFTQTQSDVSSliASNF[20]INKFNKKRPN[4]VNLRI  2746
Query_STENT        ------------   -----------------------   -----------------    ----------   -----
Query_NEMAT  2633  -LRVEFSPTMD-[3]S----QAFDWLN-PSS-QAD--S   V-QSFTVGSSEaqSSLL    FGTKPRASQR[1]--LKP  2690
Query_PISTI  2676  --NLQFTPTQD-[3]S----TAFDWLN-PTQ-ESNLQP   S-QSFAMASTQtqSLLL    FSTSQSHKRR[1]M-LKP  2735
Query_CORAL  2650  --NPQFSMTQT-[3]TGptAQTYDWLN-PSSLDTSQSS   S-ASFSVGSQP--SSLL    FKSNPNMRRF[1]T-LRP  2712
Query_HYDRA  2550  -KKLDFSQTHQ-   I---ESSIDWMS-PTIQITPQEM   ------LPG--------[ 3]--QTNSKNRK[1]-----  2593
Query_LATIM  2652  -VQFDFTPTQN-[1]GI--RSSFNWLT-GSSVDTFAEY[3]S-SSESLAS----SLLA    YGKRSEKSQR   AAMKP  2712
Query_SHRIM  2611  -ANV-FSATQAP   GA----TYNWVT-DSTFDTTLAD[2]YEATHAPTQGStsGLIF[ 1]VGASGQKDRR[3]-----  2671
Query_MCRAB  2605  -ANV-FSATQQT   GG----TFNWVT-ESTFDTSLAD[2]YAATHAPIHGSasGPMF[ 1]VGAHGKKNMN[3]-----  2665
Query_AMPHI  2639  -QRQ-FTPTQAA[4]GG----TFDMGT-DASMQEEVSF[2]LQDKANTVSQGsaGLVF[ 1]LKTN---QAR[3]-----  2700
Query_BARNA  2480  -RSLAFTPTQRS   -AasQPTFDWLT-----QTAAAP   ----AGTGTPSgaRLRY    ----------   -----  2520
Query_COPEP  2517  -ATMEFSQTQAQ   --------------GNLNNKTVF[1]NLSSNGTPMSAr-----[ 1]----------   -----  2550
Query_DAPHN  2402  sSSAAFSSTQSP   -------ANWLM----AETLSTE   -----------------    ----------   -----  2425
Query_STARF  2686  -DAHQFIATQQL[4]SR--KNAYNWLT-GSQ-DTFADY[2]SSAS-DQSS-----LLF[ 1]LGASKKMG-R[4]VSKTA  2751
Query_STRON  2679  -ATQNFTATQD-[4]SH--KNGYNWLT-GSQ-DTFADY[2]SSSA-SESS-----LLF[ 1]MGSLPKRTAQ[4]VKQ--  2742
Query_SEACU  2644  -DGQQFTATQDL[4]DG--KNAYNWLT-GSQ-DTFASY[2]SLGT-ESSS----SLLF[ 1]IGTKQKASVL[4]IKVKP  2711
Query_FUNGI  2599  eESPSMSFTPGS[1]SSlfGASMSKQR--PAVSRKEDD   -KTEASPFIHP------[ 1]QFVKPSLDKS   -----  2653
Query_DPURP  2715  -QAPQFTLTQTA[1]SEfyPS-----------SSQE--   -------SSNA------[ 1]--SQQSSQSS[1]-----  2750
Query_HALBU  2710  -QLPMFSLTQSS[1]SQfyPSGSSSSY--SELSSSQEQ   -SSSATTSTTF------[ 1]VGSAPDSSQS[1]-----  2764
Query_GLOBO  2485  qE-MVWTPTQDV[1]AKttRIQ--FSR--LQSE-----   ----------F------[ 1]SQTQSSTQ--   -----  2520
Query_PALUS  2545  qQ-VQWTPTLDT[1]IGpvQARDIFSG--SESE-----   ----------F------[ 1]SAF-TSTQFE[1]-----  2584
Query_CALIF  2610  qKELSFTPTIEM[1]TQfsYKPSYSQS--TLIVNFEIN   -NLNINTIK--------    SSNKNSISLN[1]-----  2662
Query_GLOMU  2681  nEDLAFSLTLGP[1]NAdlGSQVAGYS--LTQSNLLFS   -QSFIPSLPRS------[ 1]QGGTQK-SVR[1]-----  2735
Query_LOBOS  2614  qSTFEFSMTLDG[1]APaiRSQLGGAA--NPSSTLMFR   -NTPV------------    QGAAPKLSLE[1]-----  2662
Query_RACEM  2590  qQELAFSKTQDG[1]SGdtIASTFQAP--TPES---LS   -NVLLPTSSAA------    QASQQATDSF[1]-----  2641
Query_PVIOL  2713  -QTPQFTLTQTA[1]REfyPS-----------SNSQSV   -DSDVTMSVSS------[ 1]--SSQGSQRS[1]-----  2756
Query_BOMBU  2477  -TSHNSQ-----   ---------------LM--LQAT[1]DLDFEPTYVNDt-SI--    ----AMSI--   -----  2507
Query_AEDES  2608  -FGEEGG---LF   -------------EPTLDPSIIT[1]PATSFTLP-TQh-SLMF[ 1]INPPILDR--   -RSRR  2654
Query_ANT__  2512  -TPTFSYPRLLH   -------------GGLK--LRAT[1]DLQFEPTLDDDn-AADV[ 1]TTFNILSY--   -----  2556
Query_WHITE  2578  --SLEFAPTVAS   -------------DAVSDNFDHP[1]VTSSLNDSGVPs-SFDF[ 1]NTSKFKSR--[3]RKFEK  2631
Query_BEETL  2449  -TSLEFAPTQVV   -------------R---------   -SKQFSQYTMSe-SSLF[ 1]KPEEEKSFVD[4]NLSYK  2495
Query_THRIP  2512  -NSLIFQPTAD-   ---------------FPDGSDAA[1]SATKLKSSLSRs-SFLF[ 1]QDTSQSSK--[3]RGSFK  2563
Query_TERMI  2590  -ASLAFQPTIE-   --------------------QSG[1]SIGSLPSFSLSs-SLLF    --TTGSSF--[3]RKSYK  2633
Query_FLEA_  2509  -FKSNIAKIPMI   -------------QATQQEQL--[1]FAPTFDLQNTQp-----    -----LDV--   -SHDT  2547
Query_MIDGE  2564  -LGNDTM---MF   -------------TPTQDPMILT[1]VSQSFNFQQTQn-SLFF[ 1]LKPQYLDM--   -RSRS  2611
Query_LAODE  2583  -QSLLFEPTVNP   --------------------ADE[4]SVGSSSLHSFAvdATVS    -QNRTSHSKR[2]-----  2628
Query_LOTTI  2603  -DVQ-FTATQD-[1]GS--NAPFNWLT-GSSLDTYGDY[1]SMSSSETPS----SLLF[ 1]VGTSGSQSTN[1]--R-Q  2660
Query_APLYS  2661  -DVQQFTATMD-[2]GA--GRVFNWLT-QSSLDTFTDT[1]VIGG---DS----QLLF[13]KGSSLKGRPS[4]MAAVR  2735
Query_SCALL  2662  -DVQQFTATQD-   GS--RAPFNWLT-QSSLDTFTDY[1]TVGS-EPQS----ALLF[ 1]VGSMEASARI[3]RMR-P  2722
Query_CRASS  2623  -DAQQFTATLD-   AS--KAPFNWLT-GSSLDTFTDY[1]TVGS-ETSS----SLLF[ 1]VGTQDMALRT[4]KRK-P  2684
Query_OCTOP  2575  -GNQQFTPTQDV   ------PFDWLT-QSSLSSSMEV[4]SGIQSSLLTVS--GKAT[ 2]SQVYPTRKRP[1]-----  2633
Query_POMAC  2623  -DVEQFTATME-[1]GS--NAPFNWLT-QSSLDTFADA[1]LMET-NSLS----SLLF[ 1]VGSSSKGADS[4]ATR-H  2685
Query_WHIPW  2556  ----LLTPTAAE[4]GSp--------N--QSLETFGSA[1]----DESSGVLemDAAL[ 4]----------   -----  2598
Query_TRICH  2567  ----GFTPTAIT[4]NVg--------N--ISADTFELD[1]----DESDSSDnlTVIL[ 4]----------   -----  2609
Query_AASTA  2419  ---------QDR   --------VW-------------   ----SQTQSQQfqSSQG[ 1]ATTGSIQ---   TTSAG  2449
Query_AEUTE  2357  ---------QDP   --------IW-------------   ----SQTQAQHftQSQV[ 1]SSAPNSS---   SHRRD  2387
Query_PLASM  2547  ---------E--   ------------sDKSLR-----   ----DTMQSQLifSPQA[ 1]VRLKSSY---[3]IVDPT  2582
Query_BREMI  2530  ---------A--   ------------ySQPLE-----   ----SKMQSQLysNALT[ 1]RGSRLSY---[3]SGETN  2565
Query_OLIGA  2564  ---------QGQ   --------QWGSsTQSLDGLFDG[1]----SASSSQLisSSQA[ 1]SGSSSSQFAA[4]DGSQP  2615
Query_LIGNO  2814  vGPGMFEATQLH[4]TA------GLLA--DSQDLSVST[3]FRSSGGGGGGA--GLEV[12]LGSDANPQMA[3]AQLHR  2892
Query_CESTO  2611  ----QLLTTQQI   --------------PSNSTSTQF   ---DETLSDQSapRHVF[ 3]------KPYK[4]IDFKN  2657
Query_OVI2_  2965  ---TAFTTTTRA   ----DKTFAFTRiVTQSKTLTAG[3]IEASQQMTQPVnkKART[13]VSQAPQTPSE[4]ATPTP  3044
Query_PRIAP  2673  -VVPEFTATQD-[2]GG--RSTYNWLTqSSSIDTFADQ[2]SSVS-SISS----SSSS[ 1]PYSSPSAD--   ---HH  2730
Query_ROTIF  2600  -QTLQFQPTLKV   -------YDWLT-QSSIDTFTDT   --ALIDTKS----SLLF    ----------   -----  2636
Query_SPONG  2662  qQSLAFTPTQG-   ---qESSYNWMT---PSAQSVDY[3]SFTVGATPTSD-sSLLF[ 1]----------[1]PTMRP  2715
Query_TARDI  2674  -----FTPTQA-   ----------LA-GPANDTLAVA   ---GFTYSSQFaeSLLI    ----------   -----  2705
Query_TRICH  2627  sTNLQFTPTQD-[3]R-------NWLD-PSSVDKSIQS[5]Q-QSLSSTSSEtsSTLL    FLSSKSPRQR[1]LQMKP  2692
Query_OSTRE  2585  ---LSYKPVKRL   ---pKSAYRAKD---VSSTLHKT   -----------sfSAVI    ----------   -----  2616
Query_NITEN  2729  tQTGGRTQQQPL[1]PGmiRATQGPS-----QSLSQGG   -TFAPSLSQGD------[ 1]VSSSLMATQS[1]-----  2781
Query_CLUBM  2488  -ANASLAIVNHG[1]GGliRATASYT-----PSMTILE   -GSDEPGQTQN------[ 1]ADGSQN----   -----  2534
Query_MOSS_  2611  -RDGEFDGATHH[1]PMmiRATMTPS-----LSQAVTQ   -TQGGTLSSTQ------[ 1]VNGSVL----   -----  2657
Query_LIVER  2647  -NSKEDDEMIPR[1]AGmiRATYTAT-----LSQTQDT   -LSSYPLASQE------    ---SSL----   -----  2689
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2686  -TPG[2]FGQRKVRSDV-rVSAM------   -DEAN-KEIMRLKRRFIHD-QE-S--TAR-FFMKRNVRMKKMREDIQR  2746
Query_LINGU  2689  --TG[2]FGKGRLKMEG-tDFTD---SRE   DDSEK-EEIQNLKRRFLKD-QR-A--RST-YFAKKNIRLQAMRQEALK  2752
Query_LIMUL  2721  --AS   SPKTKGGRSLkaVSD--EVDST   SDSSSvSDILRLKRRFVKD-QE-K--HRL-VFVKRAARLKNMRDDLLK  2785
Query_SPIDE  2683  ----   ---TKTKFQIdlDEN--EKNKE   KKSTT----WMLKKRFLKD-KA-N--SEV--FARQELRRQDRRKEILK  2737
Query_SCORP  2689  --PN   IDENNRNSNQdtSKN--KIEKL   SSSNT----FHLRRRFLKD-RE-K--SRL-HFIRMEMRKQQLRQEREK  2749
Query_TICK_  2595  VQEE[2]-----LGASQ------------   -TSQKsKELGFLRRRFLRD-ED-K--VRL-QHIKREVRRKELRGEQER  2647
Query_MONOS  2637  AGTQ[2]FSQSEQDHEL--AAAMQAREKA[1]--ERR--EIGRLRRRFLQTkEGaQt---A-FWARQATRQRLERQARAK  2703
Query_ROSET  2620  VDPS[1]-------ATP---QRAAERQAR   --EAR--EIQRLRRRFVKS-KDkRaqAAI-FHAKRMAARKQEARRLIE  2678
Query_BELCH  2739  VGPG   FGSDRLSA----GRDQTDGRAA   EDSQQ-SEIMRLKRRFLKD-QD-V--TKA-YHARRQIRMQKMREERVK  2802
Query_CIONA  2708  PGPE   FGKVDLV-----KQEGQKKKQR   GFGVN-IDVQRLRRRFLKE-QT-D--SRI-NFARVQEVKKRRQETFQK  2770
Query_CAECI  2715  LGPN   FGKKRLG-LP--GDEVDSKAK-   GIDDR-SDILRLRRRFLKD-QG-Q--ISK-IFVRKGVAEQKREKEITA  2778
Query_HUMAN  2696  VGPD   FGKKRLG-LP--GDEVDNKVK-   GAAGR-TDLLRLRRRFMRD-QE-K--LSL-MYARKGVAEQKREKEIKS  2759
Query_MOUSE  2696  VGPD   FGTKKLG-LP--DDEVDNQVKS   GTPSQ-ADILRLRRRFLKD-RE-K--LSL-LYAKRGLMEQKLEKDIKS  2760
Query_XENOP  2721  LGPN   FGMRRLG-LP--GDVTDSKTK-   SMEDR-SDILRLRRRFLKD-RE-K--LSL-IYARKGTAEQKREKAIKI  2784
Query_PLATY  2703  LGPN   FGKKRVG-LP--GDEVDSKTK-   GIEER-TEILRLRRRFLKD-QE-K--LSL-IFAKKGVVEQKREKEIKS  2766
Query_TASDE  2707  VGPN   FGKKRLG-LP--GDEVDNKTK-   GIDER-AEILRLRRRFLKD-QE-K--LSL-IYARKGVVEQKREKEMKN  2770
Query_CANAR  2704  VGPD   FGKKRLG-LP--GDEVDSKTK-   GIEER-AEILRLRRRFLKD-QD-K--LSL-IYARKSLAEQKREKEIKS  2767
Query_OSTRI  2674  VGPD   FGKKRLG-LP--GDEVDSKIK-   GIDER-AEILKLRRRFLKD-QE-K--LSL-IYARKGVAEQKREKEMKS  2737
Query_TURTL  2713  LGPN   FGKKRLG-LP--GDEVDSKAK-   GIDER-AEILRLRRRFLKD-QE-K--LSL-IYARKGVAEQKREKEIKS  2776
Query_GECKO  2713  LGPN   FGEKRLG-LP--GDEVDSKTK-   GIDER-AEILRLRRRFLKD-QE-K--LSL-IYARKGVAELKQEKELKT  2776
Query_DANIO  2704  VGAG   FGSKRLT-AT--SDDTDSRTAA   ERERR-ADILRLRRRFLKD-KE-K--ESI-KFAKKEIHSQRTERERRA  2768
Query_SNAKE  2700  LSPN   FGKKRQG-LP--GDEVDNKSK-   DTDEH-TEILRLRRRFLKD-KE-K--LSL-IYARKGVAEQKQEKEIKT  2763
Query_STERL  2722  VGSD   FGKKRLG-LP--GDEVDSRTK-   VKDQR-ADILRLRRRFLKD-QE-K--VSL-IYAKKGVREQKQAKENVA  2785
Query_MILII  2712  VGPD   FGKKRLA-LP--GDEVDSKTKG   IVDQR-ADILRLRRRFLKD-RE-K--VSL-IFAKRGIREQRAEQKIKT  2776
Query_BAMBO  2704  VGPD   FGQKRLA-IP--GDEVDAKTKA   VADQR-ADILRLRRRFLKD-RQ-K--ENL-IFARKSIQEKHLAKEMKT  2768
Query_ALLIG  2671  VGPN   FGKKKLG-LP--GDEVDSTTK-   DINER-AEILRLRRRFLKD-QE-K--LSV-IYARKGVAEQKREKEIKT  2734
Query_GAR__  2718  VGAD   FGKKRLA-LP--GDEVDSRKKA   VNDQR-ADILRLRRRFLKD-QE-K--VSL-QFARKGIQAQRQEKERAA  2782
Query_TETRA  2786  EEKK[3]FGTDIQREKK--NNNLSKFFAS[6]QVNKN--STNKIRYQDSKEgKTlTntVDI----------KKRQQQLFF  2854
Query_PARAM  2642  ----   ----------------------   ------------TKKYYKL------------------KSQQPSSQMYT  2659
Query_SLEMN  2747  GAQS[4]VADDDDGFIApvAKRKGKADLA[8]VYRPQeSQGQGQRIRYVPQgSMpAhlQKQ-------EEIRKMQREQYL  2825
Query_STENT  2330  ----   ----------------------[3]---PQrTS----RIRHLVEgSQaSieDRR-------EKIKKAQQERLE  2366
Query_NEMAT  2691  VPHD   FGSRRLAPTQ--SDSQQATPAM   G-SQQ-QEILRLKRRFTKDsQT-K--DRQ-YFIRMHEKKRQMLQQTQS  2756
Query_PISTI  2736  VPGD   FGSQKLSQAS--SSRISSHQSQ   DPNEK-KEILKLKRRFLKD-QS-Q--ISA-FFAKQALKRNKFREQALQ  2801
Query_CORAL  2713  Q---   --SEQPSTSS--AAQSDDPSNE   NSSQL-QNIRKLKRRFVKD-RK-E--TDV-YFAKKEIARQKLRTESVQ  2773
Query_HYDRA  2594  ----   ------------KNEFSKPSSK   DF----------RFRHFKN-KS-E--ESY-YFASLEKKRQKLRESALE  2636
Query_LATIM  2713  VGPD   FGKRRLA-LP--GDEVDSRTK-   VADQR-LDILRLRRRFLKD-QE-K--VSL-IYARKGVQIQKEEQKMKA  2776
Query_SHRIM  2672  -RPG[1]-GNSRMRP----ADEPDGNDAA   DTMQS----KLIRRRFVKN-QDrE-kQSI-YFAKIEERRSKLREKLEK  2733
Query_MCRAB  2666  -KSS[1]-GNTCMQP----AETSTTTDSG   N-IQS----KLLRKRFYKD-QDrE-kQTI-FFAKQEERRNKMRATLEM  2726
Query_AMPHI  2701  -KTG[1]-GKTRLKPKT-fGDENDGDETP   SAIVS----SLVRRRFTRQ-QDkE-qQYR-YFAQAEEKRRALKHRLER  2765
Query_BARNA  2521  -TDD[1]-----------dGGDGDAGADA[1]---SR-----MLRRRVETDaRQvSa--RM---ARRHADQRALREDRQR  2571
Query_COPEP  2551  -GPG[1]----RRSTTP---MKVGSLVAG   -------------RRFTQEqKTaGyqKST--FSKIAERQNVRKRQLEE  2602
Query_DAPHN  2426  --PS[1]TSKAAMKKTT---ETVEKETFK[2]---------LKSWKRFTTTgPPpAk-PSY---------ENRGRSANKT  2478
Query_STARF  2752  TDAG   FGKEKLQSDQprGRDEVDAKQS   -GED--KKINRLKRRFLKD-HS-S--SQA-YFARREMKRKQLKEKILE  2817
Query_STRON  2743  -DAG   AG----------GSRDKEARKA   PSEGE-KEVQRLKRRFMKD-QS-S--QSA-FFAKRATRQKELREEVQK  2799
Query_SEACU  2712  VGPG   FGKDRLSV---qGKDNTDTGAS   KAEG--KEIMRLKRRFLKD-QQkS--SSL-YFAKRELKKQQMREELLK  2776
Query_FUNGI  2654  ----   ----------------------   ------------RKRFFKMgTPnRsqQ-------MEMIRLKNKQQQIQ  2682
Query_DPURP  2751  --SN[1]-SQQQR--------------VL   YRNPKsFENNELRKRFKKIdDKgSdeRVV-KFARMQVKKNIDREAFFE  2807
Query_HALBU  2765  --KA[1]-NKRTVNNKP-eDKAGTFKVPA   NIDPR---VMEMRNRFKKKdNRtDgeRNA-AFARLAVKKSKDRDELMS  2831
Query_GLOBO  2521  --PF[1]-DRR---------------TL-   -------------VRRSYAsKVdSt-Y----FAYGAERKKKEARNIET  2558
Query_PALUS  2585  --PQ[1]-DRARFT------------SY-   -------------TRRQVQgTYdHr-F----YAAQQERQKRARAKNEQ  2625
Query_CALIF  2663  --DN[1]-DKTISN------------SQK   DENRLlVPQRRRALRYKKTsNArKi-F----FANAAELKKKRIQKHLK  2717
Query_GLOMU  2736  --SE[1]-QRFGRR------------SFK   PYSQL-------DFKQMMAtQRkQdeI----YRTSQHAFPKN------  2778
Query_LOBOS  2663  --SL[1]-DRSQK----------------   -YRRL-------QMRHVHAtEVeQsqR----YRRAYEAKVQNKAQAIA  2706
Query_RACEM  2642  --KS[1]-GRTREK------------TNR   AYDRI-------GKRFIKTsDRdStyF----FRNRHERQRKTLQSLEA  2690
Query_PVIOL  2757  --SN[1]-KKKQLN------------VVA   SGDLRsNTYNQLRKRFKKVdERvQdgRIA-GFARLQVMRNKERESFLE  2815
Query_BOMBU  2508  ---N[2]EDNTF----E-vPKVPQPAY--   --NKK-------SKRFLSN--AsDi-SAA-IRQKEIKRNIQHAEMIKD  2560
Query_AEDES  2655  -TPQ[2]TPTT--------------RSAQ   SYDSL-------RKRILKD--KdRs-ARD-QALKAIDRYTYETVRKAE  2704
Query_ANT__  2557  ---D[2]FDRTT----I-qSTSRQSLQVA   PRQKR-------FTRILAN--TsDv-ANI-YRNEQIRKNEQRTERIKQ  2613
Query_WHITE  2632  --PH[2]TDIANVGHVI-sSQHKSYRLKK   TQS--------------KE--DdSk--YA-L---QNIKLNKAIEIAQR  2682
Query_BEETL  2496  Y---   ---------------------R   IPK----------RRFLKD--KaKv--HS-YFAQQQVKKSSEKIEKRK  2530
Query_THRIP  2564  QGAN[2]SDLLQLSEPT-eHEKSPKRTER   AS------------RYFGR--SkAv--SV-F-SKANVKRAKRREEMWR  2620
Query_TERMI  2634  MGPG[2]SSKLQ-PIPE-vSDEVDSRSDR   LAPNL-------HRRFLRR--DdQk--RI-FHASREIQRSARQQELEQ  2695
Query_FLEA_  2548  -LNY[2]FDADN-EEIK-vEDLGSQQLKD   SFLKL-------RRRLLKD--KdKa-SQQ-FAARAIAKK--KKDRRIL  2607
Query_MIDGE  2612  FTPN[2]ADDDI-EARI-aRKKEGKRHDE   AFDYL-------RHRILRQ--HeEnrSKD-YAMKAIDRRNFQEARSRE  2675
Query_LAODE  2629  --PG[1]-------DLDdnDDDGDFKVPV   ---HK-------ARRLLSD--KkKigQSA------AFMNQRRNAIMQT  2676
Query_LOTTI  2661  SKAG   FGLRRQTIVK-qLGHA-PKSEG   MLNKS-DTVWKLRRRFQKD-KD----DRL-FFMKKNVRLRKREEERLM  2725
Query_APLYS  2736  TPAG[3]-GPGDAAAVT-gGNQD--VADG   GAGSQ-TDIWRLKRRFLKD-KD-A--QSV-HFMRRNIRLRQMREEAAR  2802
Query_SCALL  2723  KPSG   FGQPRPGPSK-vKTEE-KAGPS   ESDEG-NEMFRLKRRFLRD-EE-A--SRI-FFARRTVRVQQMRQVALQ  2788
Query_CRASS  2685  GP-G   FGRPRAGASS-kPAER-S---R   NEDQD-SEMYRLKRRFLRD-EE-Q--TKT-YFIKRHTRLEKMREEALR  2746
Query_OCTOP  2634  --PG   FGQKKLSAIKdsQEDPLEENSQ   KIENKsRKFAYLKKRVLKD-RE-A--QSL-YFAKKQITKNETQVQLVH  2700
Query_POMAC  2686  HHTG[4]FGKQMARTER-tGNSAGPKNDN   NTPDE-DNVWKLRRRFLKD-RS-T--ASK-YFARREIKEKSKREERLK  2756
Query_WHIPW  2599  --AG[1]-------RLP-----QHGRAIP   DSSIG-RDFKLLKRRILKDtEQkQ------FYANLAKKKQQMLQRMAT  2652
Query_TRICH  2610  --QD[1]-------KKP-----KSSRVKK   VDDLD-GDLQFLKNRILKDtEEkA------FFVNLNRKRQKVIERL--  2661
Query_AASTA  2450  LAPR[3]FYK------S-tSNARIS-DES   SRTKT----------FFQE---------------RASRSKRQELARLN  2493
Query_AEUTE  2388  ITAK[1]FTK------H-rGN-EPG-NQE   NRTKS----------YFQD---------------QASRSKKLEIARIK  2428
Query_PLASM  2583  FDLR[3]FSTRRNFADS-sAVEDKSINDK   KLSKR----------FYQD---------------KYAMLKKHDLAKMT  2633
Query_BREMI  2566  VAPG[3]FLKQRAAVDH-tTDKVPS-NKD   KVSKL----------YFQD---------------QHGFLKKREEAQII  2615
Query_OLIGA  2616  KGKR[3]FVKKQSTVFS-sGEARGA---K   GADKR----------YFQE---------------RYALFKKTEEAILN  2663
Query_LIGNO  2893  LRRR[1]FLQQQQQQQA-aSSSSGAAASS   SSDGP------------PAtQGsSssYSA--FAQAGVERHRRQAREAA  2952
Query_CESTO  2658  LSPV[3]--RRKFSFAT---------TLA   ---AQnPDLVLRSENPIRE-----------IFKNRTVQRQHAEISSVN  2709
Query_OVI2_  3045  YELR[1]---QRLRNQLnwRESTNANHTG[1]---------------FTKE--HgKkeRNI--FRERAIARQRAEFRLFQ  3098
Query_PRIAP  2731  ETPT   EAKRRGAT----GEEPREEASA   AGGG--DDVMRLRRRFLKD-QE-Q--TRV-FFARREARRKQRREEAET  2793
Query_ROTIF  2637  ----   --------------KTESKNK-   -DSDE-KEVLKLRRRFIKDkAE-S--QSK-FFARKQIAKKNRDEDLRK  2685
Query_SPONG  2716  PLKN[3]ISAERLG--K-vGESEGGETK-   -EEKR-KQVLQLKKRALKD-KE-M--VSS-FFAMNEARKKMRREELKK  2781
Query_TARDI  2706  ---D[1]---SRRGLIT----RDEHTTEI   SRLDK--EMDTVKRRFQRDsNSdS--SNKsFFIKREYSRRRAIEETLK  2766
Query_TRICH  2693  LSKD   FGQQKLSSET--ASQPQSQSVS   QPSS---SLNHLRRRYIKS-EM-A--NNA-FYAKRERRKKQLREEILK  2756
Query_OSTRE  2617  ----   ---------P-----------K   ----------------FQDaQD------V----------QKTLGDVLL  2634
Query_NITEN  2782  --GF[1]-GGFSQVGLT-qGVTQGLGGAR   RGGALwDTQTNVQRRILKPrT--SqvLQG-VL-----GARRRREARAS  2844
Query_CLUBM  2535  ----   ---VAVV---------------   ------DPVTRLNRRLYHVkESrAgtLRN-VF-----NATRKRKALQS  2574
Query_MOSS_  2658  ----   ---YGLSQRP-gGAGYIKKTGL   EEASWmDPALHLRRRILNPrATqSqvLTG-VF-----RASRRREARLK  2717
Query_LIVER  2690  ----   ---FAMS-RK-sSVTHTQVTSQ   ETGSStDSLPGLRRRIPLAnAArSqvIRS-VL-----NANRRREARAA  2748
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2747   E   QKMKREH-QVTMYRQYRTGDFPDIQ-IKYSFLIAPLQALAQRDSTI-AKLLFVAVFQAIFTQI--ESEKTERES      2816
Query_LINGU  2753   E   QKLRREA-QVTMYRKYRTGDLPDIQ-IKYSYVIAPLQALAQRDNTL-AKMLFSALFRAIFSEL--DQVKTERET      2822
Query_LIMUL  2786   E   QRSRREA-QVTMYRRYRIGDFPDIE-INYSSIIAPLQALAQKDQNI-AKTLLTSVAKAL-CQS--EEENHPHDY      2854
Query_SPIDE  2738   D   QRSHRDA-QVTMYRQYRIGELPDIE-ISNSAIIQPVQALAMNDVSI-AKLLLKSLLSAI-CTS--VKEDFPEGF      2806
Query_SCORP  2750   E   QKKQREG-QVTIYRQYRIGDLPDIE-INYSSIIVPLQALAQHDDEI-CRMLFQSLILSV-CCY--MKSNQVTEY      2818
Query_TICK_  2648   E   RRARREA-EVTLYRSYRLGDLPDIE-IPHSALVAPLQALAQHDQQV-AQVLLTSIVKTV--------VKDGADL      2711
Query_MONOS  2704   E   RRQARHA-NVVLYRDYKDGELPDIQ-IKQAELIRPLVALTQRDTTF-ARSALVELVTGLLNAFa--SQAAEADS[ 1]  2774
Query_ROSET  2679   L   RRKARAA-QVTLFRSYRTGELPDIQ-IKQRELLQPLQGLAQRDPQL-ARMVFESLFYGVLHELp--DILSERDA[ 1]  2749
Query_BELCH  2803   E   QRAKRGN-LVTLYRQYRVGDLPDIQ-IKYSYIIAPLQALAQRDSML-ARLLFSSLFRGIFEEI--ENVKTEREA      2872
Query_CIONA  2771   E   HEKRRSN-DVTLHRRYRDGELPDVQ-IPNKSIIAPLASACQYDAVL-SKSVLSSLVCAILKEE--RKRDGE---      2837
Query_CAECIL 2779   E   RKMKRDS-EVTLYRSYRHGDLPDIQ-IAYSNLIAPLQAVAQRDPTL-AKQLFSSLFFGIMKEM--DKLKSSSEM      2848
Query_HUMAN  2760   E   LKMKQDA-QVVLYRSYRHGDLPDIQ-IKHSSLITPLQAVAQRDPII-AKQLFSSLFSGILKEM--DKFKTLSEK      2829
Query_MOUSE  2761   E   FKMKQDA-QVVLYRSYRHGDLPDIQ-IQHSGLITPLQAVAQKDPII-AKQLFSSLFSGILKEM--NKFKTTSEK      2830
Query_XENOP  2785   E   QKMKQDA-QITLYRNYRQGELPDIQ-ISYSNLIAPLQALAQRDPTM-AKLLFSSLFSGILTD-------TASDI      2849
Query_PLATY  2767   E   LKMKHDA-QITLYRSYRQGDLPDIQ-IEHSSLITPLQAVAQRDPTL-AKQLFSSLFAGILKEI--DKFKSQSEK      2836
Query_TASDE  2771   E   LKMKHDA-QVTLYRSYRQGDLPDIQ-IEYSSLINPLQAVAQRDPVL-AKQLFGSLFAGILKEM--DKFKSPSEK      2840
Query_CANAR  2768   E   LKMKYDA-QVTLYRSYRVGDLPDIQ-IEYCSLIAPLQGLAQKDPTF-AKQLFSSLFSGIFHEV--KKSKHPSEK      2837
Query_OSTRI  2738   E   LKMKHDG-QVILYRSYRVGDLPDIQ-IEYSSLIAPLQGLAQKDATF-AKQLFSSLFTGILREV--EKSKTPSEK      2807
Query_TURTL  2777   E   LKMKHDA-QVTLYRSYRQGDLPDIQ-IEYSSLITPLQGVAQRDPPL-AKQLFSSLFTGILKEM--NKFKNPSEK      2846
Query_GECKO  2777   E   LKMKHDA-QVTLYRRYRQGDLPDIQ-IEYSSLITPLQGLAQRDSTL-AKVLFSSLFAGILEEM--RKSKSQSEK      2846
Query_DANIO  2769   D   LKIRQDA-QVTLYRSYRVGDLPDIQ-IQFSSLIAPLQALAQRDATL-AKQLFSSLFAGVLVEM--ERLKSNKET      2838
Query_SNAKE  2764   Q   LKIKHDA-QITLYRKYRQGDLPDIQ-IEYSSLITPLQAVAQKDPTL-AKQLFSSLFSGILEQM--KKSKPLAER      2833
Query_STERL  2786   E   QKLKREA-QVTLYRSYRQGDLPDIQ-IQHSSLIAPLQALAQRDPTL-AKQLFSSLFSGILIEM--EKSKSQSEK      2855
Query_MILII  2777   E   KRMERDA-KVTLYRSYRFGDLPDIQ-IKYSSLIAPLQALAQKDHSL-AKQLFSALFSGILEEV--HKSKT-AET      2845
Query_BAMBO  2769   E   QLKKRDA-QVTLYRCYRVGDLPDIQ-IKFSSLIAPLQAVAQKDHSL-AKQLFSSLFSGILEEV--RKSKTAAEM      2838
Query_ALLIG  2735   E   LTMKRDA-QVTLYRSYRQGDLPDIQ-IEYSSLITPLQGVAQRDPTL-AKQLFSSLFTGILKEM--ENSKSPLEK      2804
Query_GAR__  2783   E   QTLRREA-QVTLYRRYRHGDLPDIQ-IPYSSLIVPLQALAQRDSTL-AKQLLSTLFCGILAQM--EGSKNEAEK      2852
Query_TETRA  2855   E[1]QNRMRKQ-QLQTIRKYKIGELPDIG-ILYKDIIDPLMALSFIDETL-AGQLFQLLLHSMFMKEe------SADR[ 1]  2922
Query_PARAM  2660   S[5]EKFVQNR-----VRKYREGDLPDIK-IKHKDIIDPLIVLCFEDEEL-AGRVFTALFQELYDRN------TSAET      2726
Query_SLEMN  2826   Q   RQIQRQKhQVGLYRRYRVGELPDIQ-ISYEDVLKPLQAITKSDLTI-ATEVFVEIFSELYKSQ------NDQQQ[ 1]  2893
Query_STENT  2367   Q   RINEQKNrQVNVTRSYRIGELPDIE-IRTSDILKPLANLILLDCEI-ASSIWVTICTALFTQF------DDDAK[ 1]  2434
Query_NEMAT  2757   R   LKAARES-QVVMYRRYRAGDLPDVQ-IKHSELIAPLQAVAQCDSTL-AKQLFSVLVQAILSQM--DEKLTEQDA      2826
Query_PISTI  2802   R   QKAAREG-AIVMYRKYRSGDLPDIQ-IKYSELIAPLQALAQCDNTF-AKQLFSALFRGIFSKI--DEKLPEREA      2871
Query_CORAL  2774   R   QKVAREN-KVVMYRKYREGDLPDIQ-IKHSDLITPLQALAQCDSYI-AKQLFSGIFRGIFSQI--DENVTDTQA      2843
Query_HYDRA  2637   K   RKIARDQ-KVVMYRKYRIGELPDIM-IKHSDLIVPLQALAQLDHKI-GQQLFVSIFLAILNQL--SQ--SS---      2701
Query_LATIM  2777   E   LKMKQDA-QVTLYRRYRQGDLPDIQ-IHYSSLIAPLQALAQRDPTL-AKQLFSNLFTGILTEM--EKSRSDSER      2846
Query_SHRIM  2734   E   RKFRREA-QVTMYRQYRVGELPDVQ-IPHRALIGPLQALSQRDEKV-ARLLFEVLAQGVINSA--GTAMHHREE      2803
Query_MCRAB  2727   E   RKLRREA-QVTMYRQYRVGELPDIQ-IPHRALIAPLQALSQRDNKV-AQLLFDVVVQAVVDNA--DSLMTRGEK      2796
Query_AMPHI  2766   E   RTLRREA-QVTMLREYRVGELPDIQ-IQHRAVLDPLQALSERDSSV-ARHVFVLLCEGLFRHG--EERLSETQL      2835
Query_BARNA  2572   D   EPRRQDA-QVELFRRYRRGELPDTE-IPYSDLLRPLQALAMHDPPT-AARLLAGLLVGTLHELdaAEQRPSSPT[12]  2655
Query_COPEP  2603   D[1]LKASEKN--VTLLRRYRKGDLPDIQ-IPYSDLIAPILKLCKLSPEL-SKVLFSGVIKGLLDVNa--ELLNDEGL      2672
Query_DAPHN  2479   S   ------S-QPTIYRKYRTGELPDIQ-IPPSSLIAPLCAVARNDAII-AKTLFCHLLTVLRQQVr--EIQGDDKE      2542
Query_STARF  2818   E   QRERKFS-TVTMYRRYRTGDLPDIQ-IKYSDLIAPLQALAQRDASL-ARLLFSSLFRGIFSEI--EEVKTEREA      2887
Query_STRON  2800   S   QRERRFA-QVTMYRQYRTGDLPDIQ-IKHSDIIAPLQALALRDSTL-SRLLFSALFRGILARI--GDMCTDGET      2869
Query_SEACU  2777   E   QKERRYH-QVTLYRSYRVGDLPDIQ-IKYMDLIAPLQAVAQKDAAL-ARCC-SSIFGGIFVQI--GEKKTEREA      2845
Query_FUNGI  2683   A   ARQARKK-NLVILRKYRAGELPDIQ-IKSRDILDPLMFLVKKDSNT-AQSIMAHIFAILSQDA--KQDMDLQY-      2751
Query_DPURP  2808   K   SKQAREN-QITMIRKYRTGELPDIQ-IKLQDIIKPLQILCQRDSNI-GVNIFSLLFSSIYNKT--QQENI----      2873
Query_HALBU  2832   R   AKQAREN-QVTMFRKYRSGELPDIQ-IKLQEILRPLQSLCQMDHHV-GTSVFSLLFTTIYQHC--PRDQVNKF-      2900
Query_GLOBO  2559   F   KQAASQR-SVTLYRKYREGELPDIE-IKHKDVLEPLQSLANRDNEI-ARRLMTIILTKMT--L--DNTIDLEK-      2625
Query_PALUS  2626   L   QKIARAR-TVTLARKYRIGELPDIE-IQHSEIINPLQALAMRDSEV-GQQVFSSLVASVLGHV--DDYKGMDEE[ 2]  2697
Query_CALIF  2718   L   GKQAKDK-LVYTLRNYKEGELPDIQ-IKYKEIIQPLQALAIADTDI-SKNLFIMILESIWSKK--DEIMNDKE-[ 1]  2787
Query_GLOMU  2779   -   ---------ISMLRQYRVGELPDIG-INHGDLINPLQTLAQRDLDI-SRTLFTSLFVSIYNQL--DNNIQSEE-[ 2]  2840
Query_LOBOS  2707   N   KEIARAK-SVSIIRKYREGDLPDIQ-IAYSDVIRPLKSLAEMDVEI-CRMLFSKLVISLIGQI--DSHVESEEE[ 2]  2778
Query_RACEM  2691   M   RKQQREK-AVTMFRNYRIGELPDIQ-IKNSDLIIPLQALGGRDNEI-ARILHAHIVVSFSQYA--DRYLQDKR-[ 1]  2760
Query_PVIOL  2816   K   SKQAREN-KITMIRKYRSGELPDIQ-IKSQDVLKPLLILCQKDSNI-GVYIFSSLFTTIYKNS--KQQQQ----      2881
Query_BOMBU  2561   E   GVRQRSS--VKLYRKYRIGDFPDIE-ISHATLIEPLQQLAKNDQLI-CKDLIVNIVCSLI--------ENCRQS      2623
Query_AEDES  2705   S   HKM-RQG-EVVLYRRYRIGDFPDLL-INSLALLMPLQALCRRDPPL-ARQVFVAIFSGI---LeeWDQNEVDSR      2772
Query_ANT__  2614   E   NARQRSS--VKLYRDYRIGDFPDIE-ISNASLIVPLQQLIKLDRLI-CKDVTVALFYSLIK----ETIEREKSD      2680
Query_WHITE  2683   E   EVEKRSK-EVTVLRTYKTGDFPDIE-ISYKELIVPLRVFVQQDQSS-ARHLTVALFKSFL-----DEESLQGAD      2749
Query_BEETL  2531   E   FAKEREN-KVTIYRSYRKGDLPDVQ-ITLQAVLEPLQILALRDREI-SKILYEKFLKSIL-------EKRKSDA      2595
Query_THRIP  2621   E   RKLRREG-QVAMHRKYRLGDLPDIQ-IKFADVINPLTELARHDSII-ARKIMVSLVVGVKE----VIDASDNSN      2688
Query_TERMI  2696   E   RGRRLEA-NIASFREYRSGDFPDIE-ISYSAIIKPMQELAKRDAAV-ARHLFVLLFDGL------LGELYDEKE      2761
Query_FLEA_  2608   T   IRSKMKTsKVTMYRQYRIGDFPDIQ-ITNAAILLPLKELAKRDSSI-GQIVLHSIYIGIIA-----ELDSESQK      2675
Query_MIDGE  2676   K   FRQSREGkEVKLYRRYRLGELPDFF-FNTLAVLLPLQALAKIDDQI-ARDVFISIFQSIIQILkqKESGEETEN      2748
Query_LAODE  2677   E[3]QKTRMES-AVHLTRNYTKDPMPDRT-LTNENLVVFLLNLAKKDMTI-ASQLLTSLLKGIIHGYq--KKCSEPNN[ 8]  2757
Query_LOTTI  2726   E   QKSRREN-QVTMYRKYRTGDLPDIQ-IKHSYIIAPLQALAHRDSTI-AKLLFSSLFKAIYKKM--DEVKTEREM      2795
Query_APLYS  2803   E   QRARREN-HVTLYRKYRIGDLPDIQ-ITYQYIIAPLQALAHRDSTV-ASLLFSSLFQAIFRAM--DQVKTEREM      2872
Query_SCALL  2789   E   QKTRREN-QVTMYRKYRIGDLPDIQ-IKYSYIIAPLQAVAHRDSTI-AKILFSELFKAIYDQM--DQVKVEREI      2858
Query_CRASS  2747   E   QKSRREH-QVTMYRKYRKGDLPDIQ-IKYSYIIAPLQAVAHRDSSV-AKLLFGAIFKGIFSKM--DEVKTEREV      2816
Query_OCTOP  2701   E   QKSKHEH-QVTLYRKYRTGELPDIQ-IKYSALIATLQVVALRDSTV-ARLLFNSLFKGIFADA--EQEKGDEDM      2770
Query_POMAC  2757   Q   QRLRREN-QVTLYRRYRIGDLPDIQ-IKYSYIIAPLQALAHRDNTI-AKLLFGAVFKAIYQEM--DKVKTDQEM      2826
Query_WHIPW  2653   E   AAQTKEY-DVRLRRSYRRGDFPDIQ-IPHAAIIVTLQAVAQLDGTS-AGLLLEAIMCGLSEHIs--RNERQGES      2722
Query_TRICH  2662   -   ----KQQ-SLDLRGKYRRGDFPDIQ-IPYSAVIITLQAVAELDRVF-AGALFESLMCGVAANVi--EIYNFEMT      2726
Query_AASTA  2494   Q   EK-AKGT-AVSLFRAYRVGEFPDIQ-IPRMDVLRPLLVVASLHAAT-SSALFSALLVAVLPTL-------P---      2554
Query_AEUTE  2429   Q   AK-SKGS-KVSLYRAYRVGEFPDIQ-IPRMDIFRPLLAVVASHTAT-ATSLFSIILSTVLPNL-------S---      2489
Query_PLASM  2634   R   ERHQE---HVRIKRAYRAGEFPDIQ-ITQQDIVDPIMALCEVSVET-SSLLFDELFSSIVATN-------EFEK      2696
Query_BREMI  2616   R   ERNQRQG-HVMMKRSYRVGEYPDIQ-ITQQDIVDPVMALCEISVET-SGLVFSALFSSITGNY-------QFEK      2680
Query_OLIGA  2664   R   QRKARDQ-QVQMRRKYRVGEYPDIQ-IKLKDIIDPLMALCEMHDAT-SAIVMAEVMASVVATK-------SFDD      2728
Query_LIGNO  2953   E   AAAKREH-RVTLRRRYRVGDLPDVQ-MEQGALVTTLAGLAQKDRMF-GRLLCAQLSQSMVDWAa--SALSEDDV[14]  3036
Query_CESTO  2710   S[1]LSLAEDEtTARLCRRHRIGELPDVEkLTPKALLQPLFQLAKLDSSVcGPALLRLVLESIVVSSg----KSADDD      2781
Query_OVI2_  3099   R[9]------HlQTTLTAKYRTGPLPDVEsISPASLLDPLQRLSAAQPSA-AATVFVNLYCVLLQSVs-----SQTEI[ 1]  3171
Query_PRIAP  2794   L   SRQRREG-QVXMYRKYRIGDXPDVQ-IQFCHLISPLQALAQRDARV-AQLLFSSLCDGVVRRA--GDTATERDV      2863
Query_ROTIF  2686   Q   LKAKREN-HVEKYRNYRIGDFPDIQ-IKHSELIAPMQALAQRDSQI-AMVLFESIFVSILSEI--SYLKDEQDS      2755
Query_SPONG  2782   Q   RRAARQS-YVVMYRKYRMGELPDIQ-IKHSELIRPLQALAQNDSGL-ARLLFCSLFNSIFSQL--DHLLTEEDA      2851
Query_TARDI  2767   N   ERERRSL-DVTLTRSYRIGDVPDIE-ISYEAVIQPLRNLAERDVQT-SSLLLRNLFGGIVAQArtKSEKVDEEL      2838
Query_TRICH  2757   Q   QKIAREN-KVVMYRKYRDGELPDIQ-IKHSELIAPFQALAQV--CY-LDQNYRYVNLSV-NDL--VENFNDEDA      2823
Query_OSTRE  2635   D   IDGLKQD--IKLTRSYRTGELPDVK-ISRNELILPFLVTAERDSVL-ACALLREFSSVAMEEFe-FISKNQHGG[14]  2718
Query_NITEN  2845   A   QAGARVH-KVTMMRSYRTGELPDIQ-IKTEDVLKPLMALAERDPLV-ARTLSAGLFRGAFALP--STALRDGE-      2913
Query_CLUBM  2575   T   EAFERAN-KVQIMRQYRTGELPDIE-IKHEDIIGPMGSLAERDATF-ARLLLSSTFGAFW--------QNDVT-      2637
Query_MOSS_  2718   L   QSLERAN-KVHMMRQYRTGELPDIE-IKHADIVGPLAALAERDPTF-ARLLLRILCSAVYSLP--SGGQ-----      2782
Query_LIVER  2749   Q   QASERAN-KIHMMRSYRTGELPDIE-ITHADILAPLSSLAEKDATF-ARLLVTVICRAVYSQP--AGRQEDIT-      2817
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2817  QSLVAKIESSLNKILE---TSTQYFSPFIGCVLEI---AYL-FPK-QL---KL    QP-GVVSTVCNMSDQ-QPLGILL  2879
Query_LINGU  2823  TEMMAQISRSLNAMLD---SSTQYFSPFIGCVENI---VYN-NSG-RL---KL    EP-SAVDTASLISMQ-QPLGIVL  2885
Query_LIMUL  2855  IVLCNELKQVFNNILQ---QTVQFFPPFIGFVEEV---CYH-SVQ-HL---PL    EP-EEVSKVSQASGQ-QNLGILL  2917
Query_SPIDE  2807  EAFCAQIGSCFENILQ---KSYLYFPPLMSFCLET---LN--DFP-NI---RF    DV-ETVFNACIGSSQ-QSIGILL  2868
Query_SCORP  2819  SEFCRSLSSALNEILK---KSVHYHPQLVAFCYEC---LNC-DVK-EM---DI    DY-EQICTASQASHK-QSLGILL  2881
Query_TICK_  2712  PELCRELEVALSSILV---TSDRCFPPLVAFVQET---SFS-CAG-GL---KL    PV-DAVAKASVASGH-QSLGILV  2774
Query_MONOS  2775  -QLLLTCADNFAHLLQk---SRQGISPVIAMATEV---LYSlDDD-TL--rEL[ 7]-P-ANLCRIAVASDT-MPLAALL  2844
Query_ROSET  2750  -NMKAEYNKAVNGLLQa---TRKGYTPFIAALHNL----MR-ASDpAI----V    SP-MTVSLTSAAGRS-FHTGALL  2810
Query_BELCH  2873  EDLTEQISRFLNTALS---NSEKYFPPFISSLQEI---CYR-EPR-KL---SP    DA-SVIQQTGVISLQ-QPIGIVL  2935
Query_CIONA  2838  ENTPANIGDKMFQILE---ISTEKNPSLIACVEEV---ALR-NHD-YVgfnKL    IP-AHISSSSLASNK-HALGITL  2903
Query_CAECI  2849  KDITQKLLADFNRFLN---ITTSYFPPFVACIQDM---SYH-HTE-LL---NL    DP-SNVSTSCLASLQ-QPVGILL  2911
Query_HUMAN  2830  NNITQKLLQDFNRFLN---TTFSFFPPFVSCIQDI---SCQ-HAA-LL---SL    DP-AAVSAGCLASLQ-QPVGIRL  2892
Query_MOUSE  2831  NIITQNLLQDFNRFLN---TTFLFFPPFVSCIQEI---SCQ-HPD-FL---TL    DP-ASVRVGCLASLQ-QPGGIRL  2893
Query_XENOP  2850  -SVTDKLLKQFNSFLS---NSLSYFPPFIACVQDM---CYQ-HDE-LL---HL    NP-ANISTSCLASLQ-QPLGILL  2911
Query_PLATY  2837  RNITLKLLQDFNHFLS---STVSFFPPFISCIQEM---SFQ-HTE-LL---NL    DA-NCVSTSCLASLQ-QPVGILL  2899
Query_TASDE  2841  KSITEKLLLDFNYFLS---STMLFFPPFISCIQEM---SYQ-HSE-LL---SL    DS-TTVSTSCLVSLQ-QPVGILL  2903
Query_CANAR  2838  NAIVQKLLNDFNNFLS---MSLSYFPPFIACIQEM---SYQ-HRE-LL---EL    DS-ANVSTSCLASLQ-QPMGILL  2900
Query_OSTRI  2808  KIIIQQLLKDFNQFLS---MSMSYFPPFIACIQEM---SYQ-HRE-LL---EL    DS-ANVSTSCLASLQ-QPVGILL  2870
Query_TURTL  2847  NCITQELLQNFNRFLS---MTVSYFPPFIACIQEM---SYQ-HRE-LL---DL    DS-ATVSTSCLASLQ-QPVGILL  2909
Query_GECKO  2847  NSIIQQLVQDFNCFLT---TSVLYFPPFIACVQEM---SYQ-HKE-LL---DL    NP-SSVSTSCLASLQ-QSVGILL  2909
Query_DANIO  2839  ADILKELVQTLNAFLN---KSTVYFPPFISCIQDM---SYH-HKA-LL---GV    EP-SLVSATCLASLQ-QPMGILL  2901
Query_SNAKE  2834  NNIIQQLLHDFNHFLT---TSVLYFPPFIACILEM---TYL-HEE-LL---DL    NC-YSVSTSCLASLQ-QPMGILL  2896
Query_STERL  2856  KRITEEILGDMNNFLS---KTTVYFPPFVTCVQDM---SYR-HKE-LL---NI    EP-SSVSSSCLASLQ-QPMGILL  2918
Query_MILII  2846  KDISKQLVENFNYLLK---ATTLNFPPFVACIQDM---SCQ-HKD-LL---GL    DS-TSVRSTSLASLQ-QPMGILL  2908
Query_BAMBO  2839  SGIGKELLKHFTFLLN---TTTLCFPPFVTCIQDM---SCK-HPD-LL---NL    DP-SIVSSTALVSLQ-QPMGIML  2901
Query_ALLIG  2805  THIVQKLLQNFNDFLS---ISVSYFPPFISCIQEM---SYQ-HKE-LL---DI    DS-TSVSTSCLASLQ-QPVGILL  2867
Query_GAR__  2853  RRITEELLQDMNHFLN---KSTLYFPPFVTCVQDM---SYQ-YKE-LL---KV    DP-SAVSSSCLASLQ-QPMGILL  2915
Query_TETRA  2923  ----QKLKDILSNILAk---SIVHDFNFVSTIHKT---MIMlTKE-----gEL[ 1]DS-NIIKQSGLLSHS-YYTSILL  2982
Query_PARAM  2727  ------LYEVLLGLLQr---SQCKTFIFISTVLRT---LHQ-----VLcksGI[ 3]EN-KIVKKIGLESLA-YYQSMIL  2786
Query_SLEMN  2894  ----ELLGQGVRNILNe---SILYDYGCINCIHRV---AME------L----L[ 7]EA-NIIQRTGQHSMS-FQTALAL  2954
Query_STENT  2435  ----KTI-QGLNRILEi---SENYNSDVISCIHKT---ARE------L----V[10]NA-KFVAISGIRSMC-YQSAIIL  2497
Query_NEMAT  2827  EKTVTDIQGALDNIMT---SSTQFYPPFISCIEDI---CYH-KSM-LK----V    QP-SSVSTASLASLQ-QPMGILL  2888
Query_PISTI  2872  ASVTSDLRSAMNHMLS---SSTQFFPPFISSIQDI---CFH-ECK-LR---SV    DS-ASVSTACLTSLQ-QPIGIML  2934
Query_CORAL  2844  DEIFRDIQDALNNMIT---SSTQYFPPFISCIQNV---CYN-EGK-LS----L    EA-GKVTTASIGSLQ-QPTGIML  2905
Query_HYDRA  2702  ESILLKIQNTLNIIMQ---SSIFYDSTFINAVETV---CLK-NKE-M----NL    EP-QVVSSASIISLQ-QPIGILL  2763
Query_LATIM  2847  KTILLELLGNFNGFLS---NTTVYFPPFVTCIQDM---SYQ-HAE-LL---NL    DP-ASVSASCLASLQ-QPMGVLL  2909
Query_SHRIM  2804  QEWNQELHEALNGILS---SSYMCCPQMLRTVLHL---MI--NND-VR----T    NP-ESLTAACLGSRL-EALGILV  2864
Query_MCRAB  2797  DTWIQQLHSSVNVIIA---NSYMCHPDMLRTVLHL---VI--NND-IL----V    SP-DALAAACLGGHL-EALGVLV  2857
Query_AMPHI  2836  KNWREEIRSALSTLLV---ESVQNSAPVMAAVLHL---LL--SAK-IM----A    DP-KLITNAALSSGQ-ESLGILL  2896
Query_BARNA  2656  DQLRDAVSSGLESVLRrl---RTAEPSLVSAALQT---ALD--AE--L---AL    PP-ALLASVCRLSGQ-LPAGIML  2716
Query_COPEP  2673  ----EEIMEIFNNIFNn---TTVVSTDFIGCLIDV---CIHmNPSdVL-----    DT-KKISNICLSPGL-EQIGILL  2731
Query_DAPHN  2543  --FASKLNTSLQSIVEn-yRLRSGNFNLVAAILEY----FWtNTD-EI---SM    DS-GKLVHLVKGSHL-QALGILV  2605
Query_STARF  2888  ESTKVEINRHLNTMLS---TSTQYFPPFISCVQEI---SYY-HSE-QL---TL    VP-STISTASLTSQQ-MHIGIRL  2950
Query_STRON  2870  ELHVKGIERHLNQMLT---DSNQFDPAFISCMLEI---GFY-HSE-QL---TL    DP-AAISQSCLTSRQ-LLIGIRV  2932
Query_SEACU  2846  EELVKNINQYLNQMLS---SSSQFFPHFISCIQEI---ALC-HSK-EI---SL    DP-VTVTVTSLTSKQ-IHLGIRV  2908
Query_FUNGI  2752  RGI-------IDRIFK---TS-SFTGGFCNFMLNI---CYH-QNR--I---MV    SA-PLISNCALKSYN-FHAGIML  2805
Query_DPURP  2874  RSFQQNIKKSIESIVD---QC-KYNSTLVSSILSI---SEK-NLE--F---SP    EF-SKIKDIGLNSNN-HPMAIIL  2934
Query_HALBU  2901  K---SSCKQSIEKIID---SN-ATNSALILSLLKI---VEK-NLE--L---TP    SP-ASIASASNASGN-HPMGIIV  2958
Query_GLOBO  2626  ------LTQAINGVFK---TSLMSSSPLIGSLLDI---MLH-NNGiIQ---GS    SP-SLISHVASATSN-FELGALI  2683
Query_PALUS  2698  EEYKSVIERHLQDILN---RSVNLHPPVIAAVLRV---LYQ-MP--VT---AC    SP-QLIFRASEGSAN-SELGILL  2759
Query_CALIF  2788  NQFTEKVEKCINNILQ---NTINYDFSFITAILTF---CLK-ENI------YI    DT-NILKTIGMNSNN-QSLAILL  2848
Query_GLOMU  2841  E-YKKTIGRHIQDIFS---NSTMNFPPFIGSCLRI---CYE-VND--I---EI    DP-NMIRKACSLSST-ESLGIIL  2901
Query_LOBOS  2779  S-FKDGLIRDFKNILQ---KSTTLYTPFIGSILRI---CHD-YGQ--V---DI    SA-ELVAKASIRSSN-QHLGIIL  2839
Query_RACEM  2761  D-YMQNIVDLVQDKLA---NSFLYFPPTIGAFLRV---CLD-LNA--A---AI    DS-TVVNTVSERSHN-QHAGIAV  2821
Query_PVIOL  2882  K---DEIKQAIFEIVD---KC-RFNTSLVSCLLNI---FEK-NLD--L---SP    TY-GQITNIALKSNN-HPMAIIL  2939
Query_BOMBU  2624  -SFIGKLVDSMEHIMEn---EQNSNSTITA-ILEI---LLN-AS-----itDC    SP-EIITKASRSNSL-NFLGSLV  2683
Query_AEDES  2773  -EAIGAIDGSIQKIFTe---TKNCDPNLFGALIEV---AMM-RPK----lfNL    PA-EGVTTVASGANM-MTMGVLY  2834
Query_ANT__  2681  -DFRLTIVENLKQILHk---SCERDSSFNAVILET---LLL-LSRdAAivaDC    DP-RDLVKASKANHL-NTLGALL  2746
Query_WHITE  2750  -DFYREVMASIGNILQn---SLSFNPAIIGAFFEV---SYF-CHE----kiNL    DF-DAVSTAAQASGL-LSLGTVV  2811
Query_BEETL  2596  -NFIQEISVGIQTIFNn---SSEFSRNLFRALFDA---LIF-SKS----siVF    PP-ELIVSVSQSSGL-RGLGALL  2657
Query_THRIP  2689  -DFKKLTSASLNNVLKq---SKNFEPSIIGACLEI---AY--HCQ----f-SL    DP-HLVAEVALGSNL-LALGVLL  2748
Query_TERMI  2762  -VYTQKVQKAMGSIVSt---TKDYTTMVMGAVMEI---ALC-HSR----yiTL    QA-EDITTVSRFSGL-LALGCLS  2823
Query_FLEA_  2676  -HFIKEINQSIENIFRk---SQLCNSNIMRTIFEM---ALE-YPK----lmNI    GP-DVISLVSHTSQL-LSLGALL  2737
Query_MIDGE  2749  -DFFQSINNSIMSILKqnlNSKNSDTFLMMTLIEM---TMK-SGK----ylEI    SP-DVLVNISVMNNM-MVTGVLF  2813
Query_LAODE  2758  ENVIKVIDESFRRIMGq---TKQATRGVISVVLDT---CLE-FSN--L-pvNL[ 6]DE---LYDVSNRSGL-HSLGALV  2825
Query_LOTTI  2796  EETIELINTSVETMFN---TTSQYYPPFLSCIMDI---VYV-LRS-RL---KI    SP-SNLSTCAIVSSL-QPLGIVV  2858
Query_APLYS  2873  EELTAQINRSMESLLA---QSTHYFPPFMGSIFNI---LYE-VRS-SL---KA    PA-SEVGVSAIVSNQ-QTLGMVV  2935
Query_SCALL  2859  EPIVEQIKDSLSTVLS---QSVQYHPPFIGCMLDI---VYQ-LRS-QL---SV    DV-TSVGSSATFSNL-QPLGIVV  2921
Query_CRASS  2817  QETIQQIQKSIGVVLT---TSIQYFTPFISCLLDV---LYS-LRS-KL---KV    DV-SDVSSAAMFGNI-QPLGICV  2879
Query_OCTOP  2771  ETSIQSVQNSIRNMLK---ASEQHFPPFISCMLNI---LYD-MDK-EL---PI    DI-KQLISSSAITAKlQPLAIIV  2834
Query_POMAC  2827  EETKEKISDSLGRMLS---TSIHYFTPFMGCVMDI---LHT-LRS-SL---KV    EV-SELSIAAKISNL-QPLGIEV  2889
Query_WHIPW  2723  AKFFGAISSLIADMLS---NMSACYRPLVSCIMNY---LYR-FVD-LV---TF    SP-QDVAYVALRSGQ-PDCGILL  2785
Query_TRICH  2727  EKYFDSISDGLRDIFR---SMVTFHRPLVSCLMNF---LYN-FIE-YV---HF    EP-TDIANVAIGSGQ-PDCGIVL  2789
Query_AASTA  2555  SSIQTQVFGYFETLLEsaqHNTLVVAALHQAYIGL---LCQsKSS-------V[ 3]SP-DIIGTSSISSRN-LVTGQRI  2621
Query_AEUTE  2490  ASTQSQVITKIEALLQqsqHNTFFVASLHQAYLSL---L--rHPS-------T[ 3]AP-DIVGISSIASRN-FTTGQRI  2554
Query_PLASM  2697  SGIVDELTEKIEGTLTlskASSVYVGCIVSAYVAS---MI-rHPS-LF---NI[ 3]RP-KTIGDAGLASGN-HHVSQLA  2765
Query_BREMI  2681  IGIAKRLTEKLNETLIlskTSSIYVDCIVSAYLKA---II-gNPR-LC---DL[ 3]RP-ETIGDAGLSSGK-YHICELA  2749
Query_OLIGA  2729  EGTLRTLGARMESALTtgkSSLVFTSAVHLTYYSC---IM-rKKS-AI---QM[ 3]SA-KVVGDSSLASGS-LHSGELI  2797
Query_LIGNO  3037  LIPTSPFDGKRRNGLSa-vTAATNDPALIGCALDS---LCAvASS-ACgdgGV[12]DY-GLVAAAASAAHL-PTLGALA  3117
Query_CESTO  2782  --FVKGLGASLARLLTrg--GVDQSIACLSLIWELg---MDlEREgII--lSM[ 2]EP-TILSTCAIFSEK-AVIGVSI  2848
Query_OVI2_  3172  NSTLNRFAESLHDILRl-tCRLAAISTTDSALHDVttaCLVlLSK-LVeihSI[12]DS-STVATSAVLTKS-EPVAIVL  3255
Query_PRIAP  2864  DATMEKLSEAVSVMAS---SSTSYFSPFIGCLLDV---CYR-NIS-IL---RL    DP-AVVTTVSQTSMQ-QNLGIVL  2926
Query_ROTIF  2756  YKILKELTQYLNAMLA---SSESFNSNFIACILEI---CIK-NPK-LL---KL    DP-GLVTSACLNSFR-QSLGIIL  2818
Query_SPONG  2852  AATKGVIEGNINTILD---CSTQYYPPFIGSLQDI---CY--YEN-SL---KI    KP-PSITTSSLASKQ-EFTGIML  2913
Query_TARDI  2839  EKLAKRLEGILKHQNV---TDTTFLSTLLNIILQD---SPK-ESK-LL--eSV    SA-AAVTSAAAKSSQ-PNLGIAV  2902
Query_TRICH  2824  AETKAEIKASLSNIVK---KSKHYFPPLIACVQDI---CYQ-QKE-LS----I    QS-SFASVASLASSQ-QPIGILL  2885
Query_OSTRE  2719  KKNVGGIARALKGFISki---EGGDASVVRFALSL---CEFiPKA-----tDI    -PiSDIVRLARRGES-LASGVVA  2781
Query_NITEN  2914  NDVRRPVREGVAAALA---RT-RARTSFVASLHAI---CLE-DPK--M---RV    EA-RSLGDASLRSMN-YHSGILV  2974
Query_CLUBM  2638  -TVKTEIRTGFENALE---RT-HSESTLVGTIQDL---CLK-DRD--A---WI    KP-EVAGSASYNSQN-FHSGILL  2697
Query_MOSS_  2783  -DIKAEVRLGLEAALR---QT-HNGISFVGCAEAL---CLE-DPE--I---WI    SP-KIMGSASRKSTN-YHSGIML  2842
Query_LIVER  2818  SDIKSQMTQGIGKALE---KT-SNGISFVGCIEAI---CLE-DAT--I---WI    QP-RLIGQASLKSTN-FHSGILL  2878
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2880  LE    EGLLHSPED-Q---SPASKRSRRD-RSSRITQDT--E    P   WLELAKLYKSISDYDVLRGIF    SSQI  2937
Query_LINGU  2886  LE    EQLIRDGSI-E---RRPAKRARTS--ATVPSQET--S    T   WLELARLYKSLGDFDVVRGIF    SSKI  2942
Query_LIMUL  2918  LE    EILIKSTTE-ET--EHQSKKARRE--NRSNSPAT--S    L   WIRLAELYKSLNNYDVVRGIF    SSSV  2975
Query_SPIDE  2869  LE    TQLLDGLTV-EP--KQ--KKIRHE--KHIFTQES--N    Y   WMKLAELYRSINDYDVVNGIF    SLTM  2924
Query_SCORP  2882  LE    DHLICYNNF-KS--SNPAKRIKLN--KPLISDEL--S    I   WLKMAEIYKSLSDYDSVHGIF    SNSL  2939
Query_TICK_  2775  LE    EVQRVRALSePR--PPSPKRTRKE------AAQA-cG    P   WLQMAQLYKSLGDYDTVRATL    SCSE  2830
Query_MONOS  2845  LE    RHLMLTAQHpSSg--PASKRRKQL-------GDAdpF[ 4]-[2]WLELGRLYNTLSEFDTLHGIF    -SRH  2904
Query_ROSET  2811  LE    -STLAAAEEeAAm--QPSKRSKMRrARNGLTDDEieR    -   WIYLGRLFKELGDYDTLHGIF    SNQV  2871
Query_BELCH  2936  LE    EQLVQKAGV-EE---PRSKRART--SSSPVSAET--A    T   WIELARLYKSLGDFDVLQGIF    SRKV  2992
Query_CIONA  2904  LE    QYIIDDKHS-VK--EPSHKKQKTHvSGMGNNPVL---    -   WIELARLYKSLDEFDVLRGIF    SGDS  2961
Query_CAECI  2912  LE    RSLMSLV-D-TE--EPPRKKMR---GKTELSPDV--V    R   WIELAKLYRSLGDYDVLRGIF    SGKI  2967
Query_HUMAN  2893  LE    EALLRLL-P-AE--LP-AKRVR---GKARLPPDV--L    R   WVELAKLYRSIGEYDVLRGIF    TSEI  2947
Query_MOUSE  2894  LE    EALLRLM-P-KE--PP-TKRVR---GKTCLPPDV--L    R   WMELAKLYRSIGEYDVLRGIF    SSEL  2948
Query_XENOP  2912  LE    KGLLHMK-V-PD--EPPAKKMRKEkEKAEIPPDI--V    R   WIELAKLYRSIGDYDVLRGIF    SGKI  2970
Query_PLATY  2900  LE    QALIRLG-P-PE--EPSSKKRR---GKPQLPPDV--V    R   WIELAKLYRSVGEYDVLRGIF    SGEI  2955
Query_TASDE  2904  LE    KALIHLT-P-AE--EPPSKRMR---GKPQLSPDV--I    K   WIELAKLYRSIGEYDVLRGIF    SGEI  2959
Query_CANAR  2901  LE    RALMALS-P-AE--EPPSKRMR---GRTELPPDV--I    R   WIELAKLYRSLGDYDALRGIF    SGKI  2956
Query_OSTRI  2871  LE    HALISLS-P-AE--EPPSKRVC---GRAELPPDV--I    R   WIELAKLYRSLGDSDVLRGIF    SGKI  2926
Query_TURTL  2910  LE    QALTRLS-S-EE--EPLSKRMR---GRAALSPDV--I    R   WIELAKLYRSVGDYDVLRGIF    SGKI  2965
Query_GECKO  2910  LE    QALVHLS-SeEE--EPPAKRMR---GKTDLSPDV--I    R   WIELAKLYRSVGDYDVLRGIF    SGRI  2966
Query_DANIO  2902  LE    ESLLHGAGA-SE--EPPLKRAR---GKRELPPDT--E    R   WIHLAKLYRSLGDYDVVRGIF    SGKI  2958
Query_SNAKE  2897  LE    QALVHLF-PvEQ--QPPSKRMR---GKTELSPDV--T    R   WIELARLYRSVGDYDVLRGIF    SGKI  2953
Query_STERL  2919  ME    ESLIKGL-V-PE--EPPAKRAR---GRRELPPDT--A    R   WIHLAKLYRSLGDYDVLRGIF    SGKI  2974
Query_MILII  2909  LE    KSLIHST-T-NE--EPPTKKAR---GRQELPPDTdvA    R   WIELATLYRSLGDYDVLRGIF    SEKI  2966
Query_BAMBO  2902  LE    LSLMHLS-L-PE--EPPAKRSR---GRTQLPPDT--A    R   WVELAKLYRSLGDYDVLRGIF    SEKI  2957
Query_ALLIG  2868  LE    QALIHLS-T-IE--EPPAKRMC---GRTELSPDV--I    R   WIELAKLYRSVGDYDVLRGIF    SGRI  2923
Query_GAR__  2916  LE    ESLIRGS-V-PD--EPPSKRAR---GHRELPPDT--E    R   WIHLAKLYRSLGDYDVLRGIF    SGKI  2971
Query_TETRA  2983  LE[22]KKRKEILKNnQKkaSRARGRKEMQiGEDDYEKDLdeN[22]L[4]WMDLIDIYSHLNEKDQIKGIL[ 4]-SHE  3098
Query_PARAM  2787  TE    ECILRLQEDnEN-------------------QEIrkS[ 1]-[1]WMDLMDLHQSQGNKAQIEGII    -NEL  2832
Query_SLEMN  2955  LE[22]NKKRKIQDNsGQnqKNGTGF-----------------[11]-[1]WFNLIDLYDQIGNDDALQGIW[16]QEKL  3051
Query_STENT  2498  LQ[21]-------------------------------------[ 8]-[1]WVALTKVYEKMGDSDSANGLW[ 9]ENN-  2562
Query_NEMAT  2889  LE    KQLLSQ-EA-EP--K-SHKRMRTE---HSPSSDT-tA    A   WLELAKLYRSVEEFDVLHGIF    SDHI  2944
Query_PISTI  2935  LE    KQLLEI-EG-KS--QNSRKRAKTS---TVPPSEV-tT    C   WIELSKLYRSVGDYDVLRGIF    TGHI  2991
Query_CORAL  2906  LE    KQILQMAED-PG--ERPRKRQKTG---GNPSPES-tA    T   WIELSKLYKSVEDFDVLQGIF    NSHI  2963
Query_HYDRA  2764  IE    KQLISLKNN-IE--EP--KSKRLRaDSQQENYMV---    -   WSELAKIYTSIKDFDFVQSIF    CSQI  2819
Query_LATIM  2910  LE    QGLIQLL-A-AE--EPPAKRAR---GRKELPPDT--I    R   WLELAKLYRSLGDYDVLRGIF    SGMI  2965
Query_SHRIM  2865  LE    RQALLSKR--GE--EPHPARKKMRkDAEALSQDDs--    -[1]WLSIAEMYKNLNMWDVVRGVV    QGKL  2923
Query_MCRAB  2858  LE    RQVLLSQKGlGR--DLPPPQKKLKtHLNS-QLDTs--    -[1]WINMAELYKNLSMWDVVRGII    HSKL  2917
Query_AMPHI  2897  LE    QNVLMIRKD--V--KSSPPRKRQKrGVTSELAQSseD    D[4]MMHLAELYSSLGMWDAVRGVL    QSSM  2961
Query_BARNA  2717  LE    RTLRQPAAVeAA-----TDRKRPRlGLPPPAAEParD[ 1]-[3]WSVLAELYVQYGDRDSARAAL    ----  2775
Query_COPEP  2732  LE    KYLIYRKDSaFAg---MAKKKKIT------SQDQgkV[ 3]-   -IHLSDLYRTAGEHENVKGLF[ 3]SHEM  2791
Query_DAPHN  2606  LE[ 6]--------PfPE----QSKKPKRQ--------EVkhQ[ 1]-[1]WFHVANLYGDIQNVDNVRAVV    -AHL  2656
Query_STARF  2951  LE    ENLIQKPWQ-EE---KSAKRARF--VSREPSPEI--S    T   WIELARLYKALGDFDAVQGVF    SACI  3007
Query_STRON  2933  LE    EVLIKKSGD-DE---RSRKRART--SRPESNQEM--S    T   WIELARLYKAMGEFDVLRGIF    GALI  2989
Query_SEACU  2909  LE    EMLIKKSWE-EE---RATKRSKT--SHRETTREI--S    T   WIELARLYKDLGQFDVLQGIF    GKHI  2965
Query_FUNGI  2806  LE    K-QLSYSQNlGG-----------KtEDR------ge-    -   WAELMKLYRALGDVDTRRGIF    ENKI  2850
Query_DPURP  2935  IE    K-LILKLQShKQ-----------QeKD----------    -[4]WDNLRELYKSLKEEDIIMGLT    EKQM  2980
Query_HALBU  2959  IE    S-IIESLLPtKP-----------SsKQY------aqN[ 2]T[4]WDHLRELYHALKEEDIVLGLL    ERQM  3011
Query_GLOBO  2684  IE    KGLEDDLF--LS---ADTKKLK-TfNRA--------V[ 2]-[4]WVSLAQLYKELQHIEVYESIY    ESRV  2739
Query_PALUS  2760  LE    K-LANDAP--QP---PASKRSR-TgRNA--------S[ 2]Q[4]WVQLARLYRSMNKVDVYRPLY    ETRV  2815
Query_CALIF  2849  IE    NIHLNHQEK---------------sQSI--------I[ 2]D[4]WNNLAQLYKSIGENEIYRNIE    ESKI  2896
Query_GLOMU  2902  IE    RQLLRT-S--VE---RESKRLR-TtSSS--------L[ 2]S[4]WIELSHLYKSIDGHDVYNSIY    ESKI  2957
Query_LOBOS  2840  IE    KQIQEHGS--RE---RSVKRLK-TsGAT--------S[ 2]P[4]WIELARIYKSIDEKDVFKSIY    ETKV  2896
Query_RACEM  2822  LE    AQLISGASEaAE---RPSKRKR-TgGSA--------D[ 2]R[4]WIDLASLYKSVDEAEIFQSLY    VNHV  2880
Query_PVIOL  2940  IE    K-LLNIVLNnNN-----------KkKTT------pdD[ 2]V[4]WNSLRELYRALKEEDIILGIL    EKQS  2992
Query_BOMBU  2684  LE    ENL--------IheTNIYRPPKKKiRRK---YADdsS[ 1]-[1]WLLLTNLYKSMNDVDTVLSIF    QNH-  2737
Query_AEDES  2835  LE    SKLH------------EYDFPEQSiSSRSS-SVTleA[ 1]-[1]WIKLAELNHALNEYDILGDIF    SDKM  2887
Query_ANT__  2747  VE    RSLLPH-----VqeDNLSPTTSKKmRMSDNYICNeeT[ 1]-[1]WSQLASLYKSLNDVDIVLSIF    RGRQ  2807
Query_WHITE  2812  LE    NQISAL--SeSDfsSGPSAKRQRLdRMDET-NLTeeA[ 1]-[1]WFKLAELYYSLSESDVVLGIF    RDKL  2874
Query_BEETL  2658  LE    EYLISC--DvTEveAKKSRGNQ----------LSqdT[ 1]-[1]WIKLAELYKEMEEWDVVKSIF    LEKT  2711
Query_THRIP  2749  LE    -AAATW--NeNV-------------------------    -   WLHLAELYDGLGERDVAHSTL    LHNV  2784
Query_TERMI  2824  LE    -EILIQ--KdTVseAGSSKRSKTFePSTEE-IL----    -   WLSLADLYKHMDEWDVVRGIF    HKYL  2879
Query_FLEA_  2738  LE    NRLN-----------GEFLDNECAnKRSTI-EDNdeY[ 1]-[1]WLKLAGIYQSMADYDTTQGIF    MKAL  2791
Query_MIDGE  2814  LE    SQLNYLLNNeENneTDEIEEPQQKrIKLDN-NDDikL[ 1]-[1]WLKLIELYYKMNEYEVIAGIF    TEKL  2878
Query_LAODE  2826  LE    ---------------------------NSLSKNNsnT    -   WVYLARLYRSLGFDEIVQGIF    EEKL  2862
Query_LOTTI  2859  LE    EQLIQTDST-E---QRGNKRRKTE--GTTVSTDV--N    L   WIEMARLYKSVENYDVLHGIF    SGKI  2915
Query_APLYS  2936  LE    EQLIQASSA-QA--GTTSKRARRE--AETLDADT--V    L   WIELSRLYKSTGDFDVLQGIF    SDKL  2993
Query_SCALL  2922  LE    EQLIIHNPE-E---GRPSKRGRTD--TLTISKEV--P    T   WIELARLYKSVDEYDVLLGIF    TGKI  2978
Query_CRASS  2880  LE    EQLISQDED-E----RPSKRGRSD--NAVVSTDT--L    S   WIELARLYKSVKEYDVLLGIF    SGRI  2935
Query_OCTOP  2835  LE    EQIIKGMNS-CH--PSKAKRPRQS--LNPPPTDV--S    T   WLELAKLYKSINEYDSVLGIF    KDNV  2892
Query_POMAC  2890  IE    ERLITEDSG-G---ARPSKRSRYD--VTPVTKDD--M    Q   WIELAKLYKSLGEFDVLHGIF    SGML  2946
Query_WHIPW  2786  LE    -RCVESGDI-LE--TPSFKRQRRL-------SDDsdS[ 5]S[1]WISLGRLYKALGGLENLREIF    ETHL  2845
Query_TRICH  2790  LE    -RRIEVGFV-DI--SAVTCQEKRL-------KQVneT[ 1]R[1]WVELARLYKQCGELENLKAVF    TKHL  2845
Query_AASTA  2622  LE    ELLIHSTD-------------------ENVSTHA---    -   WNQLHHVLQMVENEPYLLALA    -THE  2662
Query_AEUTE  2555  LE    EIIIRAAD-------------------SRIVSMA---    -   WNQLHHILQMVQNESYMMALS    -AET  2595
Query_PLASM  2766  LE    ELLIYETQLmAGn---------MKlKSNESLTEC---    -   WDLLHKLLSTLHKRDFLVALS    -LAC  2816
Query_BREMI  2750  LE    EILTYKIQ---Sn---------TDlELNELVTNC---    -   WDLLHKLLGTVRKRNFLTALS    -IAC  2797
Query_OLIGA  2798  LE    EFLTRGVA-----------------SSADVTAEA---    -   WDHLYKILATTQKSNLLTALA    -SKC  2840
Query_LIGNO  3118  LE[22]RGRAAKATTeQL--PTAAKRRRMEaGSQESNDDNsaD[ 8]H[4]WLALARLYRDMGEPDQQLGAL[10]EGVD  3224
Query_CESTO  2849  IE[ 8]RCIPHESSSlSPw---TSKRRRFYaSSVPIFILPqeE[15]T[4]WWQLTRLYSSIGRSEEILGWL[ 4]TSRS  2941
Query_OVI2_  3256  LE[20]---APSWSDpPElvYFPQPFQNLLmQSSQASDQSleA[ 1]-[2]WWDLTRLYRSIGRFEEVLGWL[ 1]-DRW  3339
Query_PRIAP  2927  LE    KMLMQGACD-DE--SRPAKRMRAG-SGMGVSQHT--D    T   WIELARLYKSLGDYDTLHGIF    NDKI  2985
Query_ROTIF  2819  LE    EYIINFEED-SD--APLEKKVRLN--PKDSSSEA-sI    L   WIELAKLYRSMNDYDSIKGIF    LVKS  2877
Query_SPONG  2914  LE    KQLLLNVDS-----SKSAKRSRTN--KSNEVTET---    -   WIELARLYKSLEDYDVLHGIF    TNLD  2967
Query_TARDI  2903  LE    NAILLQSAE------PEQKRARQK---ASVKADY---    -   WLGMYQLLQSDGQTDSIAGLQ    LDTV  2954
Query_TRICH  2886  LE    KQIMHL--D-TE--EPPSKRIKTS---S--ASEA-tS    N   WIELSKLYKSIGYYDVLRGIF    SGHL  2939
Query_OSTRE  2782  LE    SILISSEKDgSS-------------------------    -[2]WEALATLCKSIGHDDNALLSF[ 5]----  2823
Query_NITEN  2975  LE    NQLLDQSQDeGEerGSRSKRQKGRgGQAVPRGGQgeG[ 2]E[4]WLQLARLYREIGEDDIVLGVL    QEKV  3045
Query_CLUBM  2698  LE    KQILHGTRPnLN----AAKRHRGPsVQT------sqD[ 2]V[4]WLELARLYNAIGEDDYVLSLY    LNRL  2758
Query_MOSS_  2843  LE    KQIMNETFPePE-pNVSSKRQKGKgGQA------nrL[ 2]P[4]WLELAELYKALGENDIVLGLF    GTHV  2906
Query_LIVER  2879  LE    QELLHETFKeTAtpGSSSKRQKGAaGKV------mkE[ 2]N[4]WLELAHLYKELGEEDIVLSVY    QKKL  2943
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2938  G-SKPATR   LAIEAETRGDYRSAL   QLYNESL--AKEEWD---DGE   HpTEAEINMW    DSSRIECAE-NLT  2997
Query_LINGU  2943  G-TKKITQ   DALEAESRRDYKTAS   HLYREAR--QCDSWT---DGE   -pDGLEQELW    VDAYLLCVD-QLM  3001
Query_LIMUL  2976  N-TQTESI   QALEFEARGDYTKAL   QLYIQLY--EKTVWP---DKE   P-EQAEMDLW    DDAILKCYD-SLT  3034
Query_SPIDE  2925  G-KRDDNK   RALQHECLGQYFEAA   MLYNKLY--EEA------SDE   H-DQAEKDFW    DKSVLNCLN-KLS  2980
Query_SCORP  2940  N-IQEETK   QALEFESRGEYLLAV   KIYSKLY--DQTEWS---AEE   P-NQAEIDLW    DNAILQCYD-YLT  2998
Query_TICK_  2831  D-VDAKAK   EALQKEAQGDYSGAR   DAYRNLV--------------   ----SSGSIW    DDAVLQCCR-LLG  2877
Query_MONOS  2905  G--SGTTR   QAFACLAHDDYLGAI   KAVDEAG-----DAEv----D[1]----AEEELW    -----QQIQ----  2948
Query_ROSET  2872  L-SFEATQ   RAARLEAAGRFDEAA   EVYKECMf-KEEDEE---DDD[1]--PACEVDLW    EETLAECSL-RLC  2931
Query_BELCH  2993  G-TQTVTY   EALEAEARGDFKKAV   ELYNQAM--QETEWP---DP-   QpLAVEEDLW    DDSRLQCYD-HLT  3051
Query_CIONA  2962  E-SNKSTA   KGLALEDCGDYTAAY   KVYNEAL--HKEDWE---NDV   S--EVEKDVW    EDCLINCFK-NLA  3019
Query_CAECI  2968  G-TKPHTK   DALLAEAKSDYATAA   KLYDEAL--NMPDWS---NDE   P-TEAEKDFW    EVAALECYN-RLT  3026
Query_HUMAN  2948  G-TKQITQ   SALLAEARSDYSEAA   KQYDEAL--NKQDWV---DGE   P-TEAEKDFW    ELASLDCYN-HLA  3006
Query_MOUSE  2949  G-TTQDTQ   NALLAEARSDYCQAA   KLYDEAL--NKLEWV---DGE   P-TEAEKEFW    ELASLDCYN-NLS  3007
Query_XENOP  2971  G-AKSITQ   CALNAEAKSDYAKAA   KLYDEAL--T-ETFS---DGD   P-TDAEKDFW    ELASLECYN-HLT  3028
Query_PLATY  2956  G-TKHVTQ   IALLAESKSDYSEAA   KQYNEAL--NMQHWE---DGE   P-SEAEKDFW    ELAVLECYN-HLT  3014
Query_TASDE  2960  K-TKPVTQ   NALLAEAKNDYSEAA   KQYNEAL--NTQEWL---EGD   P-TEAEKDFW    ELAALECYN-HLT  3018
Query_CANAR  2957  G-TKEITQ   KALLAEARNDYAEAA   QYYDKAL--SQEDWQ---DGE   P-TEAEKDFW    ELASLECYD-HLT  3015
Query_OSTRI  2927  G-TKEITQ   KALLAEARSDYATAA   KHYDEAL--SKEDWQ---DGE   P-TEAEKDFW    ELASLECYD-HLT  2985
Query_TURTL  2966  G-TKEITQ   NALLAEARSDYAEAA   KQYDEAL--SKQDWP---DGE   P-TEAEKDFW    ELASLECYN-HLT  3024
Query_GECKO  2967  G-TKEVTQ   KALLAEANSDYAEAA   KQYDEAL--SKQDWP---DEE   P-IEAEKDFW    EFASLECYD-HLT  3025
Query_DANIO  2959  G-TKSITF   TALQAEAKSDYAEAV   KLYNEAL--NKEDWD---DGE   P-TITEKDFW    EIAALEAYN-HLT  3017
Query_SNAKE  2954  G-TNEITH   NALLAEAKSDYAEAA   KLYDKAL--SIQDWP---GDE   P-TEAEKDFW    KIASLECYN-HLT  3012
Query_STERL  2975  G-AKSITS   MALQAEAKSDYAEAV   KLYNEAL--NTEEWT---DGE   P-AEAEKDLW    EIAALESYS-HLT  3033
Query_MILII  2967  G-TKPLTH   TALLAEAKSDYAEAV   RLYDQAL--NTEEWM---DGN   P-METEKDFW    EIAALECYN-HLT  3025
Query_BAMBO  2958  G-TKSTTQ   KALLAEAKSDYAEAV   QLYDQAL--NTEEWV---DGN   P-TDAEKDFW    DIAALECYS-HLT  3016
Query_ALLIG  2924  G-TKEITQ   KALLAEAKSDYVEAV   KQYDEAL--SKQDWS---DGE   P-SEAERDFW    ELASLECYD-HLT  2982
Query_GAR__  2972  G-AKPVTS   LALQAEAKNDYSQAV   KLYNEAL--NTEEWM---DEG   P-TDTEKDFW    EIAALECYN-HLS  3030
Query_TETRA  3099  G-QKDMIV   -ALDLKFSGQIQNSL   KILKEISskNKESER---QNQ[1]-rMAIEDNLN[15]EYQ--DTLAl-MG  3172
Query_PARAM  2833  FkDKPAFD[3]QAMDYKHQYSYVRAN   QQLEELLdeINQEGY---DED[5]MvHYLQQEFE    -DT---CMK--LG  2898
Query_SLEMN  3052  SqDQERTV[4]EAHCLRTKGSI----[4]KVLEQALndKSLTDC------   -lEDHILCEL    ENKQLDSKA-ELL  3114
Query_STENT  2563  --------[4]KALDFKYEGEIIKAT[4]KIIDNSDad------------   -----IKQEI    MRDYLECKG-YLG  2611
Query_NEMAT  2945  G-TKPITR   LALEAEARGDYAQAL   KIYNEAM--SCDD-W---SSK   P-QQQEEDLW    DDARMQCFA-NLA  3002
Query_PISTI  2992  G-TKDITR   EAMEAEARGDYSQAL   KLYNEAI--GTDDVW---IKE   ElRQEEEDLW    DVSRLQCLA-DLT  3051
Query_CORAL  2964  G-TQSITK   DALEAEGRGDYTEAL   KLYSQAT--NCDEWP---AGS   SpLQAEEDFW    DDSILECCR-RLT  3023
Query_HYDRA  2820  S-TNDVVK   RALAAEATNDYSTAL   HIYRKAS--QKENWE---ITP   S--YEEEEYW    DTARLRCYE-NLG  2877
Query_LATIM  2966  G-TKQVTQ   KALLAEAKSDYSEAV   RLYDEAL--NVEEWP---DGE   P-TEAEKDFW    ELAALECYS-HLT  3024
Query_SHRIM  2924  GaIKEETQ   KALEAEATNNHVKAF   LLYRQSL--DTTDWE---EEP   L--AAENRLW    EEWYSNCAA-ILG  2982
Query_MCRAB  2918  GrIKPETR   EALEAEATNDSVQAF   IIYRTAL--N-TQWP---EEP   D--PAEVHLW    EECYASCAA-KLG  2975
Query_AMPHI  2962  SsLHGNTR   EALHCESLDKPLEAY   KFYLSAL--S--SPE---VEQ   S--SEEQKIW    EERLCATAS-ELG  3018
Query_BARNA  2776  GrRADLAA[3]AALEAEARGDFLEAA   RHYEQTIr-------r---GE   -------DHL[ 1]DQAYFQCLS-ELS  2828
Query_COPEP  2792  N-VHEDTK   SALNFEANGNWSEAL   KVYKNLV--ENLDLQ---ENH[1]---KQEVDIW    QKGFYQSQS-KLG  2849
Query_DAPHN  2657  TdCKESTK   KAIESESKHNWRRAF   EFYSQVF---KSDTP------   --SSAEMNFC    DEYRLKCLE-ELG  2711
Query_STARF  3008  G-THEITK   TALEAEARGDYSTAL   KLYNQAF---EMDWD---EG-   AvSQVEEDLW    DESRLNCCN-QLT  3065
Query_STRON  2990  K-TKDVTQ   EAIQAESRGDYESAL   ELYSKAI---SSDWG---PD-   GvAREEEDFW    DEARLQCCN-QLT  3047
Query_SEACU  2966  D-SHEITH   SALQAEARGDYAAAI   KLYDQAI---SYDWE---EG-   SlQQIEEDLW    DEARLECCN-QLT  3023
Query_FUNGI  2851  A-HTDISR   QAIAAEYSGNYLEAL   AEYERGT--KLLDEG---ARK[5]SpSTAEIDIW    ENNRLECFT-MLN  2915
Query_DPURP  2981  S-DIPFTK   KALEFELKGDWVNVL   KVYDEAT--SKLEQG---QLN[3]LsES-ETSLW    ENGRLECYA-KLG  3042
Query_HALBU  3012  G-ELPYTK   RALEAELSNDWVAVL   KAYDDGM--AALESG---QLA[1]TpSARETSLW    ENGRLQCYE-SLR  3072
Query_GLOBO  2740  A-YSLVTK   EAISAELLGDFETAK   NKYLEGL--GSE------NLD   -vSEAELQLW    TTQRLECMN-KLG  2795
Query_PALUS  2816  S-SNHVVK   EAITAESLGQYDRAL   RAYKQAL--ETE------E-K   -lEEHEADIC    HRGLVECLN-KLG  2870
Query_CALIF  2897  K-ANQSLK   DAIEMEVLGLYQDAL   NIYSNTL--ENL------NQT   -tDEYEVSLC    EQGRLECYN-KLG  2952
Query_GLOMU  2958  V-VSELTK   DALNAELLGDYALAC   KHYLEAL--K----S---EAD   -vDDIEVALW    EERRFDCFE-KLC  3012
Query_LOBOS  2897  A-TTEFTK   QAIEAEVVGDYDRAV   KVYFDGI--TKHFAN---EIH   -vDDVEQSIW    AHGRLECLE-HLG  2955
Query_RACEM  2881  A-SNDSAK   DAVESEMRADFVEAH   ERFMDAF--GRF------ENE   -aDSRELELW    QKEGLRCLE-QLA  2936
Query_PVIOL  2993  S-TSEYTK   KALEFELKGDWIQML   KTYDEAT--LKLEEG---AIT[1]ItDA-ETNLW    ENGRLDCLT-KLR  3052
Query_BOMBU  2738  ItNK-DMQ   IASFAQVSNKWERAK   LAYEEAY--G-TESE------   ----LMKEHC    LQGLFECLS-NLC  2789
Query_AEDES  2888  DsDP-RLR   EAIELESGGNFYKAR   EIYLQMI--RDNQAR-----V   ----AEQNFC    FQSYFNCYA-QMG  2941
Query_ANT__  2808  FfGQ-DVQ   EAALAEVDGDWIRAK   NAFTNAY--ERTEDP------   ----LIKEHC    LRGLLEAAN-NLC  2860
Query_WHITE  2875  TnGNVDLI   KATEKEQLEDWQGAR   RHYFEVY--KDETYS-----E   ----KQRQFV    TQSYFKCTA-LCS  2929
Query_BEETL  2712  QcPQ-EVK   NALLAESQKQWKDAR   DLYQSLI--ET-DDV-----I   ----ERKDFY    YDSYFKCFA-HLG  2764
Query_THRIP  2785  PlDK-TAK   DALVAHAEEQWLRAK   QQYFVAL-------K-----D   ----TKYDNF    YESCFECMA-ILS  2833
Query_TERMI  2880  NcGE-NVH   KALKWETSGHWDKAR   MEYEAAL--SSMDDQ-----D   ----GLSDYY    YEALYKCLA-HLS  2933
Query_FLEA_  2792  EcNE-DFA   RGLELEYSGDYYMAM   KCYERVV--SGTTSA---INE   ----TERQFH    KECYFNSFA-KLG  2847
Query_MIDGE  2879  NlMP-ETR[4]KAIDYESNERFFDAN   GIYDKLI--HQNLWR----NT   ----SEKEFY    YQSKFFCLE-KLS  2937
Query_LAODE  2863  A---GDVR   EALEAEAAGEWREAI   QKYQSTAk----ELQr---GK[1]---DEEADFV[ 1]-EAYFNCMA-RMG  2916
Query_LOTTI  2916  D-TKPITQ   EAMMAESKGDYQTAN   KLYDEAL--SIQDWA---GED   -pLEAEIDFW    DDSRMQCLD-HLT  2974
Query_APLYS  2994  G-TKDVTR   TAIEAESRGDYKAAL   RAYREAM--TCESWE---DGD   -pLEAEVDFW    DDNRMECLN-RLT  3052
Query_SCALL  2979  G-TQDITQ   KAVSFEARGDYYQAK   KLYDEALgySDSDWP---DGN   -pEMAEVDLW    DDCRMECLD-NLT  3039
Query_CRASS  2936  G-TQDITK   EAVQMEARGDYYKAV   KLYNQAL--SCNDW----DEN   -pPTIEVDMW    DESRMECLD-NLS  2993
Query_OCTOP  2893  G-CQEVSK   MAIKAEIRGDYSEAI   KHYVTAI--ELST--------   ---KEELCTW    HHCYIQCLE-NLS  2944
Query_POMAC  2947  K-TKEVTQ   LALIEESHSEYKAAI   ALYNKAL--ESADH----DPR   -pLEAEIDLW    DESRMECLN-RLT  3004
Query_WHIPW  2846  Q-VSEYTK   LALKLEYVGDYEKAV   SVYTEAL--SKLDDS---NQS   SlYLAETDLW    EESRLECLN-ELS  2905
Query_TRICH  2846  L-TTEKTK   LALEMEAAGDFEGAA   TIYGETF--SNKQSK---SGD   IvEKAESNLL    YEARIRCLE-MLN  2905
Query_AASTA  2663  S-TVPETT   LAVEAQLKGDLVAAL   QHYKDAEa----KYA---SMD[3]VtI-LERQRW    KSEQFHCLAt-LN  2722
Query_AEUTE  2596  C-VLSETN   LALHAQLSGDLVAAI   QHYKDAEg----KFK---LLD[4]VsL-LEQQRW    KSEQYNCLAt-LN  2656
Query_PLASM  2817  S-TNADSK   RALQEQLSGDLPLAI   AAYKKAEstLRLQME---LLD[4]ViAKNDAMRC    RWQRLNCMEt-LN  2882
Query_BREMI  2798  S-TIPESK   QALEAQLSGDLPLAI   SLYKKAEsiLHSQME---CLD[4]VsVKADAVRC    RWQRYKCLEt-LN  2863
Query_OLIGA  2841  C-VTDDSK   LALEARLAGDLPGAI   ASYRKAEsvLTSEWE---TEP[4]AqAQFEANRC    YWERLQCLEt-LN  2906
Query_LIGNO  3225  NnDDEERF   NALDWCLSGDVQLAL[4]RLFRQADesGTESAS---AAA   DsATAVNNAC    QLARLDCLT-RLA  3291
Query_CESTO  2942  EsTHLLES[2]KAILAKGFGNFEEAT   SQLREILrtPTDENEselDAS[4]IeETDVSDLW[15]REELLHCLS-RLG  3028
Query_OVI2_  3340  SiDSQLFP[3]EALVTISELNYTNAR[1]-AFEQIL-----ETP------   -------HLW[15]QDGLMACLE-RMA  3405
Query_PRIAP  2986  G-TQTCTQ   EALQAEASCDYKLAR   KLYQRAL---EMEWG---ESE   GpGQGERDLW    DISLMQCLD-RLT  3044
Query_ROTIF  2878  DlVCEHTE   RGIDFEAVNDYYSAR   NEYGAAL-----DSE---LLE   NvSQSEQDLW    LESLLKCCN-ELT  2935
Query_SPONG  2968  D-TKDVTR   KALEAEERGDYLTAL   NLYKKAId---FDWP---QGE   SpPQVEEDLW    EDSLLNAYS-HLT  3026
Query_TARDI  2955  FkRSDEVL   SAIQMEARGAYHDAA   DLYGQLI-----NTR---DDD   S----TLAVL    QESHIRCLN-ELG  3008
Query_TRICH  2940  G-TKDITQ   RALEAESRHDYAAAV   KYYNQAM--DCSSWP---DGE   P-QAEEEDLW    EISRLECLY-NLT  2998
Query_OSTRE  2824  --NQSDTM[1]-AIREEMTGNLQNSA   NIYKNLLh------N--vDDD[1]--SPREREFW    THEYFRCAE-RLG  2878
Query_NITEN  3046  A-RLESTR   EALQAQIAGDYTRAW   QGYDKAL--NEMDEM---DQD[3]-ySKLEADVC    YNQRLECLE-SLT  3107
Query_CLUBM  2759  T-TQSISH   QALESQVRGDTDRAL   VLYDKAI-------N---ETP[3]-aSKVEEDLW    YNQRLACLV-DLT  2815
Query_MOSS_  2907  A-KHEKTR   RALEIQLGGDLRQAL   DLYDQVI---AIYED---GSL[3]-iTKSEIDIW    YNQRLACLA-DLN  2967
Query_LIVER  2944  T-EKPITQ   KALEAQLGGDPGRGL   ILYDQAL--DTVDED---SGS[3]QaTQVEQDIW    YNQRLACLL-DLN  3006
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  2998  QWRDVETLSTSDL-    E---GS------SLSAVWDST---YHQ   EV-YL-PPMLRSKIKLI   LNGD[1]S-Q        3047
Query_LINGU  3002  SWEELGTEAMDCI-    E---GL-E--ETGLNKIWEDA---YYQ   EH-YL-PYILRSKVKLL   LQGN[1]D-Q        3054
Query_LIMUL  3035  RWDALEKAVVCRF-    T---DGqR---PNLDRVWEDS---YYQ   EH-YL-PYMVRSKLKKL[1]---E[1]RQD        3086
Query_SPIDE  2981  KWENLESAAQSRF-    M---IDeS---TSLLQIWEDQ---YMK   DV-YL-PYLVRSKLKLL[1]---N[1]KSD        3032
Query_SCORP  2999  QWDLLQLSIINRL-    D---ENqK--QPNLNKIWEEP---YYQ   EN-YL-PFLIKSKLKLL[1]---E[1]RND        3051
Query_TICK_  2878  EWTELQREVDSRLA[ 1]S---AG-----NKWDRLWDGG---YSE   EH-YL-PLYVKSRVKTA[2]-LGS[1]-ED        2931
Query_MONOS  2949  ----LSELFVDEST[ 1]D----EdVleAASLERIWTDR-----Q   AT-ALvDATMTSNLRDG[4]LSFL[6]LRQ[3]     3013
Query_ROSET  2932  DWGELENNTLTCIN[ 9]Dl----aEv--IDANNIWVQPrl--LQ   --rHL-PHLITVRLELG[3]----   ---[3]     2993
Query_BELCH  3052  QWKNLYEVSTVNI-    D---ED-S--PPNLDKMWDDT---YFQ   EH-YL-PFTIRSKLKL-[1]LEGE[1]--D        3103
Query_CIONA  3020  RWDQVNAVCDAELN[ 1]Y---LEdT--NCSLKQVWRVP---EYE   QR-FL-DTSVHSKLLLL   LQGN[1]EAG        3076
Query_CAECI  3027  EWKPLEYCSTVNI-    D---DS-K--PADLNKIWNDP---FYQ   ET-YL-PYMIRSKVKLL   LQGK   -DD        3078
Query_HUMAN  3007  EWKSLEYCSTASI-    D---SE-N--PPDLNKIWSEP---FYQ   ET-YL-PYMIRSKLKLL   LQGE   -AD        3058
Query_MOUSE  3008  KWKELEYCSTVNI-    V---SE-N--SLDLSKMWSEP---FYQ   ET-YL-PYVIRSKLKLL   LQGE   -GN        3059
Query_XENOP  3029  EWKPLEYCSTVNI-    D---TG-K--PPDLNKMWSDP---FYQ   ET-YL-PYMIRSKLKML   LGGN   -ND        3080
Query_PLATY  3015  EWKSLEYCSTVNI-    D---SE-S--PPDLNKIWNDP---FYQ   ET-YL-PYMIRGKLKLL   LQGG   -SD        3066
Query_TASDE  3019  EWKSLEYCSTINV-    D---SE-N--PPDLSKMWSEP---FYQ   ET-YL-PYMIRSKLKLL   LQGA   -SD        3070
Query_CANAR  3016  MWKSLEYCATVNI-    D---SG-K--PPDLSKTWNDP---FFQ   ET-YL-PYIIRSKLKLL   LSGE   -ND        3067
Query_OSTRI  2986  EWKSLEYCATVNI-    D---SG-Q--PPDLNKTWNDP---YYQ   ET-YL-PYIIRSKLKLL   LNGE   -ND        3037
Query_TURTL  3025  EWKSLEYCSTVNI-    D---SG-Q--PPDLNKTWSDP---FYQ   ET-YL-PYIIRSKLKLL   LHGG   -ND        3076
Query_GECKO  3026  EWKSLEYCSTVNI-    D---DK-Q--PPDLNKTWSNP---FYQ   ET-YL-PYIIRSKLKLL   LHGG   -DD        3077
Query_DANIO  3018  EWKSLEYCATVNI-    D---DS-S--PIRLDRMWTET---FYV   ET-YL-QYMMRSMLKQL   QMGE   -TN        3069
Query_SNAKE  3013  EWKSLEYRLTVNI-    D---NG-Q--PADLNKIWSDP---FYQ   EA-YL-PYVIRSKLKLL   FHGG   -SD        3064
Query_STERL  3034  EWKSLQYCSTVNI-    D---DS-S--PADLEKIWGDP---FYQ   EA-YL-PYVMRSSVKLL   QLGE   -PD        3085
Query_MILII  3026  EWKALEYCSTVNI-    D---DK-T--PADIENIWSDT---FYQ   ET-YL-SYMIRSKVKLL   QSGE   -VD        3077
Query_BAMBO  3017  EWKSLEYCSTVNI-    S---DK-T--PVDLEDIWTDP---FYM   EA-YL-PHMMRSKVKLL   QLGE   -ND        3068
Query_ALLIG  2983  EWKSLEYCSTLNI-    D---SE-Q--PADLNKTWSDP---FFQ   ET-YL-PYIIRSKLKLL   LHGG   -SD        3034
Query_GAR__  3031  EWKSLQYCATVNI-    G---DS-S--PVDLEKMWSDP---FYQ   EA-YL-PYVVRGMVKQL   QLGG   -QD        3082
Query_TETRA  3173  KWDDLANELIEINK    D---------------VLEQP------   EKiGV------EKMKIL[4]MQSD[1]-AD[3]     3217
Query_PARAM  2899  QWEKLDAQLKDVTQ[ 1]------------------TDR-----E[2]AN--L-PLIRRTFHKYM[3]KTGE[1]--D        2942
Query_SLEMN  3115  KWDELAKDMSE--K[ 2]-------S-------KQLSEIpf----   KH--A-EKMIRAQIRTD[2]----   ---        3153
Query_STENT  2612  QWEDVKNCLPE--M[ 2]-------L-------K-----------   ---------MKSFMRLN[2]----   ---        2637
Query_NEMAT  3003  KWGDLEEYSKTGV-    D---DEgN--PPDLDKIWTDT---YLQ   EH-YL-PFIMRSSIKLI   CEGE   -ER        3055
Query_PISTI  3052  KWKELEKCSVESI-    D---DK-N--PPDLDKVWSDT---FYQ   EH-YL-PHLITSKLKLQ   CQGS   -GD        3103
Query_CORAL  3024  LWDKMDHNCIVNF-    D---EAnG--VLDLNQIWRDD---YYK   EH-YL-PFLLCSKTKKF   CIGD   -RD        3076
Query_HYDRA  2878  MWTNLKDFTLKSIE[ 1]I---NDsD--DFDCKKMWEND---ILQ   EI-YL-PCYIRSNIKLM   VEEK[1]DLV        2934
Query_LATIM  3025  EWKALEYCSTVNI-    D---DK-T--PADLENIWNDP---FYQ   ET-YL-PYLMRSKVKLL   LLGQ   -GD        3076
Query_SHRIM  2983  QWSELENFVEKRFL[ 1]D----DdG--EVDLNRVWLLP---KPT   AT-AL-PALVHSKL-MN   ILDG[1]ETD        3037
Query_MCRAB  2976  QWSELENFVENRLL[ 1]D----EaG--QVSLNQVWTLP---RPA   VS-VL-PSIVNSKL-MS   ILDG[1]ETD        3030
Query_AMPHI  3019  MWGVLDSFLDQRLL[ 1]D---PEsG--AVDTERLWQLP---RPA   A--AI-TAFVNSRV-MK   LLSN[1]ESD        3073
Query_BARNA  2829  EWGRLSRAVDQRLT[ 1]D---HDqQ---PRLDSVWLAGnd---E   QRdHLlPWLVVSKTQLA[3]-DGT[1]-TQ        2887
Query_COPEP  2850  KWDAVQTAVSKQFN    ------------NLDEIWNVS----HK   SS-VL-QAFMSSSMHLS[2]----[1]DSN[3]     2898
Query_DAPHN  2712  QWDEIGSLNPKKEP    -----------------WNIR----FR   ---------VRNALQNF[4]AARE[1]--D[3]     2752
Query_STARF  3066  QWKELQNYSTVNI-    D---DN-S--PHDLGKIWSDT---FYQ   EL-YL-PHMMRSMLK--[2]MEGG[1]--D        3117
Query_STRON  3048  KWDDMESFSTDQI-    E---GK-G--SNDLSQIWNDD---YQQ   EI-YL-PYMMRSKLK--[2]MEGS[1]--D        3099
Query_SEACU  3024  QWKELENYATMNI-    D---DS-E--QQDISKVWDGG---YSQ   DL-YL-PYMIRSKIKLL[4]IEGA[1]TAD        3081
Query_FUNGI  2916  RWDDLYDNTMVEVE[ 2]P---QE-------LYRE---G---LRD   -P-YL-YYYVISSIKS-   ----   KNR        2958
Query_DPURP  3043  NWNALKENFYSYYN[ 1]P---IS-------VFKE-NN-----CD   -L-LL-SYLFEFNLKV-   ----   KEN        3084
Query_HALBU  3073  EWSNLKANFIAYYP[ 2]P---QS-------VFEEQDEA---TRH   -E-MM-SHFLQYSLKV-   ----   KEN        3118
Query_GLOBO  2796  EWSLLAASVLQDVD[ 2]L---NS----------LWEPE---NVD   -P-YL-KLFLRSHLKLR[4]LADD[6]PAD        2853
Query_PALUS  2871  LWDELAHYIMADLN[ 2]T---EH----------LWEAD---YQD   -P-DL-RLFFRAFLRLR[4]NDAG[6]EDR        2928
Query_CALIF  2953  SWSDLVDNVTDISN[ 2]I---ES----------LWDNN---KKD   ---LL-NYFIRGYIKLH[4]DREK[6]DNN        3009
Query_GLOMU  3013  QWDDLAETVLVDIE[ 2]L---SN----------LWEDE---LQD   -I-YL-QHFVHSFFKLS[4]-NSE[5]---        3065
Query_LOBOS  2956  DWDYLEANMMSDLD[ 2]S---QE----------LWTED---YQD   -P-YL-HYFLTSYIKLV[4]-ESE[6]AEN        3012
Query_RACEM  2937  QWDDIAMQVKDTVE[ 2]Y---FC----------LWDVE---KKD   -P-YL-RAFVHSFTRLR[4]IDQS[6]DDN        2994
Query_PVIOL  3053  NWTHLRDNFNAYYP[ 1]P---MT-------IFEDYNQ-----RE   -T-LF-HYFLQYHLKV-   ----   KEN        3095
Query_BOMBU  2790  SWHEIDRHIKGKLD[ 1]------------NLDNIWNNP---WK-   DW--MfPWVFDVHVRKL   INED[1]TEE        2838
Query_AEDES  2942  MWEDLVSILERQVD[ 1]------------NMEEFWTDE---WN-[1]EN--IlPKYVHSNTRLN   LAGD   ER-        2989
Query_ANT__  2861  DWSAIDRLVKDRAG[ 2]------------DLNNIWNDT---WR-   DW--IiPYACDAYL--Y   ASEE[1]NWD        2908
Query_WHITE  2930  NWNELKQSVLVALD[ 1]------------DYDRLWEEDkklW--[2]EN--IaPWLFSSELHMF   IRSG   ---        2978
Query_BEETL  2765  AWNDLINSVDSIVS[ 1]------------NPTETWNNL---WDQ[4]EK--LlPWYISGHLKDI   LFT-   DSS        2816
Query_THRIP  2834  DWDQMNDVIREEVD[ 1]------------ELDNLWNNS---W--[2]QK--LlPRLLHSEVLGL   VQRN[1]DSK        2883
Query_TERMI  2934  DWKGLNNQVRRQVE[ 1]------------NLDQLWTDE---W--[2]EK--LlPQILLSEIQMM   LDKG[1]DSK        2983
Query_FLEA_  2848  HWDDLNIGIKRQVD    ------------SLDEMWTEQ---WA-[1]EK--LfPLLITSEIHKA   INSS   EN-        2894
Query_MIDGE  2938  DWRVITEEIRNQFN    ------------SYNDAWDEK---FP-[3]ET--LlPHLMKSELRMI   LNND   QDV        2987
Query_LAODE  2917  EWNGLKEALVNVNE    -------------VDEMWNDS---WSK   EN--LlPWLFKAELNSA[1]MSSE[1]-KD        2964
Query_LOTTI  2975  QWKDLETVCLTAI-    D---NT-E--LPNLDKVWEDT---LYQ   EH-YL-PYIIRSKLKLL   LLG-[1]D-Q        3026
Query_APLYS  3053  QWQDLEKVSVRGV-    D---SS-D--EPNLDKVWEDT---FYQ   EH-YL-PYMVRSKLKLV   LAGE[1]N-Q        3105
Query_SCALL  3040  NWNELDTVSTSGI-    E---DG------NLSSVWDDT---FYQ   EH-YL-PFIMRSKVKLV   LQGD[1]H-Q        3089
Query_CRASS  2994  QWKELEGVAMSGI-    S---SN------SISQVWKDT---YYQ   EH-YL-PYIIRSKVKLM   LQGD[1]D-Q        3043
Query_OCTOP  2945  RWDKVEKECVHGI-    D---PEnG----DFTKVLDDS---FNQ   EN-YL-QPLIQSKLKLM[1]---Q[1]DIN        2995
Query_POMAC  3005  EWTKLEMVTTDAL-    S---QE-G--EISLASVWDDT---YYM   ET-YL-PYLIRSKLKLL   LHGD[1]K-Q        3057
Query_WHIPW  2906  DWNRMTALVPCYD-    ----DAgR--T-QYDEIWKDD---HSI   DF-YL-PFLLRSKIKLM   MTS-   APQ        2956
Query_TRICH  2906  DWKGLETWSTTFS-    ----ASgK--TAKLNQIWETD---YLI   NM-HL-PVFLRSKLKIA   LTSD[1]AAL        2959
Query_AASTA  2723  RWDTLLDDVH----    -----P--------DRLWDQHep-FLE   KH--T-KCFLEASIAIQ[2]----   ---        2761
Query_AEUTE  2657  RWDKVLEEVE----    -----S--------LDLWQQNep-FLA   NH--M-GYFLEASMALN[2]----   ---        2695
Query_PLASM  2883  SWHALNSEITEAAN[ 1]----DT--------NYLWKQRpp-YLE   QG--V-GHYLRSCLGLL[2]---S[2]NDL        2933
Query_BREMI  2864  NWETLEAEISIATK[ 1]----NS--------NFLWKQRpp-YLE   QG--V-GHYLHSSISLS[2]---N[2]DSI        2914
Query_OLIGA  2907  NWDKLQKELG---S[ 1]----ED--------GALWNEEpp-YLE   QG--V-KHSLHSTLALI[4]DPEG[2]SEM        2959
Query_LIGNO  3292  DWSRLSGVSKKLLL[ 3]D---GEeG--SGLADRLWSRV---GVQ[3]KNaLS-SYMLSSTRLVC[4]SAIG[1]DSE        3358
Query_CESTO  3029  SWEDLDVYAMDSAE[10]D-kaSTsSvpADELASLWKDP---YLD   NS-IL-PHIIHSRLQIC[4]INKI[6]DTT[3]     3110
Query_OVI2_  3406  SWDDLDVLTSRGLK[ 4]DtsaGSqQlhNVSFSSVWPHLhn---D[4]ET--LlPLLFRARLKRL[4]VQSQ[6]PSS[4]     3487
Query_PRIAP  3045  DWTALDAHTKECLI[ 1]E---GS-S--EADLQLLWEDT---YYE   EN-HL-PFVMRARLKRL[1]VEGA[1]--D        3099
Query_ROTIF  2936  DWKAMCEWSMNGN-    ------------TLDQLFEDA---YSI   EN-VF-PYAFRSKLKLI[3]DENE[1]KKH        2986
Query_SPONG  3027  KWKDLEKASTCNIG[ 1]E---TN-----SKLEQLWKDS---YCQ   EH-YL-PFVLSSKLKLW[1]---T[1]DSD        3078
Query_TARDI  3009  NWKQLRTATIPLQM[ 1]Q---NSqS--SEDMTALWDNN---GHV   ET-ML-PYALRATFNLI[3]-SAE[1]EKA        3067
Query_TRICH  2999  QWDKLEYLSLVNV-    D---NN-D--PPSFDKIWEDN---YYQ   EN-YL-PLFLKSKAKVM   CEND[1]AGS        3052
Query_OSTRE  2879  DWDEIGEMLADVRS[ 3]E---ADiAv--ADCGIAGEASgg--SR[1]STa------FRVAIRQP[2]---D[2]---        2932
Query_NITEN  3108  RWQELVDESLEQVT[ 2]D---AHgRpmDADTQKLWADE---NRD   -P-YL-RLFVLSASKV-   ----   PTQ        3160
Query_CLUBM  2816  RWGEIMEDALLQVK[ 2]E---TD----EYDLEKLWHPK---LQD   -P-YL-GLFVKGALKV-   ----   PTH        2864
Query_MOSS_  2968  EWEVLMGEIEYESM[ 2]D---NS----EPDFKRLWEPT---KQE   -I-YL-GYFIRGAVKE-   ----   AKY        3016
Query_LIVER  3007  NWQLLLEDTLTQVR[ 2]D---SA----EPDLQKLWEPN---LKD   -P-YL-GLFVRGAAKL-   ----   STY        3055
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3048  DN-LLSFIEESL---KTPVHK   AFIEGSFSSELALVCIHQ-ELY    DKAKH---YAN-MSLQAFIQDWANLST---  3108
Query_LINGU  3055  QD-MLDFVDNSM---NVEEQK   IFLEARYSEELALLYLFQ-EDY    ARARH---YTS-IVFKTFLQDWGSIDT---  3115
Query_LIMUL  3087  QS-LLSFFDNSM---QDPDKK   EYIEEHFTEELALLYVIQ-EDY    DRARF---YST-SCLQKFLQDWSSLSI---  3147
Query_SPIDE  3033  PS-LLNFFDESR---KDDENK   THLEIHFNEELALLYVVQ-DKF    KIARQ---YSS-SCLHKFLKEWQTLSQ---  3093
Query_SCORP  3052  QS-LLCFIDSAL---KEENKK   KYLEENFSEELALLYIIQ-ENY    DRAQY---YSQ-LCLQDFLKEWSSLNV---  3112
Query_TICK_  2932  QF--LESIRQWL---EDPARK   RLLEAECVEELVLLFLLQ-KDM    GKADY---YVR-CFAEQFLEEWSGLSV---  2991
Query_MONOS  3014  KEaSVAFYKLREf-vTSSSQQ[4]ELLARRHAYELAVTELLH-DKI[10]ARARHl---VN-TGVQAFEQSWPLLPA---  3091
Query_ROSET  2994  -AcLQAFVEAARp----HDDR[1]AHLQYHFAMELAMLSVLR-EDF    DLARY---HLS-QALQRFLDDWTTLSP---  3054
Query_BELCH  3104  QS-LLNFVDSAM---QKPDRK   ALLESRYCEELAMLYLYQ-DDY    GRAKY---YVG-NCTDAFLQEWSGMDR---  3164
Query_CIONA  3077  NElCEFFDDSMQ---GDDARK   NMVEKNYSEHLAMMYIYK-GDY    DSAKF---YAD-KAADSFSHQWSNLDS---  3138
Query_CAECI  3079  QS-LLTFIDNAM---KVEQRK   ALLETNYNQELSLLYILQ-DDF    DRAKY---YID-NGIQVFMQSHSSIDT---  3139
Query_HUMAN  3059  QS-LLTFIDKAM---HGELQK   AILELHYSQELSLLYLLQ-DDV    DRAKY---YIQ-NGIQSFMQNYSSIDV---  3119
Query_MOUSE  3060  QS-LLTFVDEAM---NKELQK   TVLELQYSQELSLLYILQ-DDI    DRATY---YIK-NGIQIFMQNYSSIDV---  3120
Query_XENOP  3081  QT-LLTFVDEAM---KVEQRK   VLMETFYSQELSLLYILQ-DDF    DRAKY---YIN-NGIQVFMQNYSSIDC---  3141
Query_PLATY  3067  QS-LLTFMDEAM---KKEKQK   TLIETHYSQELSLLYILQ-DDF    DRAKY---YID-YCIQIFMQNYSNIDA---  3127
Query_TASDE  3071  QS-LLTFIDEAM---KKEPQK   TLIEVHYSQELSLLYILQ-DDF    DRAKY---YIE-NSIQVFMQNYSSIDA---  3131
Query_CANAR  3068  QT-LLTFIDEAM---KTEQKK   AIIEMHYSQELSLLYILQ-DDF    DRAKY---YIS-NGMQIFMQSYSSIDS---  3128
Query_OSTRI  3038  QT-LLTFIDEAM---KTEQKK   AIIEMRYSQELSLLYILQ-DDF    DRAKY---YIS-NGLQIFMQSYSSTDT---  3098
Query_TURTL  3077  QS-LLTFIDEAM---KIEQRK   ALIEMYYSQELSLLYILQ-DDF    DRARY---YIS-SGMQAFMQNYSSIDT---  3137
Query_GECKO  3078  QS-LLTFIDEAM---KTEQKK   ALIEMYYSQELSLLYILQ-DDF    DRAKY---YIS-NAMQVFMQNYSSIDS---  3138
Query_DANIO  3070  QD-LLSFVDAAM---KTEEHK   IIMETHYSQELSLLYILQ-EDY    DRAKY---YAN-NCMQVFMQNYSSIDP---  3130
Query_SNAKE  3065  QS-LLTFIDEAM---KAEEKK   ALIEMYYSQELSLLYILQ-DDF    DRARY---YVK-NAMQVFMQNYSNIDS---  3125
Query_STERL  3086  QS-LLTFVDNAM---KVEERK   SIIETSYSQELSLLYILQ-GDF    DRAKY---YVN-NCMQVFMQNYSSIDT---  3146
Query_MILII  3078  QS-LLTFVDRAM---KVEQRK   AVIETYYSQELSLLYILQ-DDV    DRAKY---YIN-NCMEVFMQNYSSIDT---  3138
Query_BAMBO  3069  QS-LLTFVDNAM---KVDQRK   TVIETRYSQELSLLYILQ-DDF    DRAKY---YVN-NSIQSFMQNYSSIDP---  3129
Query_ALLIG  3035  QS-LLTFIDEAM---KIEQRK   ALIEMHYSQELSLLYILQ-DDF    DRARY---YIS-NGMQVFMQNYSSIDA---  3095
Query_GAR__  3083  QS-LLTFVDNAM---KVEERK   SVMETNYSQELSLLYILQ-DDY    DRAKF---YIN-NSMQVFIQNYSNIDA---  3143
Query_TETRA  3218  WEeLKDYVKLWM---NTEYGK   SVLEGPLSYEMALITILD-GDI    DRTRY---YHK-LSQNQFLRVWSNFGS---  3279
Query_PARAM  2943  -------LQKAKyl----ERR[1]EHLADEFLLEKTLILISE-QDI    DQSRY---FLQ-KAQRRFLTQWSSLLSqne  3000
Query_SLEMN  3154  WNnLTKQIAFWQ---KDQVSK   AYLDKNFNYELAINSLTC-KDY    DRARF---YLD-KETNELLSQWKNLTK---  3215
Query_STENT  2638  -SqLTEEVC--------KLPS   NSLSLRFPYEMSLLNLTQ-DDT    DRARH---YLD-LEFTHFVEKWQSLNP---  2693
Query_NEMAT  3056  PC-FLEFLRHSM---MNEERK   AFLESRYGDDLALLFILR-DDY    GRARY---HVT-NCFQAFLKDWSGLSS---  3116
Query_PISTI  3104  MT-LNEFVAKAM---KTSERA   ALLRSRYSDSLALMYVLQ-DNC    DSASY---YAA-LSEQSFLSDWAGLDS---  3164
Query_CORAL  3077  EV-FFNFIFDSL---KNEERK   VYLENKYSDILALLFILD-EDY    DRARY---YTS-ICLQAFLEDWSGLDS---  3137
Query_HYDRA  2935  HKeFQSFIDSSL---SDEYSK   CQLQMRFCEDLALLKILQ-DNY    GAAHH---FLK-IAETSFLDEWSLTSQ---  2996
Query_LATIM  3077  QS-LLTFIDNAM---KVEERK   ALLETHYSQELSLLYILQ-DDF    DRAKY---YVG-NCMQVFMQNYSSIDT---  3137
Query_SHRIM  3038  GN-LIAFIDQSM---GDHRHR   SMLEGDLPLQLGVLAAHQ-EKF    EQANV---YAN-TATSMTLLTLSQYSV---  3098
Query_MCRAB  3031  RN-LNAFIDSAL---GEPNHR   NLLESTLPLPLAIMSAHQ-QKM    AQARH---YIT-LATSKVLLTLAQSSL---  3091
Query_AMPHI  3074  GG-LLQFVNNAM---KDEPRK   MLLEDNVSLQLGVLSVYQ-ESP    TIART---FLL-SATNALVGKLGHLSH---  3134
Query_BARNA  2888  --qLHQFLDAGLa---EEASR[1]-LLESRFPEELALRHVAC-GQL    TRARH---YCH-MAQESFVRDWSAAEAagg  2950
Query_COPEP  2899  ----YNFLNNSL----TDKTK[1]SVLSGTCALQLAVLFGAR-GKL[10]KESKR---------LQLFSQNILLSGDp--  2962
Query_DAPHN  2753  ----LPTVDPVLl--------   --------ADLAALALRN-GNG    DRAAIc---IS-KATDNFSSRYSQLSP---  2797
Query_STARF  3118  ET-LLQFIDASM---SDADKR   AFLEARYSEELSTLYTIQ-EDY    DRAMF---YLA-NCRQSFLHDWSGLGP---  3178
Query_STRON  3100  DS-LLGFIDSSA---KDQEKR   VILESRYTEELAALSILK-EDY    NRAKY---YLA-NSTSSFLQEWSGLGP---  3160
Query_SEACU  3082  ES-LLQFVDNSM---QQQDRR   NILEVRNSEDLTGLFVIQ-EDY    QRARY---YLD-NCTKSFLEEWSSMNP---  3142
Query_FUNGI  2959  WEqLWQFISTSF---EDKTKR   GILESEYLKEMAFVAVTQ-NDF    DRGRY---YVS-KFWHQFMSLWSKLQS---  3020
Query_DPURP  3085  WNqLYQFIGKL-----EPHQY   QYLEDKFSGELAFLEVTR-SDH    NKAAY---YVS-KFYQNFKQQWSSSHP---  3144
Query_HALBU  3119  WPdLYQFMGSL-----TPSQH   QYMESSFPGELAFLEITR-SDF    NRSQY---YIQ-KFYRLFRDQWASTHP---  3178
Query_GLOBO  2854  DNpIKKFLDEAK---KDAKKR   SVLELLFSDDLLTSAVIA-RDL    NLARY---FVD-ISWKQFVAAYAELNT---  2915
Query_PALUS  2929  PNpLYDFISKAM---EDAAKR   DSLGSLFSKDLSLALVAS-GDL    NRALH---YVQ-RAQDEFLQGFATLNP---  2990
Query_CALIF  3010  KNpVIQLLEDAL---KDNNKR   DLVNTNYPYELSVISMME-NNI    NQTRF---YIY-QTYEMIFKLLSNSNY---  3071
Query_GLOMU  3066  ----FDFIDGAM---NNQYQK   SLLTTQYPIELALSSVTK----    KDLEQgrfYVR-KAFDNFLFAWSTLHP---  3123
Query_LOBOS  3013  PNpLFQFIDEAM---NNPSHR   QILTTQYQPELALAAVMR-RDF    KQASH---YVR-RSYNRFLSIWSNLHP---  3074
Query_RACEM  2995  PNpIYVWMSDAF---QSKDKE   EYLLQNHACDVALVTMYQ-GKY    ERALR---YLQ-HSYDDMLTVWSSLHP---  3056
Query_PVIOL  3096  WPsLYQFISSL-----NLEQY   QFLETKFPGELAFLEITR-SDH    NKAQY---YIQ-KFYQTFKEQWASSHP---  3155
Query_BOMBU  2839  FQnDFKIMESWL---NDTTKI   KHMKRFFGEECSIFFLHN--QL    EVARE---FLL-NSSDEIREQWIRLHP---  2899
Query_AEDES  2990  GRrFIALLETWM---HVPDRM   DYIKSNFGEEIAMLQIAS-GEY    SRAKM---YSN-QVMRQFLEEWSYLDV---  3051
Query_ANT__  2909  SNnDLEVIQSWI---YDRDKL   QTLMPVTGESLVIFLLKK--DI    KKATD---LLN-KLLDMTGKQWVELSP---  2969
Query_WHITE  2979  NSeFPQKINQWI---HDKDKE[4]NELLNYVGEQLAVWNAIEvDSF    ERGKL---YAK-GHLSNLLNDWAHLNP---  3045
Query_BEETL  2817  TQhFLSSLNSSL---ESSHLA   -HLETHYLEEITMLWIFK-NDI    QEAEH---YLK-IYINNFLEDWQQLDS---  2877
Query_THRIP  2884  --qFLNVINRWL---KDDAKS   STLSSRFSEELSIICLMK-KKF    PQSRF---YGS-KTLSKFVDEWSELNP---  2943
Query_TERMI  2984  REqFLQQVDRWL---RDRDKS   DYLSSHYCEELAAIFLIQ-DKL    DQAQH---QVE-HSLTMFLDSWAALNP---  3045
Query_FLEA_  2895  -SeFSFFLSSWL---SDSEKA   AYLKTNFSEELSIFNIIN-GDI    FGGGY---YLE-QKLIMFLKDWSNLNF---  2955
Query_MIDGE  2988  DDeFMKILEEWI---NDDVKC   KYLKENFPEEIIMLHILD-SKI    AEGQL---EAE-KAIRFCCEEFSGLER---  3049
Query_LAODE  2965  ----ISFFEKVKvwlSDSDKS[1]-YIKTNIGEELSMLYLVN-KSI    DRASE---VSQ-NNRKLFLEDWMQLPA---  3025
Query_LOTTI  3027  QS-LLTFVDTSM---KTTEYR   AVLESRYSMELSLMYLWQ-DNY    DRARH---YAN-LALQKFLQDWSSTDT---  3087
Query_APLYS  3106  QP-LLTFMDRAL---SDEGKK   QYIQSRYCEELALLYLWQ-EDF    DRARH---FTS-LAFQRFLQDWSSTTS---  3166
Query_SCALL  3090  QS-LISFVDEAM---KIPEHK   SLLENRYCGDLALLAVWQ-EDY    DRARH---YTA-MAVQRFQQEWSSTDT---  3150
Query_CRASS  3044  QP-LLTFIDDCM---KQTEQK   ALIESRYSEELALMYIWQ-ADY    DRARH---YST-LATQQFLLDWSSTDS---  3104
Query_OCTOP  2996  QTeFLKFIDNAL---KSPFKS   NILEHQFSKELSLMFLWQ-ENY    DKALI---YIN-NAIQQFLLEWSNTDI---  3057
Query_POMAC  3058  QS-LLSFIDDSM---KNPQHR   AHLESRYCEDLALMYVWQ-KNF    DLARH---YAR-LAVQSFFQDWSSTDS---  3118
Query_WHIPW  2957  EE-FISFVERGE---ADEANR   QILLSRHAGCLALLNIVQ-RDY    VKAKF---LTH-TAFDTILMTYPSLRV---  3017
Query_TRICH  2960  EE-LSAFVEQNQ---TVPKNR   EILENRHSGLLVLLNILR-EDF    TKSQY---FLN-VAIDSFLKCYPTLRV---  3020
Query_AASTA  2762  -SlLREFIDPILqptRQQDTK[1]QYLVTRFPDLIACANMQL-LQW    SHAQA---AVD-EFYATCLHQW---SS---  2823
Query_AEUTE  2696  -DrLKSFFGSI----NNQSDK[1]RYLNSRFPDILAAIYLQC-QQT    SDAQA---AVE-DFYQTCLNQW---SS---  2753
Query_PLASM  2934  LEsLQNFIGDA----VRDSAK[1]KLIQSKFSVELCLAYLAS-GDK    NQTRV---CVE-NFYLNFLKIWRQTSP---  2995
Query_BREMI  2915  LAsLRNFIEGS----SRDLRK[1]ELVQSKFPVELCLVHLLS-NDN    TNARF---CIE-TFYSTFLTMWRHTSP---  2976
Query_OLIGA  2960  ISvLRFFLDDA----HKEPKR[1]DLLQTNFGVELSLSSLVF-GDA    NQARV---FVE-KLFSSFLARWNGTSR---  3021
Query_LIGNO  3359  SAeFMDFLRCTT---DNDSVR   TQFESEHAELLASVLIKT-GKY    HNAMA---YVD-SALDRVLNNVSSVPA---  3420
Query_CESTO  3111  -SkFARIMDSYL---RSTSQD[1]---EPKYACELAILSILR-NDL    SKAQF---YCK-EAFKYLAQTWSSTEG---  3169
Query_OVI2_  3488  NMdLCDFVESAL---EHDNVR   SLFEVRFSEELSIFYCWK-GDF[ 1]-RALY---FQE-RALKQHASDWRESIAcdp  3552
Query_PRIAP  3100  QT-LLSFVDAAT---GHPDRR   AVIETWHGEELALLYIVQ-GDY    TRARH---HAA-RCLQAFLRDWASLDA---  3160
Query_ROTIF  2987  KD-LIRFINGL-----EVEKK   KYLEQSFCQEMALISLHQ-NDL    SAANY---YAH-LAIQKYLTEWSSTNK---  3045
Query_SPONG  3079  HG-LASFLDKSL---SEEANR   SVLESHYPQELTLSYVLR-DDY    DRAYY---YLS-LSLKAKQQIWSVQDG---  3139
Query_TARDI  3068  GDaLDAFIKDAM---EISGRR   DFLQRKFGFELFIYQLLA-SNL    SEADF---FLNtHAVKNWLEEWAGTTQ---  3130
Query_TRICH  3053  DD-FIQYLNKSL---ANDEHL   SVLENRHCDCLALIHVLL-EKY    DRARY---YLD-FSIQKFLQDWSALDS---  3113
Query_OSTRE  2933  -----AFFANIFa-nGESALR[1]---HRSVGVELSLHELTN--NR[ 1]EQARNa---LE-HVFEGFLEAWSATCAss-  2991
Query_NITEN  3161  HVaLQEFLEASM---DDPAKR   ARLTDDFGLQLTAMALAG-GQY    DRARF---HLS-NCYRNFRRKWGALHP---  3222
Query_CLUBM  2865  QSlLTKFLQTSL---VDSTRR   DVLLQRFGTQVAASAVFN-DEI    DHARC---YLA-ECYKTFRKRWSALHP---  2926
Query_MOSS_  3017  HGlLTEFVQSSS---EILPKR   QFLLNEFSTQLATLAAIN-KEW    DRARF---YLG-QCHLIFRRQWSALHP---  3078
Query_LIVER  3056  GSlADEFVQSSF---KEPLKR   EVLLKNYGTQLATLAVVN-DQF    DRARY---YIS-YCYKSFRNQWAALHP---  3117
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3109  TRTSTRA--TKLQKLQ   EFTEMQEFLSVIK-TT--sGEDFPR    QIDRTLHKWE   ARSPDRLL   DSVVIWDN  3170
Query_LINGU  3116  LMTTTRI--GKLQILQ   KLTEMQEFLDFVS-NE--dTFQDSQ    AALSLVTRWQ   TRSPDPIV   EPCNIWDD  3177
Query_LIMUL  3148  MTPGVRE--SKLHSLQ   KLVELKQFLDFMAEEN---NFNSER    PAKSLLKQWE   LRLPSVA-   DPLHVWDD  3208
Query_SPIDE  3094  VALEMQH--SVLQNLQ   KYVELDEFLNLIDRNN---SYQ-YK    SAIKTLKYWQ   KRLPNAI-   DPVDIWDD  3153
Query_SCORP  3113  LMPIIRK--TKLQSLQ   KHIELQEFLNFIDTYR---NID-EY    HFKNLTDRWL   KRYPDIF-   DSVNIWDD  3172
Query_TICK_  2992  LTPALLEg--KLESLM   VLSDLAEYVKHHTRRG------TED    DVDRLLSKWR   LRVP-SNA   DSVLLWND  3049
Query_MONOS  3092  LQTGIRL--TRVEHVQ[1]-LIELDEYLDFMGTARN---FEGSA[ 1]PMEQLARRWR   TRDLTQES[2]--FDAFEG  3154
Query_ROSET  3055  LMTETRR--ASLQRLQ[1]-MVELQEFLDFVGQETN---LASET    SVNTLLSSWS   KRTPTPYK   DPIHVWDE  3116
Query_BELCH  3165  LMTSSRA--AKLQSLQ   KLTEMQEFLELISNDR---NFETPW    PVTRLLEKWS   SRYPDNRM   DPISVWDD  3226
Query_CIONA  3139  LMKD------RKENLQ[4]RLNDINDFLDLVCGENE----ISSQ    SVKNLVSCWS   NSLPNTNF   GSITAWND  3199
Query_CAECI  3140  LLNKSRL--TKLQSVQ   ALTETKDFIHFISKPS---NLTSQT    SLKRLLRIWS   SRYPDSKM   DPMNVWDD  3201
Query_HUMAN  3120  LLHQSRL--TKLQSVQ   ALTEIQEFISFISKQG---NLSSQV    PLKRLLNTWT   NRYPDAKM   DPMNIWDD  3181
Query_MOUSE  3121  LLYRSRL--AKLQSVQ   TLAEIEEFLSFICKHG---DLSSLG    PLRRLLKTWT   SRYPDVVT   DPMHIWDD  3182
Query_XENOP  3142  LLYQSRL--TKLQSVQ   ALTETQDFISFIRKPG---NVSS-S    SLRKLFQGWM   KRYPDSKM   DPMNIWDD  3202
Query_PLATY  3128  LLLQSRL--TKLQSVQ   ALTEIQDFTNFISRPS---NLSSQA    SLKRLLRIWT   NRYPDAKM   DPMNIWDD  3189
Query_TASDE  3132  LLHQSRL--TKLQSVQ   ALIEIQDFVNFMSKQG---NLSSQA    PLKRLLRIWT   NRYPDAKM   DPMNIWDD  3193
Query_CANAR  3129  LLYQSRI--TKLQSVQ   ALTEIQDFINFMTKRS---NLASEA    SLKRLLRTWT   SRYPDAKM   DPMNIWDD  3190
Query_OSTRI  3099  LLHQSRM--TKLQSVQ   TLTEIQDFIHFMTKTS---NLASQA    SLKRLLRIWT   SRYPDAKM   DPMNIWDD  3160
Query_TURTL  3138  LLHQSRL--TKLQSVQ   ALTEIQEFINFISKPS---SLASQV    SLKRLLRIWR   SRYPDDKM   DPMNIWDD  3199
Query_GECKO  3139  VLHQSRL--TKLQSIQ   TLTEIQGFVNFISRSG---NIASQA    SLKRLISIWM   SRYPDSKM   DPMNIWDD  3200
Query_DANIO  3131  LLNRSRL--TVLQSVQ   ALTEIQDFLNYITGD------VSVN    SLKFMIRRWT   SHYPDAKL   DPMNVWDD  3189
Query_SNAKE  3126  LLFNSRM--IKLQSVQ   ALTEIQDFINFMSKES---SLTFQA    SLKRLLNIWR   SRYPDTKM   DPMNIWDD  3187
Query_STERL  3147  LLPGSRL--TKLQSVQ   ALTEIQDFLNFMTKD------VSQT    SLKSLIRIWT   DRYPDAKM   DPMNVWDD  3205
Query_MILII  3139  LLHQSRL--TKLQSLQ   AVTEMQDFLHFISKEY---NFISEV    PLRKLLRTWT   GRYPDAKM   DPMSIWDD  3200
Query_BAMBO  3130  LLHRSRL--TKLQSLQ   TLTEIQDFLHFITKGS---NLTSEV    PLRKLVRTWI   GRYPESKL   DPMSVWDD  3191
Query_ALLIG  3096  LLHQSRL--TKLQSVQ   ALTEIQDFINFMREPS---NLASQA    SLKRLLRIWT   SRYPDAKM   DPMNIWDD  3157
Query_GAR__  3144  LLYGSRL--AKLQSVQ   ALTEIQDFLYFITKD------VSTA    SLKTLIRTWT   HRYPDDRM   DPMNVWDD  3202
Query_TETRA  3280  FATASKHe--YLSNIY[3]---ELKEYLHLDGKIEA----TYED    -FNSLFKNWQ   NRKPSNTY   DNQICWDQ  3339
Query_PARAM  3001  LLSKQLRh-KHLEQLQ   MMHEMGEIQSIYSGVEI--DIDKKM    ---QILNTWA   VRSPSLQD   -SLYTWNR  3060
Query_SLEMN  3216  LSQVAQH--FLVQRIQ   KIYEMREFLQSIKQEGLqtKSDIVP    QVIKSIHKWM   ARVPSYAF   DTIVVWDD  3280
Query_STENT  2694  LSLIAKH--KLVQKLQ   KIFELSEFVTNIKPKKDeeEQEHKQ    NLISILNSYK   TRTPSIAL   DDINTWEE  2758
Query_NEMAT  3117  MMVSSRA--SKLQSIQ   KITEMQEFLEFVVKEG---KLLKKT    -LEDLFSRWS   VRFPNSKL   HPISVWDD  3177
Query_PISTI  3165  NLRSSRS--IKLQSLQ   KITEMQEFLEFIAEEG---NFSSPM    PVLSLLNKWY   SRLPHPIL   HPIGVWDD  3226
Query_CORAL  3138  IMTSSRS--AHLQSLQ   ALTEMQEFLDFISHED---NFSTIS    SVKSLLKRWS   SREPDPKL   HSVGVWED  3199
Query_HYDRA  2997  FINDDCMi--KLNKLQ   RVEEMISFVSLMTSDNY----KDTA    KVQEFFSYWS   NRLPDPKK   HSTIAWND  3057
Query_LATIM  3138  LQHKSRL--TKLQSVQ   ALTEMQDFLGFISKPD---HLNSRA    SLRRLLRNWS   DRYPDAKM   DPVNIWDD  3199
Query_SHRIM  3099  LTPNPLL--DTLQNVQ   LLTELSDCLNNVKYAD---RDFYSS    KVKQTVTSWK   KQKCRIDD   -NPLMTQC  3159
Query_MCRAB  3092  LTPKPLT--GTVRDIQ   LVTELGEFLETLDYTH---NENYPN    KVRKTVKNWK   TQETCLTD   -SSVLMQF  3152
Query_AMPHI  3135  LSPKPLA--AALRHLQ   LILEMDQYLSILQENS---EAGKKA    LASKLCSSWL   HNLPRKED   -DAVFAQF  3195
Query_BARNA  2951  ---QSRTp--LLQRLR   TTTELQQFLDAVTAPDP----TAAA    --SSLLDHWE   RNYPAAQE   -SLAFWER  3005
Query_COPEP  2963  ------I--SELLNLQ   AFCELEQYLTQLNLSNP--------    --NKILKLWE   QNAPVSST   -DLNEWDN  3010
Query_DAPHN  2798  LNFGGRR--NLLEHVC[1]-ISEIQTFMEH-----------SDI    ----------   --IPNASH[4]DSLLVWDS  2843
Query_STARF  3179  LMISSRR--AKLQCIQ   RLTEKQNFLKLAIKPDLqeAIGCQK    S-LELLDRWS   GCHLDSKK   DPINIWDD  3242
Query_STRON  3161  LMLTSRL--SKLQDIQ   KLTEAEQFLQLANKPF---SEGLKS    SATRLVSDWS   RNLPDPKR   DPISVWDD  3222
Query_SEACU  3143  LMTKTRS--SKLQGVQ   RLMEIHQFVDFMNKSDVtqSEGCRK    S-RQMLERWS   SCLPDPKK   DSVTVWDD  3206
Query_FUNGI  3021  FDFQAKQ--IQLQQLQ   NVVELEEFLEF---VSR-eAHLENT[ 1]ALNTLMDGWL   QRYPSSEV   DPLTVWDN  3082
Query_DPURP  3145  LAIASRH--RILQPLQ   KIVEVEEFLNLT------sNQNISN[ 1]QLDHLLSQWK   KRFPSK-Q   DDIMVWDD  3203
Query_HALBU  3179  LALASRH--RILQPIQ   KVVEVEEFLGL---VTN-dKRAIEV[ 1]KMEKLVSLWK   TRYPSS-F   DDILVWDD  3239
Query_GLOBO  2916  SALVSRI--DLLASLQ   HIYEFDEYLRILKLPGL-aSTKD--    --EQLCSLFA[3]HRYPSA-T   DSISVWDD  2977
Query_PALUS  2991  LAYEARL--SKLSQIQ   VVMELRSFLQLSQDHKS-rSLRF--    --EALIKQWA   -RYPSIDL   DSIEVWDD  3049
Query_CALIF  3072  FTNNNHL--IKASEIQ   SILEISEAVDFIENINS-dHSKYIF    --NKMISKWD   TRYPSNNE   ISIDYWFD  3133
Query_GLOMU  3124  LAIGTRL--GKLSSLQ   QIVEIEEFLNLIQSSEK-dQI---W    --DSMISHW-[1]RRFPSKTL   DSSNTWDD  3182
Query_LOBOS  3075  LSEKPRL--HELANLQ   RVVELEEFLGAVDKALR-dPMASFL    --ASLISKW-[1]HRFPSVST   DTMVTWSN  3136
Query_RACEM  3057  LAHVSRM--ARLSNVQ   RTVELEEFLNIVAEYQR-kNKALST    --ADRVTNYI[4]ARYPDPKR   DGTDVWDD  3122
Query_PVIOL  3156  LAIASRH--KILQPIQ   KIVEIEEFLSLSLFSIN-sDVTINT[ 1]QLNSLLKQWK   HRYPSL-M   DDIEVWDD  3219
Query_BOMBU  2900  LSTQLRI--CKLQKLR   IINDVNKFIKILKSAESstDLNE--    ----ILKFWD   NSMPSSLD   -AILPWDK  2957
Query_AEDES  3052  LSDKLRV--SKLLDIR   KVSEIYSFSSLLSNR---mKEDQ--    -LKKLVFNWS   SSFPTASD   -SPTVWDT  3109
Query_ANT__  2970  LCIELGI--CKLFKLQ   IMNDLDASLKILRCTKKtdYLNKTV    -SKALFNFWS   MKTPTIRD   -NLIQWNK  3032
Query_WHITE  3046  LSHQLRT--KKLLDFQ   LLSDI-LNMLFATDSFAltQMTK--    -MERLLKDWN   KHSPAISD   -SLTHWDT  3105
Query_BEETL  2878  FYANLRY--MKLLKLR   NMIDI-DKFLDVLNNISieNVVE--    -IDGLVKYWN   RAAFENLP   -SILLKED  2937
Query_THRIP  2944  LFRKRRE--VMLLGIH   SITD---AMAFFVEGKElqSVEE--    -SRNVLKKFE   KVLPVPQD   -SLMSWET  3001
Query_TERMI  3046  SSHNLRT--KQLLDLQ   KTAELKACLAFATNGEHskQELL--    -AKELVKRWN   KNVPSAFD   -SLLRWDI  3106
Query_FLEA_  2956  LSNELRL--MKLMDIQ   KTSEIQMFMRTTKKSDKieYAAT--    -VRKLIKKWL   NNMAKPSD   -SLLIWES  3016
Query_MIDGE  3050  FDEKHKC----LLKAR   NIAELTNFIHLTTSTDVsmIEKK--    -VKQLLHSWK   TSKPIPSD   -SLVLWHD  3108
Query_LAODE  3026  LSHNLIM--DKCLKLQ   RLMDIESAIAVLRASKL----GTVD[ 3]NCIDLLEKWD   SNPPMKND   -SLSHWEL  3088
Query_LOTTI  3088  LIVASRT--SQLQTLQ   PLVELHEFLTYIS-NE--nDITHTT    -ANQLLKSWQ   KRQPHKLL   DPVNIWDD  3148
Query_APLYS  3167  LMVSGQT--DILRRLQ   GLVELQEFLDLVGIDG--gGGVSPS    GVRVLVDKWQ   HRWPHSLL   DPATIWDD  3229
Query_SCALL  3151  LMMSSRV--NKLQQLQ   PLVELQEFLDFIS-RE--sNFSSPG    PGNRLLEKWE   RRSPDMLL   DSVRVWDD  3212
Query_CRASS  3105  LMSVSRK--SQLQRLQ   PLVELQEFLSFMA-SE--gNFTSVT    PANSLIERWQ   ARALHPLL   DTEAVWDN  3166
Query_OCTOP  3058  LMKSRRK--YLLQNLQ   VLSEIQNFLSFFKNKE---KIKSCS    SINRLFNCWE   KQTPSLLE   DPVKVWDD  3119
Query_POMAC  3119  LMTASRT--HCLQRLQ   PLLELQEFLDFVA-DK--nNFTSFE    PSHQLIARWE   QRYPHVQL   DLVDIWDD  3180
Query_WHIPW  3018  TDLQARR-----QVIQ[3]VLAEMLAYVKLA-ASGE----ATIP    RIGALCTLWA   NSGSDLTT[2]--TPVVDD  3077
Query_TRICH  3021  CDVTARL-----KLVQ[3]PVSETAKFLHLLHASEH----LKLE    SVVRLCRFWL   RYGLDVET[2]--TTTADD  3081
Query_AASTA  2824  MPASHRL--RQLQRLP   NMVFLDDVL-----DVIktHGGGAM[ 2]----CVEKWK   PIVPLVHQ   DSLDAWST  2881
Query_AEUTE  2754  MPTAHQK--RQLERLP   NAVFLDEVL-----NTVr--NASSI[ 2]----CILKWK   SITPLVQA   DTLKTWSS  2809
Query_PLASM  2996  MATISRM--ELLQSLS   LIVEIDEALIRLNEDGSdiCKN--E[ 3]---GFVNEWK   QKPPFTGE   EGMVLWSQ  3058
Query_BREMI  2977  LASSSRL--QLLQALS   SMVQIDELFHRFGRDTFemRKTKLQ[ 3]---GFVKLWE   RFSPSMSE   DGIELWSQ  3041
Query_OLIGA  3022  IANLTRL--ELMQTLS   SAVELDDLLSISQKSVVktESTTAM[ 5]---SFVEGWS   RALPSQGE   GSMALWSH  3088
Query_LIGNO  3421  LQRQFHL--SNMASVH   RLTELRESLTLLTGDRR--RQDVVS    STADLLAKWR   QRRPDFIR   DACDSWER  3483
Query_CESTO  3170  ---KEKC----LERVQ[1]-ATELKEYLEASTSGPD----SDAT[ 1]---KILFSWQ   ERCSKIKC[1]------EE  3218
Query_OVI2_  3553  ----FRSv-GKSARAQ   TLTELSEFLNNMLNKDI-----SFN    KIQSIVDTWH   NRPPHPSL   DPLSAWND  3609
Query_PRIAP  3161  LMLASRA--ARLRSLQ   KLTEMNEFLDAITATP--dAFTRND    TARALVDRWM   DHMPHPTL   DPVSIWDD  3223
Query_ROTIF  3046  TIKQNQI--NKLQSLQ   SIVELNEFLKFINSNQA-fTVNLAK    KIDNLIDVWT   DSMPNMQT   DPPCTWDD  3109
Query_SPONG  3140  ISSSTGAvsSTVQSLQ   RLTEMNEFITFMKSND---NFSSPG    PVNNLLRKWR   KRFPDPKL   DPVPIWDD  3203
Query_TARDI  3131  LNLQGRRk--QLPMLQ[1]-MTEARRYLDVRHQIAH-----SPQ    KGRELSEQFH   HFLANSER[5]DPVGVGND  3195
Query_TRICH  3114  TLLYSRA--SKLQSLQ   TLTEMQEFLQFMDDGR---NFESVT    PVLSLISKWS   TRFPNEVI   DPINVWDD  3175
Query_OSTRE  2992  IAIRRKI----LQALQ[1]-AVEISETLRAMEEINDilCLSSTG[ 9]-VSDLDSIWR   SRWPSAVH   DPSEAWER  3062
Query_NITEN  3223  LAVGARH--AQAEQLQ   IMVEFDEFLSFVSSSSG-dQLAAVT[28]QVGDLLDTWQ   HRWPSWEH   DRVGVWDD  3314
Query_CLUBM  2927  LATGARR--LQVRILQ   RAVELEEFVDFSSRPDR-sNISR--    ----LLSKWR   QRWPSTER   DDVVAWDD  2984
Query_MOSS_  3079  FALGSRH--QRVRQLQ   KIVELEEFITFISKPEA-gELSR--    ----LLSGWH   KRWPSPNF   DDVEAWDD  3136
Query_LIVER  3118  LAIGARH--RQVQKLQ   KITELDEFVEMMTNQQS-eQLPR--    ----KLVEWR   HRWPSSKL   DSVEAWDD  3175
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3171  IILNRNVY-MDHFAE-HL--    ----TsIdAKI---------DPFGKQRVVMSLAQVDAASLQGNHSVA---RKLLKS  3226
Query_LINGU  3178  IVTNRCLY-VDQLCH-KL--    ------SsVQE----EFEEEDKFREKRLQFILAQAESAVEQHNFPVA---KKKLKI  3236
Query_LIMUL  3209  IVTNRIMF-MEKLRE-KVVG[ 9]DSSMKtLsQEElSRTTLD----IIEQEFLLKLRFVDSACKQGCIAVA---SKHLKW  3284
Query_SPIDE  3154  VIANRQLY-IQKLGD-KFGN[10]DSLNNdSiHED------DFKRIALNTSVHMKIRFAENCAAKGNFPVA---VSTLKN  3228
Query_SCORP  3173  IIMNRSLY-MEKIVS-KLQE[10]EESMDtSiVDSePNSYFDSEKFLKEQRLKMKLKFVEAALLQGNLQVA---VKHLKW  3253
Query_TICK_  3050  VISNRCFL----VKS--LGQ[ 2]-----------------------AGEKARLLNAFAGAMLQRDNIPTA---LRCVLE  3095
Query_MONOS  3155  VFLLRSLF-LGQIAL-EYSR[ 3]--------AILgDHRAEDFDLTDSTDAASQAKRIFVREQAQHNYRLA---LKAVQI  3220
Query_ROSET  3117  VTIFRRVY-FSKIIH-----[12]EMLQQvDlTDSrAPLPARIRRRLVVEDAEQCLRMAHQSSEQGPLKVA---EKYHKI  3195
Query_BELCH  3227  IIANRCLY-LDKLND-KF-E    QMAARgVlDTEvG--GEPFSHRAQREQFLLRLVLADSAREQANFPVA---LKHLKN  3294
Query_CIONA  3200  CIVYRSFF-MDQLQK-KLNI[ 6]-------------HEDIQWME---KTKFQLRLNMMESARLQGNHVVVariRSSLKK  3263
Query_CAECI  3202  IITNRCFF-LDKIRE-KLPH[ 9]DGGGDaSdRMElGGEEEDITSMIKSCKFNMKLKMVESARKQNNFSVA---MKLLKD  3281
Query_HUMAN  3182  IITNRCFF-LSKIEE-KLTP[ 9]DQDGDpSdRMEvQEQEEDISSLIRSCKFSMKMKMIDSARKQNNFSLA---MKLLKE  3261
Query_MOUSE  3183  IITNRCFF-LSKIEE-RLTA[ 9]DEDEEsIdR-EvYEPKEDVRCMLQSCRFTMKMKMIESAWKQSNFSLS---MKLLKE  3261
Query_XENOP  3203  IISNRCFF-LDKIQDvAVGH[10]DDLADgNeAMEv-DRQEDIAVMINKCRFTMKMKMVDSARKQNNFSVA---MKLLKD  3283
Query_PLATY  3190  IITNRCFF-LSKIQE-KLAG[ 9]DGGGDtTdRMElEKQEDDIYSMIESCKFSMKMKMIESARSQNNFRVA---MKLLKD  3269
Query_TASDE  3194  IITNRCFF-LNKIQE-KLNC[ 9]EGDGHfIdKVEvEKQEEDIYSMIRSCKFSMKMKMIESARKQNNFSVA---MKLLKD  3273
Query_CANAR  3191  IITNRCFF-LDKLQE-KLPS[ 9]DGEYSaGdQMEvDQQGENIHSMIKSCKFNMKMKMIESARKQNSFAVA---KKLLKD  3270
Query_OSTRI  3161  IITNRCFF-LDKLQE-KLPR[ 9]DEKDDiGdQMEaDQQGEDISSMIRSCKFNMKMKMIESARKQNSFSVA---KKLLKY  3240
Query_TURTL  3200  IITNRCFF-LGKIQE-KLPN[ 9]DGRGDvGdKMEtDMQEEDIHSMIRSCKFNMKMKMIESARKQNSFSVA---MKLLKD  3279
Query_GECKO  3201  IITNRCFF-LNKIQE-RLPN[ 9]DGGGDaNdNIEgGNEKEDIHSMIRNCKFSMKMKMIESARKQNSFSVA---IKLLKD  3280
Query_DANIO  3190  IITSRCFF-LDKILK-RL-K[ 9]--DG-----ADqG-SGEELGVLVKTCKFNMKLQMADSAWKQNNFPVA---SKLLKE  3260
Query_SNAKE  3188  IITNRCFF-LDKIQE-KFSS[ 9]D--GDaTfSMEiDNENQDTHTMIKNCKFAMKMKMIECARKQNSFSVA---LKLLKD  3265
Query_STERL  3206  IITTRCFF-LDKIAG-KLTN[ 9]DGGEDtGvKMEvD-EPEDVGSMIRGCKFSMKLMMADSARKQNNFSVA---TKLLKE  3284
Query_MILII  3201  VITNRCFF-LDKIQE-RLVS[ 9]DREGEvEtRMEvDKEEEDINMAIRSCKFNMKLKMAESGTKQNNFAVA---MKLLKE  3280
Query_BAMBO  3192  VITNRCFF-LDKIKE-RLSL[ 9]DDEG-aEtQMDiDKEEDDITMLIRSCKFNMKLKMLESARKQNNFSVS---TKLLKE  3270
Query_ALLIG  3158  IITNRCFF-LDKIQE-KLLN[ 9]VEAGDiGdEMEvDQQEDNIHSMIRSCKFNLKMKMIESARKQNSFAVA---KKLLKD  3237
Query_GAR__  3203  VITNRCFF-LDKISE-KLVS[ 9]EEDGDkEiKTEaG-EWEDPDFLIKDCKISMRLKMAESARKQRNFSVA---AKLLKA  3281
Query_TETRA  3340  ILNSREIFaLALYQKyNLED    --------------------QFVNKIRSNWYIQNVKGMLKKGIHGAAd---RKMKK  3392
Query_PARAM  3061  IYQNRLIM-ANQTHP-VLGM[ 3]-------------QFAKGVVQMNYFQQAANLLSICFKIYKCENWSTY----STLAK  3120
Query_SLEMN  3281  ILTARNLYi-DLY-KfRLKD[ 3]--------------TSLKQHKDLADIGAILQVQCAKGSYKMGLFDSS---DRFLKR  3340
Query_STENT  2759  ILFARLLFy-DTI-NqNINL[ 2]--------------------EELNDYASSLYIGTAECALKHGWLDVS---EKLLKD  2811
Query_NEMAT  3178  VITNRRIFgLTVCVY-RP--[ 5]DGEWScP------SV-DALRERITREEVALQLKMVQAAREQSNFPVA---HRYLRE  3245
Query_PISTI  3227  IVTNRVLF-MSKFLQ-KF--[ 5]E-ERSgI------CL-DDIKDKFFKQEVELKLCMANAAKEQGNYAVA---HRCLKE  3292
Query_CORAL  3200  VMTNRSVF-LDKILL-KFGQ[13]D----gL------NT-TQLEEIFLRERVLFSLRLSEVATGQVNFPVA---KKQLQS  3272
Query_HYDRA  3058  IVLSRCAF-IKKF---KSVY[ 6]-------------CQLVYLNEIPFRKVFKMVFGLTTAPCLMGNFAVA---KQYLRD  3119
Query_LATIM  3200  IITNRCFF-LDKIQE-HLCK[ 9]DGEGDaSpRMDvD-EEPNISAMVQSCKFNMKLKMAESARKQNNFSVA---TKLLKE  3278
Query_SHRIM  3160  LASYRDLY-LHLMEK-KLPE[ 3]-----------------DVSNLIKDTKNFTHRFVIRAALHNSNYHLA---NRHLKK  3216
Query_MCRAB  3153  LSSHRELY-FQFLEK-ELPQ[ 3]----------------EDIQKLMKDAKSSLHKSVIKTALMNKNYHLA---NRHLKK  3210
Query_AMPHI  3196  LASYRGIY-LHFMSK-M---[ 2]----------------REIADKVNETKANSHIGVVEVALSTGNSSLA---TQQLSR  3249
Query_BARNA  3006  LVGRRHLY-VDQLPP-----    -------------------GTRQQRYLRRSRLRLVEAALAQRNATAAg---RQLGH  3053
Query_COPEP  3011  ILTTRCLY--NEKREkKMED[ 3]--------------------NLYEKALGRCYANLALSAVYQKNALFA---QRCLKK  3064
Query_DAPHN  2844  ILALRTFFaQRKKTErNLAS[ 3]-----------------------------LRLKLATIALEQKNTELA---------  2884
Query_STARF  3243  VVSNRLLY-MDKLGS-RF-I    RSESQqA-GE----EETALEKRILERLWLELKMADSACQQGNFSVA---QKHLKA  3307
Query_STRON  3223  VTQYRCLY-MDKLCN-LL-G    DSGGGgGgGDSmETEDQKFDDLVRREKVLLSMKMADSACQQANFCVA---AKHLKS  3292
Query_SEACU  3207  IVTNRCLF-MEKFAT-RF-D    VADTQgSdSE----VKVEFEEYISMEKTRLSLKLAESASEQANFPVA---LKHLKS  3272
Query_FUNGI  3083  IVTNRGIY-MEKIHE-RQS-[ 1]TVRGQ---NEInEHKLSELKEILIDKRSKVYREIADAARKQNNLYVS---EAYLRN  3150
Query_DPURP  3204  LVFYRSVL-LEKIYE-RFST[ 3]DKSSE-----------DRVKSKLIQERAILYHKMSKGARKLGNIVVS---ETYFRL  3266
Query_HALBU  3240  LVSFRSVL-LEKIYE-RFI-    ---GQ---NNNdEAKEDGVKNLLIRERAELYHQMSKGARKLGNIIVS---EAYFRL  3303
Query_GLOBO  2978  VISGRKLM-IEKVSA-MLN-    NPPDQ---ENNvYKFHQVIVSDDLLYSRR-M---AKAAAKQGLFSVA---DRWLDT  3040
Query_PALUS  3050  IMLTREVT-AEKWAE-VGF-    SPEEK---S---------IVNSNLQDTRRIM---AKAARKHQRFYMA---ERYL-T  3103
Query_CALIF  3134  ICENRETI-LNQIKK-AIE-    NDSNI---NNNdYYN-----KKLIEHKKNIWLTCSKAALEMKNFFVV---ATCLSK  3195
Query_GLOMU  3183  IVTTRQSI-LKEMAE-RIK-    LNRFN---R--------MKV-SLESQSLLKMSSTARIQRNFEVARVC---LNKIRN  3240
Query_LOBOS  3137  VLDDRKLM-IQRLED-YTA-    YSRRQ---E--------TLVTNHRIRSFMKM---SAAARNQGNFYVA---NTCVAS  3192
Query_RACEM  3123  IIKSRNMF-IEELKG-VID-    QKDLP---N--------PVMLAMDECSKSNILAEISAAKQQNNFTLA---YNTISS  3181
Query_PVIOL  3220  IVSFRSVL-LEKIYE-RFTT[ 3]NIESK---NQDqESIVNKVKSTLIQERAELYHKMSKGARKLGNVVVS---ESYFRL  3290
Query_BOMBU  2958  LTSYRTYF-INKLLD-DKLK[ 3]ESLIQnDsDSNpDENENSITYRLRTIIFDMRLKMVEASINQKNKYIA---KNYLRQ  3031
Query_AEDES  3110  LLVYRKFV-LNKLET-LLEA[ 2]DD----------RLHQIDFVKQLTDAVFETELRLLDASFVQSNYQFA---RKIIKR  3172
Query_ANT__  3033  LAAYRAYS-SMLFYD-SCQK[ 3]KR-------------------RLFQINYQLRLDIVDAALNQKQRYIA---ERHLNC  3087
Query_WHITE  3106  RLLYRCLF-LNQVWK-ELG-[ 1]DL----------ETGLCKSLKDSR---LHL----IDITLKGKNSEFA---KQQLRL  3159
Query_BEETL  2938  RVLYRKQF-LRTLVR-KIES[ 3]DD----------FANTLTQLNDMS---LKLDEDILEAAIGENNFYMT---KKYLNL  2998
Query_THRIP  3002  RLTYRFAC-MKYMPT-EME-[ 1]EV----------EKS---AL-RMQ---LSL----AEAALQQGNQNFA---KKYIVD  3051
Query_TERMI  3107  RAMYRKKF-TNRIDQ-MLE-[ 1]TG----------ASN---ALSRTE---LAM----VDSALYQENFYIA---RKYLRE  3157
Query_FLEA_  3017  KLAYRLMF-SNHIKK-QLAT[ 2]QD----------FAFINEELDNCK---IKMELSFLDAAIKQKNLGLS---RIYIRK  3076
Query_MIDGE  3109  LLAYRRNF-HRILEE-NYEK[ 2]NV----------TEDSFDYLLNI-------QRNISNVAFAQKNCDAA---KFIING  3164
Query_LAODE  3089  RTKYRAHF-CGAVMK-KLKN    --------------------PEIASFENRYYMSLLDLALNHENAHLA---KKVLRK  3139
Query_LOTTI  3149  VISQRNVY-LDHISQ-CL--    --AGKnKeSQE----MD-EGDMFHQSKLRLYLAMAESCKEQNNIALT---LRILKK  3210
Query_APLYS  3230  VIVNRTVY-LTQMSE-RL--    RKGQEeEnSEEpTWREAAVAERLEEMKLWMLLKMAESCRVQNNVTLS---LQLLRN  3298
Query_SCALL  3213  IVTNRNVF-LDHIGA-RL-M    AREQNsEdSME----VD-MGDTFLESKLQLHLKMAESCKDQNNYKLA---LQILGE  3277
Query_CRASS  3167  VVTNRMVF-LDHISN-RLST    SLVKKeEdAME----ED-EEDLFLASKIHLRLKMADSSCHKRNHKLA---LNILAD  3232
Query_OCTOP  3120  VITNRLTY-IDQI-------[ 9]SDAMDtN----------DGNYFPKEFKLTLQLLRAESCCEQNNFKLA---RTILKK  3183
Query_POMAC  3181  VVTNRNVY-MDKIRD-QM--    ----LsSnLME----EDHDSSIFEDIKLRQQLKLAHSCQEQNNFNLT---LRILKE  3241
Query_WHIPW  3078  LVCNRNLF-IEQADD-----    -LAGGkTaMAK------EIYPSLAHFSVRYRFQYAKNALHRMNYDTA------LRQ  3134
Query_TRICH  3082  VVLNRKTF-LEMFGK-----    -GLVDsGiSVS------DVRQSLLPFRLQFGYQYLKCAVRQNNHGVL------LQR  3138
Query_AASTA  2882  YHLIRNVG-YESLVA-----    -------aTDEsKYGGRSIAQDMASIRVHTMLSYAEAAVASNVLALA---GRLLAD  2941
Query_AEUTE  2810  YYLIRNLG-YDVLTA-----    -------hSQGqAKGDDSIIREVMAARVHAMLSYAEAAVESNVLALA---GRLLTE  2869
Query_PLASM  3059  HTMVQNAI-TSLMLN-----    -------rAHDqGVLSNEIRLVVLGIKGHALLQYANAAVSCNILSLA---SKVLKI  3118
Query_BREMI  3042  HTMVQNTI-VDFLLK-----    -------tASDqGMLSDEIQLSLRRGMSNTILQYASAAIICNIPALA---SKMLKS  3101
Query_OLIGA  3089  FYLVQQVV-RSFLLH-----    -------sGSEdGLISDTVERVFNQEHAQLMLKYAQAAVANDLLAVA---SKYLKD  3148
Query_LIGNO  3484  TVAVRLLC-ANKMLL-----[ 1]DGLDDlAaGKR----------RFNAFAGEQLLLLAEAACLQSQPAVAli---HLGR  3541
Query_CESTO  3219  VINSRLFY-LSRLVE-----    -------------SGSLSLEEQFLPACIDFRIKCCNDFSSSGNAKRA---LNILPQ  3272
Query_OVI2_  3610  IILNRLFY-LTIMST-KCGP[ 3]------AsSALgERGEAVNPHFLDRTIFSLRMAEVNACITQNDPVLA---LNHLKN  3677
Query_PRIAP  3224  TVINRCLY-MDHLSR-HF-H    GNQPAgSdPF-------------LEQKFNLFLAYQQSAMQQRNPQVA---LKHLRT  3280
Query_ROTIF  3110  VITNRSIF-FEFIEE-KYLN[14]GDSVMnDfDEDlSNYSKKLKKKMTKTTFLMKIRFAQAAQSQGNFKLS---LDKLKQ  3194
Query_SPONG  3204  IITTRWMM-LRKIST-KLQD[ 5]----QcEyLLDdDGSSCSVDSMMKNENINMFMRAADAARTQSNYPVA---ENYLKI  3275
Query_TARDI  3196  LATLRCYY--SHVLAsQPKN[ 3]-------------------EALMASFRVHEKLAVARTAIRQNNFDLS---QRLLTK  3250
Query_TRICH  3176  VIVNRCMY-LRKILL-QYRV[14]DGEDRvI------YTFPALKKRIAEEEVSLYLKLADTARLQSNYPVA---IKHLRQ  3254
Query_OSTRE  3063  VVTFRSRM-MSTIVG-VIPE[ 3]--------------------TSIEALRTGMWLCASDGLRQVGQVALS---RKFLRQ  3116
Query_NITEN  3315  VVRHRALC-LERAKG-HMGM[ 3]AGHVA-------NYDGTKLAADLAASSSSYFKQAAAGLRKRGEYDVA---QDYLKE  3381
Query_CLUBM  2985  ILQNRLVF-LQKLQG-LFNT[ 3]AKA---------DQARQNFTKQLDEERAFGFLQASRGVMKQGAYCVA---QKYMDE  3049
Query_MOSS_  3137  IAQNRLLF-LKYIQE-HFNN[ 3]RNSF-------dDKFCNSISSLIDAERAHLHVAVANGLMECGALPVA---QSYINR  3203
Query_LIVER  3176  IVKNRLVF-FHKIQG-LVNR[ 3]TGSVLlGaTDTrDKKSEALASKLDQERAHIFHQAAVGLRKVGAYNAA---QSYMEE  3249
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3227  AKDLVKVVDSEE    LSVLFTHS--HANANK--RKCAS--[1]SEPSERL   NLLVEVKKR-LDKC--SESSTLKLNA  3288
Query_LINGU  3237  IESEVKSLSESP    VLIQFYHC--YAKLHN--K--ITEV[1]MDASATL   SPVLGALGL-LERK--RDHKLVKESN  3298
Query_LIMUL  3285  LNKKAEHLGNSD    HRTTWIHK--YCEIYH--QQAVM--   LRGAEKL   KILLRAKEQ-LMKL--DDTVARSNIS  3345
Query_SPIDE  3229  IYRDSRSLDGST    -WSKFIEA--YCSVNN--QRVLQ--   SHSEKSL   DIAVGSYKQ-LGEL--EKRCGSG---  3285
Query_SCORP  3254  SQQIAKELKADH    HWINCIHL--YCKIDL--KKAKN--   QNYPEQL   DTLLCAHKQ-LAKL--DLHFPKK-LD  3313
Query_TICK_  3096  IEKMRTEGQLEG    -SLEWSYAEtYCSALC--RAAEQV-   -PLDKKL   ASHLRTLQK-LEAT----RAEGLDDA  3155
Query_MONOS  3221  NNPKQSSQYFRL[ 9]FSRDDMRDLlLLKSQA--RAAHNNL[5]MDRGSSP[4]QSILRCLDA-----ysDQMRFTSGNS  3301
Query_ROSET  3196  VHHILNTYTFDD[ 5]LNLQQAHGLvAYHLHEf-----NLG[2]----EQK   -HYFRAFARdLDRLnkCSADLPPSKS  3262
Query_BELCH  3295  TLKDFT--EDDP    LRLRWTHC--YARLHQ--KKAS---[1]LTGADKL   NNVLTTFEQ-LDKF--SSCRVFEEDL  3353
Query_CIONA  3264  EHELLNSDKN--    LSFKLTHEAaLSYLAQ--AESLALN[1]KPNNNKI   KYFIDSVEC-LNEQ---NEFITESAT  3326
Query_CAECI  3282  LHRESKTRDDW-    -SVRWIHS--YCHYTH--SRSQS--   QTSPDQV   LSVIKTVSL-L-EE--TKADYLRENT  3339
Query_HUMAN  3262  LHKESKTRDDW-    -LVSWVQS--YCRLSH--CRSRS--   QGCSEQV   LTVLKTVSL-LDEN--NVSSYLSKNI  3320
Query_MOUSE  3262  MHKESKTREIW-    -RVQWLHS--YSQLNH--CRSHT--   QSPREQV   LNTLKTITL-LDES--DISNYLNKNI  3320
Query_XENOP  3284  LHRESKTNEDW-    -SVKWIHS--YSRYSH--SRSRD--   LTCSEQI   LTALKTIPL-L-EE--SKTEYLTKNT  3341
Query_PLATY  3270  LHRESKTKDDW-    -LVRWNHN--YCRCSH--SQSQS--   QSSPEKI   LTVLKTVSL-L-DD--NLSSYLSKNI  3327
Query_TASDE  3274  LHRESKTREDW-    -LVKWNHS--YCRFSH--SRCQI--   QSSSEQI   LTVLKTVSL-LADE--NLSSYQSRNS  3332
Query_CANAR  3271  LRKEARWREDW-    -LVRWNYA--YCRFTH--SSSRN--   LSCPERV   LSVLKTISL-L-ED--TKSDYLSKNI  3328
Query_OSTRI  3241  LHREAKTREDW-    -LVRWNYA--YCRFTH--SCSRN--   QSCPEQV   CSVLKTISL-L-ED--TKSDYLSKKV  3298
Query_TURTL  3280  LHRESKTREDW-    -LLRWNHS--YCRFSH--SRSQS--   QNSPEQM   LTVLKTISL-L-DE--SKSDYLSKNI  3337
Query_GECKO  3281  LHRDSKTREDW-    -LVRWNHS--YCRFSH--SRSRS--   QSNPEQI   LTVLKTVSL-L-DE--ASSGTLSKDR  3338
Query_DANIO  3261  LHRHAKIDDAR-    -LLRWVHS--FSRFTH--KRIAR--   LGPSEKI   NALLKTVPL-L-KD--AERQSEALSA  3318
Query_SNAKE  3266  LHGDSKTCEDW-    -RVKWNHS--YCCFSH--SRSRS--   QSNAEQI   LTVLKTVFL-L-DE--IKLSSVSNNA  3323
Query_STERL  3285  LHRESKNQEDW-    -LVRWVHS--FSRFSH--TRSHT--   QSSSEQI   LTVLKTIPL-L-EE--RKIDYTKATT  3342
Query_MILII  3281  LHKESKARDDW-    -AVKWTHT--YSRLSH--RRSQR--   QARPEQV   VTVLKSIPL-L-VE--INVDYLNTHT  3338
Query_BAMBO  3271  LHKESNIRIDW-    -RVRWIQS--YSRFSH--KRSRS--   MKRAEQI   VTVMKTIPL-L-GE--IDVGNLNIQR  3328
Query_ALLIG  3238  LKRESKTRDSW-    -LVRWNHS--YCHVSH--SSCQT--   QSSPQQI   LTVLKTVSL-L-DE--NVPDHLSKNI  3295
Query_GAR__  3282  LHKECKARSHW-    -NLRWAHG--FCRFKH--LYSQA--   QDSAEQI   LNVLKTVDV-L-ED--SKTDPEKETA  3339
Query_TETRA  3393  AAGYRKVMNGLP[ 3]----WDYYKiTSKLKLkqLEVDLKN[4]ENPFEKF   SKLIKIFEQ-------EKEKYIHQLD  3458
Query_PARAM  3121  LKLKEERFESDQ[ 3]----------ETIYQK-----LNDN[1]----DRY[3]----KKLKLfFSQQa-------QKNT  3167
Query_SLEMN  3341  ALKIRADED---[ 4]-NMKIVAPI--IKLKCeqFKVDTMN[1]-----NF[2]KG--DKLQK-IVKVfdS-KMQESFGE  3402
Query_STENT  2812  AINKRKDKT---[ 3]--VSLLSSV--IRLKAktVIRNSEE[1]-----GF   -D--LAIQH-FQRLyeSIKQYKSNNE  2869
Query_NEMAT  3246  SLKSI---EDN-    LHVKWTHA--YAVTQQ--KKAHG--   LMAIDAV   QTALSTVSK-LGKP--SINTLAN-DP  3301
Query_PISTI  3293  ALYKIPQGNEN-    LNISWSHT--YAEVHF--RKIQT--   LTAVGAV   ETALLVAGQ-LDKF--SASKILEIEP  3352
Query_CORAL  3273  SLHLIRDRLKND[ 2]LEVLWTHS--YGDLNC--KKCPN--   LLPSEAV   DTAISALDQ-LDKI--KDSKVLQENA  3335
Query_HYDRA  3120  SHDVILKEVNNV[ 2]VGLHWRHLYaKVYCKT--AKKFQND[1]GANFTAL   VSVIGHLEK-CQSF---DKTIATNLL  3186
Query_LATIM  3279  LHRESKSREEC-    -LTRWVHS--YSRFGH--SRSQS--   QSCPERI   QTVMKTISL-L-EE--SKADYLNKNT  3336
Query_SHRIM  3217  ISSLSED----E[ 1]SKAQFYFLM--AETHI--LRGQG--   KQ-TDRL   KYLVEAWAK-YLGRv-SSMSILDQNP  3274
Query_MCRAB  3211  LRPLCGD----D[ 1]ELAKFYFLM--TETNI--QRGRS--   RP-QNRL   QYLVEAWVK-FLGKv-TSMPVLEQDA  3268
Query_AMPHI  3250  VDPLVKKGLLTE[ 1]LTVKYHFLC--ANTSI--IIGQN--   KRGSERI   KSLVEAWCR-QLGEv-EKLQERFTSP  3312
Query_BARNA  3054  LRSQQWDADDAS[ 3]------------EYQL--LRCQTAV[2]--GRDPV   -RTWRVLDD--------QVRWRDSAP  3103
Query_COPEP  3065  LKPLMDGDVELA[ 3]-------SXiISKVAI--LNFTTRS[1]LTAVDNF[2]--LFKSMNS----------SQLSEEN  3119
Query_DAPHN  2885  -RKLLEGATDLR[ 4]-----FHYLiQSKVLA--MESSLDS[3]ASNLDRL   ILAKETINR-----avAASTGVPKDG  2948
Query_STARF  3308  THGSLS--AHED    LKPKWTHS--YVQMNQ--QKIL---[1]IELKEQV   MTLLTTFDP-LAKL--SNTPALTGDY  3366
Query_STRON  3293  THRGLV--EFED    LRPAWTHS--YVSMTQ--RKAL---[1]LKPYERI   TSLLKALDH-LGKL--PVSHNLGSDL  3351
Query_SEACU  3273  TYGHLG--DHED    LKLLWTHT--YVKMHQ--RKAS---[1]LTAAEKV   TTLLTTFDP-MSKL--ENTESFKTNK  3331
Query_FUNGI  3151  ALKSKSK---KK[ 3]--FDCFRSL--VKLHN--KKALDSK   -SLIETA   DKLTKALK--FVQNk-EEEEIILNNK  3210
Query_DPURP  3267  SVKSYPK---TK[ 3]LAFPLVSSL--LDIYC--IKARTSA   -SPIETL   DRFVKALK--FIESk-KNEESIVSIN  3328
Query_HALBU  3304  AVKSYPK---TR[ 3]LAFPLVSSL--IKIYC--LKAKNST   -TQVETL   DRFIKALK--FVESk-SDEDSIKNNP  3365
Query_GLOBO  3041  ANSRPGE-----    FVPEFMQQY--YKQTL--LKSDREL   -GFANRF   KPLE-------NLLk-HMDYYKKSIR  3092
Query_PALUS  3104  QSVQNG------    FDDIFNYNA--LKLSF--VRM-KQL   -QGQEMF   EIAG-------KAMq-NFAFYKTGIE  3153
Query_CALIF  3196  SKNYGL------    SKMEFSYEA--IKYIV--TELKTLK   -DSNERL   KKIV-SLG-----Is-LSSLYKKLEK  3247
Query_GLOMU  3241  LRLIDD------    --ERLSYCQ--FKLNL--REAQNTI   -DTEKKA   EILL-------ALSe-EFKKVKSS--  3287
Query_LOBOS  3193  MQKLQA------    SSYNIHYSQ--LKLDL--ARATLEM   -DRVVKA   GILA-------DSLs-KFEEYISNDR  3243
Query_RACEM  3182  FSELEV------    TKAEKQLVY--LQLHL--QWVLSAE   -DSSRAS   QLLGRGIGLtYEQEe-PLRPLLASSD  3239
Query_PVIOL  3291  SVKSYPK---TK[ 3]LAFPLVSSL--LNIYC--MKAKNSP   -SPMETL   DRFVKALK--FIDSk-KDDETIASSK  3352
Query_BOMBU  3032  VEQSV--EAPSI[ 3]---EFLLFC--GKLKY----LMGDT[1]TDMKKKL[3]SSSWKHCHD-LLRQns--IVDTTNVS  3094
Query_AEDES  3173  L-DVVAEEQSER    ------------GYRW---RIAHLK[1]RRLGE--   --LKSA-----------GEPAVSYQK  3212
Query_ANT__  3088  VRQTV--CE---[ 1]---ELLWLD--ARIKR----LHADV[1]TNMKKKM[3]TDSWETLHK-LLKSnk--LDTISSTA  3145
Query_WHITE  3160  VKQHIPESVNTL[ 3]---KYHLKC--AQFAE---------   -NENEKA[3]NQSWNALDS-MRKMai--EDDLQCMR  3217
Query_BEETL  2999  SNNRN-------[ 2]---KTQLIL--SKLAY----IKSLM[1]NNSEQKL[3]LQGIKTLVP-IF-Eet--ESFLQLSA  3054
Query_THRIP  3052  MKGKIQSSSSPT[ 3]---RWLIAR--SSERHqwAKREQNS[1]TRLDNLL[3]RQLDTVLAS-DD-Ldd--HPDLSVIA  3119
Query_TERMI  3158  AREKM-SNSGTD[ 3]---QCSLAY--SKAK--wLQAQLKN[1]T--DEKL[3]LDSWSVIED-QR-Feq--YPMLHAET  3220
Query_FLEA_  3077  LKDLGKHEDPEK    ------------NYRW---RISRCQ[1]IFLQGIL   ------------------ETDIDKKV  3114
Query_MIDGE  3165  LKDEIRKKSSNE[ 3]---KYKLAV--GKYSM---MVADQK[1]TSVDEKM[4]NGMKSLVNG-IIKAeiEGEVEVPQIK  3233
Query_LAODE  3140  CNKNEPMVIAGK[ 3]---ELLLAR----MYE--TRDLNEM[1]----KKI   ADSWKTIDQ------vNRLPKLDSQV  3194
Query_LOTTI  3211  TNSAIKSVDNEE    LKVEWSQL--YASTHQ--KKIIS--[1]DWSEDGI   NGLLTTI----GKF--NDSDVLKSKP  3269
Query_APLYS  3299  IRTSCKRASNKQ    PLLEWVHL--YAKTHH--TKAGNTP[1]PWADDTF   GNVLSTLDE-LDKM--KGSPALYERP  3362
Query_SCALL  3278  NKSQCKSSGSVG    LWVKWGHL--YTSTHH--KRAQSVQ[1]SWTDDTF   TSITSTVDL-LGKN--STSDTLEETP  3341
Query_CRASS  3233  IHKNYRDKENEI    LNLEWSHL--YAKTHQ--LSAMSAD[1]DWTNDIF   SSVITTIDR-LGKF--KDSSLFEREA  3296
Query_OCTOP  3184  ISGDCKHISLMK    LKLKWNHL--YVQSCQ--KNTLL--   SN--DTC   INLVNNLKI-LCEH--DLSLLPPEEQ  3242
Query_POMAC  3242  TYSKCKREKSEL    LVVEWSHL--YATTHQ--KKALSCS[1]IWSKDML   YNVLSTLDP-LSKV--AESKMLREKP  3305
Query_WHIPW  3135  MEEARKISIDDD[ 7]-----NHEA-YSICQM--VQSASV-   --HSDKK   VQVFRRAVE-LFENckISHETFSRYG  3199
Query_TRICH  3139  LQ--------DK[ 7]-----NYET-WALLQL--ANCANA-   --VKSRT   AYLMQ-ALN-LFDLkrISDPEVNVTE  3194
Query_AASTA  2942  YRTLCNQANLPK[ 3]HMVQVYSAQ-VSKLAS--RQLQAAI[1]LTDDH--[1]QRVLAQVAN-----yyAALTRLFDND  3006
Query_AEUTE  2870  YRTLCDENSLPK[ 3]HMVKVYASQ-VSKLAS--QKLQAAS[1]LHGND--[1]ARVISQVTN-----yyAALTRLFEND  2934
Query_PLASM  3119  YRELCNTHQLPK[ 3]RMIEVFVSH-VLKLID--RQEHKSS[1]STS----   ---IKLITK-----yyETATKLFNNV  3177
Query_BREMI  3102  YRELCNTHQLPK[ 3]QMIEVFVAH-MLKLVD--RQGYQSS[1]PSS----   ---NTIITR-----yfNTATKTFDNA  3160
Query_OLIGA  3149  YRELCNAENLPK[ 3]LMVDVFVSH-VQKLAE--RQAQKSS[1]MASHGGL[1]SDAINTITR-----yyQTSTKMFDNT  3215
Query_LIGNO  3542  ARKYCTQTSATS[ 6]TSVAWRAAC--AKAWT--TAALSL-   --DTDRN   FDAMSVAEK-VAKV------------  3597
Query_CESTO  3273  LHQLIGETGNDC[ 3]-RVAWCSAF--SRAWK-aEYLKNVV[5]EKSSVDL[4]EQCLSSLSQcIDVYdlLKEPSSIKSL  3350
Query_OVI2_  3678  IRMLAQRGNAVE[11]LWLRWCEAFaDAWVVW-sDNAVTVN[5]KSVESQS[4]VCAIEDLVKgLANAglHLVNCSTNRP  3766
Query_PRIAP  3281  TNQLLKKLDSDS    LAVQWTHS--YVRTWQ--QIATGH-[1]RDSADRL   HRLLDTIKE-LDKH--AGKAVLHQQR  3343
Query_ROTIF  3195  SRSILKIKNSDF[ 2]LQLIWNHC--YLNTHL--NKCKQAN[1]KESLESF   FESMVLREI-LKYDsiSDTFKNNQKN  3262
Query_SPONG  3276  VEKSILKDFPDN[ 2]IQLRWLHS--LIKVRL--DRIHQDP[1]PTDMEAV   ISLTCQMDN-FKGRlgGSGKGAISSK  3343
Query_TARDI  3251  LHPVACKAHQGA[11]LEDRWGGTFaAAAIAK--SRSNKT-[1]QQAQSNF   AGLMQSLGL-IDKLh-GKSTAANAVP  3327
Query_TRICH  3255  TLYHINQELDKD[ 2]LTVKWTHI--YTKVHC--DKSLL--   TPPDERL   KTLLTANSQ-FEKV--KDLSVILENT  3317
Query_OSTRE  3117  YIESVQRTEASE[ 4]-----YEALfNLKLAT---------   -GTSNSI[1]----KALEM--------CRNQLHEEK  3164
Query_NITEN  3382  YDAGRQSLGNVS[ 3]FDWDFLRLA--AKISS--LRAERRT[1]VNPLEAA   GQLSQVMR--ELSDk-AAG--AAGVL  3446
Query_CLUBM  3050  YVAVAE--KTSF[ 3]V----FKRL--VKPRC--LQALRTS[1]RDPKSAA   AMLEQVLD--YVRKq-ASKPSISGCD  3110
Query_MOSS_  3204  YMSSKF--EFSA[ 3]LDFRNFKAL--VRLQC--LRAERDE   -SANNRI   KKLQQLLV--YIETe-RRK--LSKHP  3264
Query_LIVER  3250  HMAAHQ--GTVK[ 3]LDFKFFKGL--VKLRC--LQADRVK[1]IEASEAT   EMLQQVLN--YVKSt-AQKPWVESSH  3314
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3289  A   LLCDQHLLQG-EH--LVTMAT--IL-   QNQV--KDIDALDQKLL-EKITES--VGSPAL-SHQE    -IVASL  3345
Query_LINGU  3299  V   LTRRHLVLQG-NI--FDSVTS--SL-   LS-L--PDFSSVSPRDI-QKLVEA--ASCSMQ-NADM[ 1]EIINGL  3356
Query_LIMUL  3346  E   AKTHHSYLHS-FN--AEMIAE--SLI   -EDD-vSVVSELDYEDQ-NRLKSI--A----G-STDS[ 1]RFIPKF  3401
Query_SPIDE  3286  -   --VSQKKLLS-ET--LELIAR--MFM   -K-D-kSLFDTITNENM-AVISSL--S---NG-KSDL[ 1]EIVSEI  3338
Query_SCORP  3314  V   STMNHYLLQS-EI--IEMAAD--IVK   -YND-vHAEKIQELTDT-FSIQS-------------K[ 1]KLEIDL  3363
Query_TICK_  3156  A   STVRYCLLKA-QA--VAGLAKd--L-   --DAg----AVLNKEQM--AELQL---ASTSS-DPRK    -MSLEL  3206
Query_MONOS  3302  E[1]FSRTSLIMDGaGT--LVDVARalHLL   --ERpqPGTPDLDEQQTaMLLEAAhnAGLKLH--ADA[ 2]-LAGQM  3367
Query_ROSET  3263  G[2]SLYKAEVLLS--------LGKivPLC[1]----------ATDENKV---LWDHiiAEVDAP---DA[ 4]---EKC  3312
Query_BELCH  3354  A   VAKQHHTLTS-LS--YEIVAH--AL-   IQEG-gAILTGLKDTRHaQLLTAA--GVSPDT-SVDQ    -TISSL  3412
Query_CIONA  3327  N   IERIHHLLAL-AT--THHLTSs-NL-   NNGE--------DADIE--QVKEI--FGTDIN-ADNI[ 1]---QTL  3376
Query_CAECI  3340  R   EFRNQNLLLG-TT--YSIIAN--AL-   CKDP--RCLQRIDKNKA-DKVLQL--SGTMSE-NPD-    EVAASL  3396
Query_HUMAN  3321  L   AFRDQNILLG-TT--YRIIAN--AL-   SSEP--ACLAEIEEDKA-RRILEL--SGSSSE-DSE-    KVIAGL  3377
Query_MOUSE  3321  Q   ASCDQSILLG-TT--CRIMAD--AL-   SREP--ACLSDLEENKV-NSILTL--SGSNAE-NTE-    TVITGL  3377
Query_XENOP  3342  K   ACRYQNMLLG-DT--YRIMAD--AV-   CKEP--DCLYKIEDGKA-GKVKDL------SE-SPE-    NVVGGL  3394
Query_PLATY  3328  V   ASRNQKLLLG-TT--YRIMAN--AL-   SKEP--ASLEQIERNKA-RKVMEL--SGASFD-NVQ-    KVIAGL  3384
Query_TASDE  3333  L   AFRNQNILLG-IT--YNIIAN--AL-   SNEP--TCLAQIEESKA-KKVLEL--SGASSE-DAE-    MVIGGL  3389
Query_CANAR  3329  M   AFRNQNLLLG-TT--YQIMAN--AL-   SQDP--RCLEQIEKEKA-RKISLI--SGESLE-NPK-    KVLAGL  3385
Query_OSTRI  3299  T   AFRNQNLLLG-TT--YRIMAD--AL-   SHDP--RCLEQIEEEKA-RKVSVL--SGESLE-SPE-    KVLAGL  3355
Query_TURTL  3338  V   AFRNQNLLLG-TT--YHILAS--AL-   SKDP--GCLEKIEEEKA-RKVVML--SGERPE-NSE-    KVIAGL  3394
Query_GECKO  3339  L   AFRNQNLLLG-TT--FHIMAN--AL-   SKDP--RCLEQIVEEKA-RKVLVL--SGEKTD-DIG-    KVIAGL  3395
Query_DANIO  3319  R   MLRDQRILLG-TT--YDLMAG--AA-   DRSP--FALETLGEEKV-QKILQL--SQASSI---A-    QVVEGL  3373
Query_SNAKE  3324  I   TFRNQNLLLG-TT--FHIMAN--AL-   SKER--SCLNQIGEEKA-RKVFML--SEERSD-NVE-    KVIAGL  3380
Query_STERL  3343  R   IVRDQNILLG-TT--YHIMAN--AL-   TRDG--LALEHIGEDKA-GKVLDL--LGTTSD-NLE-    QVVVGL  3399
Query_MILII  3339  R   ILRDQNTLLG-SS--YCILAD--AL-   NKEP--TCLQRIKEDQA-ERVIEH--SGSSSE-EIQ-    EVVIGL  3395
Query_BAMBO  3329  H   AFRDQNTLLG-TT--YYIMAN--AL-   NKDP--SCLQYIEEERA-KNVKKL--SGATSD-NPV-    QVSIGL  3385
Query_ALLIG  3296  E   ILRNQNLLLG-TT--YRIMAV--AL-   SKDP--RCLVELEEDKV-RKVSML--SGEKPD-SPE-    KVIASL  3352
Query_GAR__  3340  Q   NFRNQKLLLG-TT--YHIMAS--AV-   SRDP--SVLEKIGPEKA-GRVLKL--SGASSS-NPQ-    QVTAGF  3396
Query_TETRA  3459  Q[4]KFEQLNIQVNeKFi---DVYQdyTLT   ------QIVEDIFLEKYaDVINNI-yWGYNKL-----    -FQNDI  3516
Query_PARAM  3168  E[2]VQFRFNMLPI--------ISNqhYLK[1]----nqDLVFMFNQESR------------KHNnRYDF[ 3]------  3213
Query_SLEMN  3403  D   ---QQF-----QE---DLRTEikCIL   -------LKQKLEKKML--------------------[11]EI----  3441
Query_STENT  2870  D   --------------------------   -------------KDAI--------------------[10]ET----  2886
Query_NEMAT  3302  V   IALRHHILRS-RC--FDILTD--AI-   AGAG----------RST-LATYKLicLINQPV-DKSY[ 7]QIVSQL  3360
Query_PISTI  3353  L   RGIQHHILRS-KT--FDILCK--AM-   LDGG--VVIATLESQAK-NALVKL-cRIDETE-SQK-    KVVSDL  3410
Query_CORAL  3336  S   TGIRHHLLKS-RN--LDAITN--VL-   SIPG--TW-DSLEKEAK-EKLHGL-cSSKNPE-QLD-    KILGHL  3392
Query_HYDRA  3187  E   FDGWRCLL----------LCNs-RL-   QNDE----FKLITCEIK--KLATL--SQFEGD-KLEF[ 1]---KTI  3233
Query_LATIM  3337  R   ALRDQNILLG-TT--YNIIAN--AL-   NKTP--TSLQDIGEEKA-RKIQDL--SGAQSD-HPE-    QIVISL  3393
Query_SHRIM  3275  R   VEIQYLKLES------HLCLE---IC   --RAiqDMGDDWTEDNQ--YMQVL---SKKFE-NA-N[ 3]SWQREL  3329
Query_MCRAB  3269  T   IETEYLKLES------SLCVE---IS   --EAmkDMGNSWQEDDK--YIKSL---AEKFP-NA-K[ 3]SWYKEL  3323
Query_AMPHI  3313  F   VNMEYLCVES------KLCVE---LC   --NAiqSLGGNFDAQNA--HLALV---AKKFP-EVQN[ 3]SWYQSL  3368
Query_BARNA  3104  P[4]LLLRRARLEA-------ELAErcLLE[1]-AAVadPFVSQLLEESG-----HWtgDGARAD-----    -----L  3155
Query_COPEP  3120  R[4]WNIHQRDINE---clFTIVDKnsQLT[1]-----sILRRQMKTDHDlMWFSSKygTVSDDN-----[ 1]ALRSVL  3182
Query_DAPHN  2949  R[4]LERFEASLCA-QV--FQLVDN---IQ[1]---------KELTSRHA--KLVEL--------wNVQP[ 2]-----I  2995
Query_STARF  3367  C   LTRHHHTLTS-KS--YDLVAI--AI-   QKGG-aDILTGLRNVGKlDKLLGL--VGQNAT-RPEQ    -ICSTL  3425
Query_STRON  3352  F   LLCDHRILTS-KT--YQSMAS--AL-   QEEK-gEMLTRLSDAGKlDKLLSK--AECQES-KPEK    -VCQSL  3410
Query_SEACU  3332  N   NFQSHQILSS-TS--FEMVAD--AL-   NSDTeaNMLGKLRETGKfDKLKML--AGGQSS-SISA    -ISSGL  3391
Query_FUNGI  3211  S   YQSMFTTLQA-NV--LAEVSR---LS   -LT-eeNVVREAIDKNQ---W--------QWK-PTGE[ 1]-ISNEL  3260
Query_DPURP  3329  E   NLQKYLMMHG-DI--FWDIYQ---LD   -KKLgsKFVNESFKKNE---LPS----HLTSI-PI-N[ 1]-LKDEL  3382
Query_HALBU  3366  D   NLQRYKRLHG-NI--QWEIYQ---LD   -QKLgsNLVMENLKKNG---L-T----SLSSL-PT-T[ 1]-LSSEL  3418
Query_GLOBO  3093  E[4]VQLKFITLET-RT--LSRLLS---LT[4]ANDS---FFSDLGSNKG---VTKF--VGQRI--QSKP[ 1]-LMQFF  3154
Query_PALUS  3154  K[4]QRAKFLSLEG-RL--YDTIVQ---AV[2]EKDLaeLFLSNKHLKKT---IKSL--GFEQP--HDGT[ 1]-LLRIL  3216
Query_CALIF  3248  D[4]LLAKLYQLDS-------NLSS---II[3]KKSMhqSLLSECQNNSK---FSKF--IPENFT-GDTN[ 1]-LIYYF  3308
Query_GLOMU  3288  -[3]DDIKMRIIES-ET--YGLLV------   -----------SELDKR---VPNR--NLLKIK-HNKN[ 1]-FI---  3330
Query_LOBOS  3244  D[4]QIAGLNMLGS-KA--YSLLRE---LI[2]DRTIw------SHLKSN---VWAQ--SLISSK-AGEA[ 1]-IFRKL  3301
Query_RACEM  3240  Y[4]FLTGSKILETaRL--YLQSDP---KM[2]QMRVs------KSFAKW---LKSR--GLHDAK-SVAN[ 1]-LVLSG  3298
Query_PVIOL  3353  Q   NLQKYLMMNG-DI--LWELYQ---LD   -RKLgsELVSDSIKKNQ---IGG----SLSST-SV-N[ 1]-LRDEL  3406
Query_BOMBU  3095  S   --RKQISMIAsKI--IKFSEEneTFA[2]---------------------LREntIILKEI-NAE-[ 5]SIRNTL  3144
Query_AEDES  3213  L[1]--------RIwQR--LKAVVEeaGSE[2]KAVKvaGLHELFQTTEMlREIVRSnpQLGQD--SMEA[ 8]SLQKSI  3281
Query_ANT__  3146  I   --REHICTMAsKI--ESLSREniVFA[2]---------------------LACnsTILRDI-GITE[ 7]NIREYL  3198
Query_WHITE  3218  F   ---SLMADVIsSF--RRIMPKisADN[1]-NSFekKVRSELVSHELlFST----------------    PLEENL  3266
Query_BEETL  3055  H[4]KHLDYISSLLsQN--EDLLQNrqSDL[1]-TL--tEVTQSVSSKDVaFFA--------AN------    KLKSFL  3110
Query_THRIP  3120  H   ---SVESNLAlTI--RETLSAepHSL[1]-TL----IKNQSDGHNWlCNKLQLqpGVSCE------    DIKSTL  3174
Query_TERMI  3221  L   ---QHVSVMAaGI--LQQLEQqpGVL[1]-KLPaaEVKKLADRIG--------apDCSQD------    TLMQNL  3271
Query_FLEA_  3115  Q[3]--------RSwDM--LESLKEidNLK[2]PNLYidVHHSIVDVTKLfKHLIFEnaNILENYpSLEN[11]SHSKTL  3190
Query_MIDGE  3234  I[3]SCASEISLRLwRI--YKKCSEdeTIE   --IPadFSREILKLLDVsDDVDTSn---------VTD[ 8]SL-KFG  3300
Query_LAODE  3195  K   LMSQKHVTEIcSTlh-SVVSTnrNEL   ----------ELFINLFkYKFPKI----DESD----C[ 2]-VLAKL  3246
Query_LOTTI  3270  D   LYRRHHVLTA-TS--YELLAE--SC-   LA-A--ELPQSLSSKTC-DKLKNL--LKTDNL-DKTQ[ 1]NIVSDL  3327
Query_APLYS  3363  D   MGLRHSTLGG-RG--VDLLLG--GL-   LA-L--QSTEGLSENTR-TKLAAQ--ANAGAG-TYEQ    -LVNGL  3418
Query_SCALL  3342  A   FGRHHFVLLG-QT--HELLVD--GI-   RS-V--SGLGCLGDRNR-QRLMDI--TGVSKI-DETE    -LIESL  3397
Query_CRASS  3297  G   LGRRHFVLLG-QA--FGLLVK--GV-   QS-E--NCFSSLSDKNS-EKIMGI--SKASKL-NPRE    -LARCL  3352
Query_OCTOP  3243  S   LRWKHNVLYG-EH--LSSLAS--VLK   NEGD-eFQFENLPLDIQ-QKISYY--AGYESG-TLST[ 1]ELSQKL  3303
Query_POMAC  3306  A   LGMRHMILES-RT--FDAVVT--GL-   MS-V--GNLEELGDKTC-EKLQQY--TK-SPL-NKGD    -LIKEL  3360
Query_WHIPW  3200  H   LTLDHITLKA-VL--YNSVADc-FLN   STKVgnHFLKEIEHFVQ------S----SRAT-TKAS[ 3]NVADCL  3257
Query_TRICH  3195  G   SSADQFALIG-ML--FDNVAEt-VVA   SSE---NTLAEFCLSRW------I----SKTD-TADI[ 3]-VTMCY  3248
Query_AASTA  3007  E   ------VLVFiPTlpATTQADiyGWQ   --------ARALTEAAAfHREYG------DAS-----    -LGTTL  3050
Query_AEUTE  2935  E   ------VLELiSTmpPDAQAEicGWQ   --------ARGLMESSDfQREYG------DPS-----    -TSVKL  2978
Query_PLASM  3178  E   ------IMHMmKIatARDQVAvgYLE   --------AKTFAKAAAfYAFNDL-----DDD-----    -LSKKF  3222
Query_BREMI  3161  E   ------IMQMmETaaVSDKVAmgYLE   --------AKTFANAAIfYASQDI-----DDN-----    -LKEIY  3205
Query_OLIGA  3216  D   ------ILELmESlpLHDRVTmgSLE   --------AKTFGQAAQfYMTADL-----DKS-----    -MTEEF  3260
Query_LIGNO  3598  -   --RRQLISPS-PTwe-ELLAA--GLC   --------------DRAsAALLRIqlASSPPS-TGQA[ 5]--AEAF  3648
Query_CESTO  3351  D[2]-----IVLFShAIstAQVLSVfhDLS   -------LSGRLSGAASkHLVSSF-----QSTlKARF[10]KATDFI  3415
Query_OVI2_  3767  H[4]YSSGSHADMApFTf-FTKVGRllSRL[4]DLPAddGQLSPQASDRLnLLLRQLgtPPFSSY-GCDT[ 4]TLSARL  3846
Query_PRIAP  3344  G   LARENLLLKG-RS--IEAVAR--EV-   LTRP--ASLHTLPAATR-EKLSLV--AGAREH-DHTA    -LVSGL  3400
Query_ROTIF  3263  Y   LYQEQRVLHS-KF--SKFLIE--SLM[3]NQTDfdSGLQGIITEKK--HVLQL-sDYVKND-DIND[ 2]KTISSL  3328
Query_SPONG  3344  P   FISQHHVLLG-KL--LNKMAEifAT-   PEDG-mLALSALSNDGN-TKLNSL-iGGTPST-STKA    -ISEGL  3404
Query_TARDI  3328  L   RSLEADILLA--------VAD--TLH[2]---G-----SSVSLETAdRQRLGL------SS-GTDV[ 2]-----L  3371
Query_TRICH  3318  T   LMREHLLLRS-RL--LDMTSK--IL-   AESD--ASILDLLNEKK-ESLYQL-aNVNQHD-SKE-    KIAAKL  3375
Query_OSTRE  3165  W   --------QGdDHin-------cSLL   -------------EARMcEHLASA--GGDDSI-----    ------  3193
Query_NITEN  3447  G   VSAAVRLLQA-DL--AAQLAR---VC   -VQHhgEIAERLRSGTG---TESW--GGFSSG-QA--    --VVGN  3499
Query_CLUBM  3111  D   WRRVTMLVES-DT--SAQLGR---TK   -LD--------LLIDKD---LDT------NAL-SQ--    --SMAL  3151
Query_MOSS_  3265  E   FLPEFQMEQA-DV--NARVAW---AT   -LS--------SQSGEG---ATS------QDI-TA--    --CFAL  3305
Query_LIVER  3315  D   WQMGLRLLEA-TV--SSQLAW---MN   -LD--------QASKLD---TSS------STL-DV--    --SLSL  3355
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3346  HEAAYNSFKSGVGVA[23]LRQREND-----   ------------------------EIS    GTK---[ 1]ESEED  3402
Query_LINGU  3357  VDRGYQYIKEAV---[ 1]IYEEDSSNDK--   ----STVETHMALAKYCDKYLKLREED    DSL---[ 2]ESIDS  3412
Query_LIMUL  3402  LDVSYQYLQKGVHYI[ 2]LQDALHHTQMvS[1]SVH-CTTDPHMSLAFFCNKYLRKKEDA    TEDNWI[ 1]NSLPS  3469
Query_SPIDE  3339  LDQSYSSLKNITMGT[ 2]LANPVN-----D[1]NRK-SVAAAHMELANFCNRYLCAENAK    NFDLKK[ 1]PSFNE  3401
Query_SCORP  3364  FCKNYEYKCRAIDII[ 2]VEQ-----KMiE[1]DQA-TIGETYFKFANYCNKILRMQEE-    -EEKTL[ 1]LSFET  3424
Query_TICK_  3207  REQALQLHLSAAHCK    ---ATPVSVR--   ------AEAQLRLARFCNTALQQEEGT    LFPDDR[ 2]----S  3258
Query_MONOS  3368  AEQALRTAHEACRQL[46]PLDVVPDMAAvP[4]YPE-RAARHHVRFADFCFQALAQLEAQ[ 3]ASVMPS[19]DLYGR  3503
Query_ROSET  3313  IVRAFDVLLETAQRA[46]PLDVLFPSATaY   GSE--TLNAFVRLAEFCADVLQDRIAM[ 3]------    ---SM  3415
Query_BELCH  3413  LQKGHACLQKAVKLA[ 2]EEKARQDTSL--   DKS-GIVDAFMAMVGFCDKILRHQEDD    E--EG-[ 3]SSIPG  3476
Query_CIONA  3377  SNKAFKFYLKAAKFG[ 3]NNIGEGSLQQ--   -------KSYIALAKFCDEQLRNGDQA    EGLLSP[ 7]----E  3438
Query_CAECI  3397  FKKSFHYLDSAVKKA    EEEVQSRAME-N[1]DTT-GVIKAYMTLVDFCDTHLLKKEDT    TAVIDH    SHLQQ  3460
Query_HUMAN  3378  YQRAFQHLSEAVQAA    EEEAQPPSWS-C[1]PAA-GVIDAYMTLADFCDQQLRKEEEN    ASVIDS    AELQA  3441
Query_MOUSE  3378  YQRAFHHLSKAVQSA    EEETQLSCWG-H[1]AAA-ERAHAYMTLVGFCDQQLRKVEES    ASQKTS    AEMEA  3441
Query_XENOP  3395  YRKSLHYFTNAVRKA    TEEEQSHSTD-Q[1]DVR-GIIKAYMTLVDFCDSHLRKVEEE    SAVMDR    ADYQN  3458
Query_PLATY  3385  HRRAFQYLTHAVRKA    EEEEQSHSSE-H[1]DTG-GVIDTYMTLVDFCDKHLRNEEEG    TSVIDA    VVLQA  3448
Query_TASDE  3390  HKKAFECLFKAARRA    EDEVQSHTTE-H[1]DIA-GVIDAYMTLVDFCDRHLRKEEES    TSVIHT    VELQT  3453
Query_CANAR  3386  NKKAFQCFSCAVRKT    EEEVQSSSVD-H[1]DMV-GVVDAYLTLISFCDQYLRREEEG    LLEINT    VDLQQ  3449
Query_OSTRI  3356  NKRAFQCFSSAVRKS    EEEVQSHSME-H[1]DVA-GVIDSYMTLVGFCDQHLRREEEG    LLEINT    ADLQL  3419
Query_TURTL  3395  NKKAFQYFSTAVRKA    EEEVQSHSME-H[1]DLT-GVIDAYMILVDFCDKHLRKEEEG    SSDVSA    LDLQT  3458
Query_GECKO  3396  NTRAFQCFSSAVRKA    EEEIQSLSFD-H[1]NTK-SIIDAYMILANFCDTHLRKKEEG    SADINA    VDLQM  3459
Query_DANIO  3374  QIQALELLRSAACKA    EEEEQSFSQQ-H[1]NTH-GIVEAYMTMANFCDRRLRESEQK    EEAIS-    SKLQS  3436
Query_SNAKE  3381  NKRAFLCFSGAVKKA    EEEVQSMSFD-H[1]DIS-GIIDAYMTLANFCDSHLRKEEQN    SADVNN    EDLQI  3444
Query_STERL  3400  HRKALQHLQNAVKKA    EEELQSRSTE-H[1]DIP-GAIEAYMTLVNFCDKRLREEEES    AEVSTF    SELPT  3463
Query_MILII  3396  HRKTLHYLERAVRKA    EEEVQSLAGE-H[1]DIA-GVIEAYVNLANFCDRCLREKEES    ATAINF    SEQQT  3459
Query_BAMBO  3386  YRKAHEYLQSAVKKA    EEEVQSLSEE-H[1]DTV-GVIKAYMTLANFCDRCLREEEES    ENINNF    PELQT  3449
Query_ALLIG  3353  NKQAFQYFSNAVRKA    EEEMQSRTME-H[1]NLT-GVIDAYLTLVDFCDKHLHKEREG    LSDISA    VDLQT  3416
Query_GAR__  3397  HKRALGLMQSAVKKV    EDEVQSRSAD-H[1]DTA-GAIEAHMTLANFCDKRLREEEEN    DQVSEF    SELPT  3460
Query_TETRA  3517  MKKDIDLSSQSQSLS[20]NNNFNSSEQIfE[3]PMR---LKSYKRFANFCEQLLRIIESE[17]ETVGSS[19]QKSNK  3637
Query_PARAM  3214  ----YQRYADAIMQT[ 7]-------QIQlH   DENiNQSKFYNKAAMFCERCLRVIEPD[ 3]SNDLRK[ 7]-SLDM  3283
Query_SLEMN  3442  --KHLQRYQKDCTTY[11]IKKVQTAAQFdD   QMKkTMAKTYFKFAQFSDKILRKME--    NDQCRY[10]KSEEQ  3523
Query_STENT  2887  --KMLLSYKTEAADV[ 7]------AFYYlN   QVRhTISA--LKFCRFCDMILRSYEEE    NGENKY[10]VTGDM  2958
Query_NEMAT  3361  ITRSFTSLQEANKAA    K-QAATSSSD-G   STT-CMVEAMMAMVPFCDRALRRKEDE    SDGPSF[10]FVFQR  3432
Query_PISTI  3411  MTRSFVSLQSAIKAA    K-NAEGKCNT-T   QTS-GTVDAMIAMVNFCDNCLRKKEVS    ------[ 5]VDTKL  3471
Query_CORAL  3393  TTKSFTTLQRAIKIA    QTNENNQKNA-E   SVR-NLVTSSMTMANFCDRLLRMKEED    SEQ---[ 3]LDMKL  3455
Query_HYDRA  3234  KEKAWTALSLAYEKS[ 5]STDLDSKLHH--   -------KAYMEMFLYCDELIENEENS    E-----[ 3]----C  3288
Query_LATIM  3394  YRKALSYLQTAAKKA    EEQVQSRSTE-H[1]DVA-GVIEAYMTLVDFCDKRLREVEES    AEVSHF    SELQT  3457
Query_SHRIM  3330  LSCSYKSLQLAVNCA[ 1]GKLVKSDALNqD[1]SKK-SDSDAHLALVKYCE---ECVIKW    KDHVDV    ---SQ  3389
Query_MCRAB  3324  LCCSYKSLKSAVNCG[ 1]EKFMQLNNL-yE[1]AHS-HSKDIYMTLAKYCE---DCLEKW    KESIDV    ---SD  3382
Query_AMPHI  3369  LMCALNAQVQA--CE[ 1]SDKINSEMLDgN[1]VSS-YHIQTHESLAVFCE---DCISAW    SESIQV    ---DE  3426
Query_BARNA  3156  VQLAFGRLSAAVEAA[ 7]------------   -------DACLEAARFCGRQLEALEET    ------[ 6]AELRG  3208
Query_COPEP  3183  MNQVLENYRDAAESG[ 7]------------   -------KMLLEGSTFAHRLLTAEGNS    ------[ 1]-----  3225
Query_DAPHN  2996  IQTGLKHAERQAQIA[ 7]------------   ----DMAEAYFIFAKYFSEWKNVISDE    ------    -EYRL  3044
Query_STARF  3426  YSKAYSELKQAVSLA[ 2]DDAVTSQS----   -ER-SMTEAYMALVSFCDRMLRLKESD    E-VPM-[ 3]PQTDN  3487
Query_STRON  3411  FMKSKKYLDDAVGTI[ 2]DPASHSLSSSaK[1]EEQ-SMAKAHMALVSFCDQLLRQQDDD    DCNP--[ 3]PLMEA  3478
Query_SEACU  3392  YTKALKQLQEISDHS[ 2]SDSVEGES----   -KT-A--ESFVALGKFCDKILKMKDAD    DNLPS-[ 3]PLTAN  3452
Query_FUNGI  3261  YKKALEYHVKASHLA[13]------------   -------KSQLKLALFSDEMIKSGSKD[ 2]LGDH--    -----  3314
Query_DPURP  3383  FNLTFSCYSESIKLN[26]NNNPSSQTSEiN[4]IEKtKIKSAHLKFANFCDNIL--KDK-[ 1]QSNN-G[ 7]----N  3477
Query_HALBU  3419  FSLAYQSYSQSIQLY[ 8]------------   DTH-KSKSAYLQFGNFCDTELQRRGDQ[ 2]DSN---    ----R  3473
Query_GLOBO  3155  LDRAKHNILAIHA--[ 7]------------   --F---QKGLIEVGMFVDKTLRCVEAD    EVSTKF[ 7]-----  3209
Query_PALUS  3217  TSRAYTCLQQA----[ 5]------------   --K---DKHRLRISIHCDRILRRMEEE    WGVSGD[ 7]---CL  3269
Query_CALIF  3309  QGRGLKRLLETLKIY[ 8]------------   --N---KKIYLSIATFCDQSIKRIENS[ 1]NENTLK[ 7]---NE  3369
Query_GLOMU  3331  -ERGYNLFNNATKEI[ 8]------------   --H---AKILVKFAKFCDYFLR-----[ 2]DFTEEQ[ 7]---AQ  3386
Query_LOBOS  3302  QQRGYECLRAAA---[ 8]------------   --V---RKLRLNTASYCDKVLRHHENE[ 3]EGGESS[ 7]---AV  3361
Query_RACEM  3299  TRYLQDILKNAVE--[ 4]------------   --R---AECVWRLAEYCDNALRANEDP    ---SSK[ 7]---DH  3349
Query_PVIOL  3407  FSLTFKSFSESISLS[ 8]------------   VDK-KHKSANLKFANFCDNVL--RDKQ[ 2]QSNSIF[ 7]----Q  3469
Query_BOMBU  3145  ETYSFNNLKTCCDMT    --TKN-------   -----IKECYFSLAKYCYDRLLHDTHD    ----IQ    -LSKE  3190
Query_AEDES  3282  EVLSEDMFDDSTARS    --N---------   DVA-LLADCHFKMAQFCYDQL-EVEVL    ----GE    -TLDL  3328
Query_ANT__  3199  LHYSLNNLRYCCSQD    --TTK-------   -----LGEHYCTLAKYCYGRLNTEPEN    ----DE    -IFHD  3244
Query_WHITE  3267  RKLGFLCLKEALEWA    --GRTESVNQ--   -ET-VLADAHMKIVRYCFGVYNQQEEA    ----PE    -EYN-  3319
Query_BEETL  3111  HNYNTSDVCTDAYI-    ------------   ----QYADAYVQLAYFAK---------    ----NQ    -GKKY  3144
Query_THRIP  3175  DTLGLNSLIHSKNIA    --KKHKDASSfL   -TE-QTAEAYLQIAQYCR-NLLENQEG    ----SS    -GTNT  3229
Query_TERMI  3272  KQYGLSSLKTAAEVV    --TPRGNSE---   --Q-AVAKAYLQVAKYCRAHSNTSRED    ----NF    -GTD-  3322
Query_FLEA_  3191  DIFGFKCLQKSIKYA    --EAAHANNDiN[1]EIS-TVTNCYKKMIQFNMHNL-KYAED    ----DL    -KSVY  3247
Query_MIDGE  3301  RKLAQKILDTNYSME    --N---------   ET--MLGEVHHKLGQFYYQVY-ESGAL    -----T    -SKEI  3345
Query_LAODE  3247  YETGLESLKTAVQIA[ 7]------------   SKA-DLAETYLTLVKYCQQ-----END    ------[ 3]-TVES  3296
Query_LOTTI  3328  LQKSHQELKSAI---[ 1]DSVVTRS----R[1]ESF-TEADAQIALSKFCDKFLRMKENE    EE----[ 4]KVIEN  3386
Query_APLYS  3419  ISKGFHSLKAAV---[ 1]EEGSRSGGAEvR[1]TAM-TCEEANLELAKFCDKYLRLLEDD    ELP---[ 2]CDTSQ  3480
Query_SCALL  3398  VKRGFENMKKSL---[ 1]YEDDRNS----K[1]TGY-TRDMAFMAVAKFCDRILRMVEDD    ----EL[ 4]ATIKT  3456
Query_CRASS  3353  VEEGYNQMRKSL---[ 1]Y-DGRGT----R[1]CRY-GLEEAYLALAQYCNKFLRMAEDD    TDKESL[ 4]ECQKN  3414
Query_OCTOP  3304  IISGCEYLEKNVISD[ 2]EEKSVSV-----   -RE-SIGETHLAIAKYCNQFLQDS---    ------    ---PS  3351
Query_POMAC  3361  IHHGYESLRKVI---[ 1]EVPEDSS----R[1]AVC-LQQDVQLALAKYCDKYLRLAEEN    DGL---[ 3]ESMEV  3419
Query_WHIPW  3258  LCNAFDELKVAISLC    -----SKNRE--   SYR-QLTEPHLVMTEVCSKILR-----    ------    -----  3298
Query_TRICH  3249  -GQAFRYLKLSCQYS    -----SQLNI--   ESP-KLTEPLMELSSFCDRILRKNGHT    SLVNSE[ 3]-----  3302
Query_AASTA  3051  ERGAMEIFQYQTAKA    -TTMTSM-----   ---------HVMYIEYLDKCLERSTAG    -TFGSL    --AKA  3097
Query_AEUTE  2979  SSAALEIFQYQTSKA    -TNVTTM-----   ---------HLTYIEYLDKIIERTPSS    -T--DL    --VNQ  3023
Query_PLASM  3223  FYRSLDRFKDSCRRI    -DAVLSK-QSqD   DAPpAFLRCRLTYIEFLNDLLYKRQLE    -KLANF    --ADS  3282
Query_BREMI  3206  FSRAVDMFKASSRRV    -DAVMSN-DLrE   GGSsRFLRCRLTFIEFLNDLLFKPQLE    -KLAKL    --ADR  3265
Query_OLIGA  3261  FTRAIDVFQRSCRSI    -PGEITMVALsR   ESEqSFSRCRVTFIEYLIDILFSEKRA    -EWMTL    --VEQ  3321
Query_LIGNO  3649  DKSALSNLQSAASKA[ 5]-ADVGSSSLQ--   --------CHLELIRFCDSCLLEAAPN    QQPQQF[ 2]----A  3705
Query_CESTO  3416  QASAFAFFKRCVDLA[15]--SAENLLCRrN[4]KPE----EALLEVANFCNSRIDEEAAI    ------    ---RA  3484
Query_OVI2_  3847  SSYTLSLSSQATRMA[ 7]--SKMTPSLIqQ   -----LSEPMIVLADFCDRKLISASPS    ------[ 9]EQLST  3914
Query_PRIAP  3401  ISAGFRELREAVDVG[ 2]EESSASLASSp-   ------SAAHVAMASYCDYFLRVREDD    G---T-[ 3]P-TDG  3458
Query_ROTIF  3329  MKHGVDNLLKIKDSL[ 2]-------SLE--   ------------LATYCDFYLRIAEND    NNDFEK[ 6]MIRSQ  3380
Query_SPONG  3405  WIQALSSFNAGVRSG    -TTLNPLTKD--   -PV-PTVQSLEELFGFCDNALRQEDHD    DEQSTG[ 1]-DTKS  3464
Query_TARDI  3372  YVRVFSEFAEACERA[ 1]---------D--   ------AVPFLRTFQYCDQLLRKGILD    DVVLVE[ 2]-----  3417
Query_TRICH  3376  TSHACQALHQAVKAS    K-------EE-R   GGS-SYIEAMMMMVNFCDDILRHEEDG    RAN---[ 5]IDTKI  3433
Query_OSTRE  3194  -TYGIAAIENFTAIA[ 7]-----------H   -----AVDASLRLAHFCDAALKKLHGD[ 3]ASISSN[ 9]AENHS  3260
Query_NITEN  3500  LGAAFRGYSDAVSIL[21]------------   -AR-RQSKASLKFARFCNDVLQLATGE[ 2]QTQTNS[19]PPVSQ  3592
Query_CLUBM  3152  LGSAYKTYSELLRNA[ 5]------------   -SR-EMARTSFKFACFCDDLLNTADDS[ 2]KSEVRR[11]TSFSV  3220
Query_MOSS_  3306  LGSAFKTYSTALDND[13]------------   --H-DTAKALLKFGLFCDELLKLSQQP[ 2]RSQMTQ[ 6]PQASD  3376
Query_LIVER  3356  LGSAYGAYS------[17]------------   -SR-QLAKISLSFGLFCDELLKAAKSM[ 2]RDDIRS[11]PPASA  3430
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3403  --FPLALVTSLLS-AMKNSS   KDARQLFPRLLQI--IERFP-QTREV    FIRKSRD---IP-AWMFLGWVSQICAI  3465
Query_LINGU  3413  --YPEAVVTSLLR-CIELDH   MDARQRFPRLLQL--VEYYP-DTMQG    FINKVSV---IP-CWMFMTWIGQLVAL  3475
Query_LIMUL  3470  --YPTTLVTSLLT-AMRQGS   KAAMQLFPRLLQV--LELYP-GCSED    FVKKTNE---IP-SWMFISWINQIVAL  3532
Query_SPIDE  3402  --FPRILIESLLD-ALKYGS   TDAHLLFPRLLQI--IENNP-SCLQV    FEEKVSK---MP-CWLFIGWIDQMVAL  3464
Query_SCORP  3425  --YPSVLTNSLLN-ALKNGN   LNAIKLFPRLLQI--IEVYP-DTIPE    FIAKSKV---VP-CWMFLSWINQLLTL  3487
Query_TICK_  3259  --LPLSLVRSVLT-AMRLGS   DAAIDMFPRLLQV--LQSHP-LCGKS    FAEECAK---VP-CWMFLGWIGQMLPL  3321
Query_MONOS  3504  slLLQLAVEHSFA-AARSGL[1]-EGRALLPRVLALleVEEERPPWLPQ[ 2]----CAAgcaAP-AWIYLRWLPQLLAY  3572
Query_ROSET  3416  grVKQVFVETALA-AMRSGN[2]--MRQYFPQLLNLg---ERPTDSDHL[ 3]LASATTS---VP-CWMFLNWIPQMLAI  3483
Query_BELCH  3477  --LAAAVVSNLLQ-ALRHDS   PEARQRFPRLLQL--VEIYP-DTMDL    FIEKVSH---VP-SWLCISWLGQMVAL  3539
Query_CIONA  3439  --YAVTMTTYVLR-AMRIQS[3]SKASWSFPRLLQL--VQDFKGNILDS    FIQESSA---LP-SWMVLSWLGHLVAL  3505
Query_CAECI  3461  --FPTIVVEKMIK-SLKLNS   KDARLKFPRLLQI--VELYPAETLDL    MTREVSS---VP-CWQFIGWISQMMAM  3524
Query_HUMAN  3442  --YPALVVEKMLK-ALKLNS   NEARLKFPRLLQI--IERYPEETLSL    MTKEISS---VP-CWQFISWISHMVAL  3505
Query_MOUSE  3442  --YPALVVEKMLR-ALKLNS   SEARLKFPRLLQI--IEQYSEETLNI    MTKEISS---IP-CWQFIGWISHMMAL  3505
Query_XENOP  3459  --FPEIMVEKMIK-ALKLNS   SEARLKFPRMLQI--IEQYPSETLDL    MARENCT---VP-CWQFIGWISQMMAM  3522
Query_PLATY  3449  --YPALVVEKMMK-ALKLNS   REARLKFPRLLQI--IERYPEETLDL    MTQEISS---IP-CWQFIGWISQMMAV  3512
Query_TASDE  3454  --YPALVVDKMLK-ALKLNS   KEARLKFPRLLQI--LERYPEETLNL    MIKEMSS---IP-CWQFIGWISQMMAL  3517
Query_CANAR  3450  --FPAIVVEKMIK-ALKLNS   REARLRFPRLLQI--IERYPAETLGL    VTQELSS---VP-CWQFIGWISQMMAL  3513
Query_OSTRI  3420  --FPAIVVEKMIK-ALKLNS   REARLRFPRLLQI--IERYPAETLGL    VTRELSS---VP-CWQFIGWISQMMAL  3483
Query_TURTL  3459  --FPAIVVEKMIK-ALKLNS   RDARLKFPRLLQI--IERYPAETLGL    MAREISS---VP-CWQFIGWISQLMAV  3522
Query_GECKO  3460  --FPAIVVEKVIK-ALKLDS   KEARLKFPRLLQI--VEKYPTETLGL    MAQEISS---VP-CWQFIGWISQLMAM  3523
Query_DANIO  3437  --LPEHVVKMMLK-ALKLSS   EEARLKFPRLLQL--VEVYPAETLDL    MVREVVS---VP-CWLLIGWISQMMAL  3500
Query_SNAKE  3445  --FPAIVVEKVIK-ALKLNS   NEARLKFPRLLQI--VEKYPEKILRL    MAQEISS---VP-CWQFIGWISQLMAM  3508
Query_STERL  3464  --FPAVVVDTMLK-ALKLNS   KEARLKFPRLLQI--IELYPAETLDL    MAKEVSS---VP-CWQLIGWISQMMAL  3527
Query_MILII  3460  --FPACVVEKILK-AMKYNS   KEARLKFPRLLQI--VKLYSAETLDL    MKKEVSS---VP-CWQFIGWINQMMAE  3523
Query_BAMBO  3450  --YPASVVENMLK-ALQYNS   REARLRFPRLLQI--VELYPVETQDV    MTRGVST---IP-CWQFIGWINQMMAS  3513
Query_ALLIG  3417  --FPTIVVEKMIK-ALKLNS   REARLKFPRLLQI--IERYPAETLSL    MAQEIST---VP-CWQFIGWISQMMAV  3480
Query_GAR__  3461  --FPAIVVEQVLK-ALKLNS   SEARLKFPRLLQI--VELYSAETLDL    MIKEMSS---VP-CWQLIGWISQMMAL  3524
Query_TETRA  3638  mqMEVETGEKIQEnPEKSQN[6]QNMSNLVLKMLRI--IQQIPNMTGQK    FKEAIELqi-IP-SWIFISFLPQMMHI  3712
Query_PARAM  3284  --IADQFVRYTIS-SFKFNA[3]NKIETLLIKVLKV--VQKFPNQIGPV    FQDAVEKgh-LS-TWIFLQYLPQLIHL  3352
Query_SLEMN  3524  --VARDSIKYGL-lSLNHEY   QNASDIIPRLLDI--TSKYKDHVQGD    FIA--YSke-TP-TWYFLRWINQIVAV  3587
Query_STENT  2959  --LAEIIVSSTL-qSMKLGN   QRAHDMFPRLIGL--F-KYQN-VAKI    FEN--ELed-FP-EWMMIRWIDQVLAI  3020
Query_NEMAT  3433  --FPEFVVRYVLK-AMEHGS   EEARNRFPRLLQI--VELYPH-TLDS    FVKKVET---VP-CWMFIGWINQMVAL  3495
Query_PISTI  3472  --FPGIVVRYMLK-SMCYSS   KEARQKFPRLLQI--VHKYPD-TLEA    FVKKAAD---VP-CWMFIGWINQMMAI  3534
Query_CORAL  3456  --FPGIVIRFVLK-SMAFES   AEARQKFPRLLQL--VELYPGGDIEA    FKRKSSD---VP-CWMFIGWISQMVAL  3519
Query_HYDRA  3289  --YAKTMVDSLLK-AMAMNS   SKARIAFPRLLQV--VNLFPSTG-EL    FEKQSSQ---VP-CWMFLGWINQIVSM  3351
Query_LATIM  3458  --FPAVVVENLLR-ALKLNS   REARLKFPRLLQI--VELYPAETLDL    MAKMGGS---VP-CWQFIGWISQMMAL  3521
Query_SHRIM  3390  --YSKSLMVSVLR-AMANGS   RDAHFHFPRLINL--MNEDA-SLLSS    FKEEYEK---VP-IWMFLLWLSHILIY  3452
Query_MCRAB  3383  --YADSLVMSVLR-AMSLGS   REAHFHFPRVINL--MSDDS-SLVAT    FKKEAKK---VP-VWMFLLWISHILIY  3445
Query_AMPHI  3427  --YLKILVTSTLK-AMTAGS   RVSHFNFPRLIHL--MEDHA-ELIEV    FQREIST---VP-VWLFLLWLPHILIY  3489
Query_BARNA  3209  p-----FVRCTLW-ALAAGS   AEACDWLPRLLQL--ADEHPPAAAAL[ 4]-----EA---VP-EWRFLAWLPQLFSY  3268
Query_COPEP  3226  -eYGELCVSYHLK-SMKAGC[1]-AARELFPRILSL---------VLEN[ 6]FEDNWST---VP-SWMFLPWSNQLISC  3289
Query_DAPHN  3045  g-----LTRSVLS-SMRYGS[1]-DGQGLFPLVILAl-----KDDCAFT[ 2]FETLRCD---VP-LFMFLPWVNQLVSY  3104
Query_STARF  3488  --FSQTVVHYTLK-AMHYSS   QEARQRFPRLLQL--VEACP-DTMQL    MIKEVSD---VP-CWMFIGWISQMVAL  3550
Query_STRON  3479  --YPSTVVNSTLQ-AMNLNS   PEARQHFPRLLQL--VEKYP-DTMQS    LVKKASI---IP-SWMFIGWISQMVAL  3541
Query_SEACU  3453  --FPSTVVKSFLT-AMRYSS   SEAKQRFPRLLQI--IETYP-ETLKD    FIKETSS---VP-CWMFIGWISQMVAL  3515
Query_FUNGI  3315  ------VVRGVVD-AMIFEY[1]-SASDYFPRLLEI--TGSYP-SSNRL    FSKLIRE---IP-TWQFIRWTPQIIAL  3373
Query_DPURP  3478  --LALTVIKSTLE-AIREEL[1]-GSIDKFPRLLEI--LTQFE-QYQQV[ 3]FTEHVSN---IP-CWMFIRWISQMFPY  3543
Query_HALBU  3474  --LAVSVVNSIMN-AIKQEI[1]-GAIEKFPRLLEI--ISQYE-KAAK-[ 1]FVERTKQ---IP-CWMFIRWQSQMFPY  3536
Query_GLOBO  3210  ---AKVIIHSIFR-AMSMGS[1]-KAVEFFPRLLQI--LETYP-EVAND    FSVHSQD---VP-TWMYLRWLPQLTSL  3271
Query_PALUS  3270  --YARMVITNTLA-AMKSGK[1]-DAIEAYPRLLQL--IELYE-ETADD    FQRMATS---VP-SWTFLRWMPQMTAL  3332
Query_CALIF  3370  --YSKLEIKYLLK-ALKYGS[1]-EGIQLFPRLLQL--IENDE-SNQQA    FIKNIEG---IE-PWLFIRWIPQITAL  3432
Query_GLOMU  3387  --YASNVVESILQ-AMENGS[1]-EATELFPRLLQI--IEKYH-NTQSI    FKEKVES---IEsVWKFIRWIPQIVAI  3450
Query_LOBOS  3362  --HANLVIKNTLL-SIRDGE[1]-GATELFPRLLQI--IEIYP-SSQKP    FTEMVTD---LSgSWTFVRWIQQMVAV  3425
Query_RACEM  3350  --YAQTVIQNYYE-SMNMDV[1]-EAIERFPRLLEL--IERYP-EVQKS    FKVHAEA---FNvSWMYIRWLPQLVAI  3413
Query_PVIOL  3470  --LAKSTIRSTLE-AIREEI[1]-GAIDKFPRLLEL--LQQFD-KSLEI[ 3]FKELIST---IP-SWMFIRWISQMFPY  3535
Query_BOMBU  3191  --F----IHSILK-AMSYGS[1]-EAAHYFPCLLKP--EYFNDQETKDI    FMKEIES---IQ-IWLFLSWQAQLFSH  3250
Query_AEDES  3329  --E-RQLVTALLA-AMQYGS[1]-SARQLFPCLLQL--PFLQDGTLHGC    FNEASHP---VP-EWMFLRWIPQILSF  3391
Query_ANT__  3245  --F----LSSTLK-SMFYDY[1]-EAMHYFPCLLKP--ERFKDEQTRNT    FVEECAK---LR-PWLFLRWRDLLFSF  3304
Query_WHITE  3320  ----QYFIRSLLS-AMEYNS[1]-EATQLFPCLLQL--ECLEISEMSEL    FKNQCKR---IP-EWKFLEWVPQLLTA  3381
Query_BEETL  3145  --F-KDFMLFVLR-AMKFNS[1]-EGKQLFPCVLME--QEFS-LEEQQL    FLTETEQ---IP-VWMFLKWIPQILVN  3206
Query_THRIP  3230  --IARHLVKAILC-AMKCGS[1]-RARQLFPLLLQL--PTLS-SNLSKD    FAEKSSS---IP-VWMFLGWIPQLLAM  3292
Query_TERMI  3323  ------IVTSILR-AMRYGS[1]-DARQLFPCLLQL--PGLD-TSLSEL    FHKESSA---VP-VWMFLGWVNQLLAC  3381
Query_FLEA_  3248  --I-LSMIKSILK-AMALGS[1]-EAKQLFPVLLEY--DEI-SADMKDV    FLELSAQ---VP-EWMFIQWAPQLLAQ  3309
Query_MIDGE  3346  --Q-IKILESIFR-AITHGS[1]-DAKFFIPWILQL--PYLKNNELTSQ    FNNQLTL---VP-EWNFISYISQMMSN  3408
Query_LAODE  3297  k-----MIESLLR-AMKLGS   KEARHLFPIILSL------PRLAQSH[ 3]FFEESKQ---VP-AWMFLSWANQMLAR  3356
Query_LOTTI  3387  --YGETIIICLLK-AMKLGS   MEAIERFPRLLQF--VELYP-DTMPT    FIKMVSEnrlFP-LWMFLLWINQMMAL  3452
Query_APLYS  3481  --FPRTVVLCLLR-AIKGGQ   LEARERFPRLLQL--PELYP-DVVDV    FNKKVTE---VP-SWMFIMWVSQMTAL  3543
Query_SCALL  3457  --FPDTVVVCLLN-AMRFDS   MEAKQRFPRLLQL--VEHYP-DCMGT    FVDKCEE---VP-CWMYLLWISQMMAL  3519
Query_CRASS  3415  --FPETITVCLLN-AMKLDS   MEARQRFPRLLQL--VEMYP-EIRST    FVEKSGQ---VP-CWMFIMWISQMMAL  3477
Query_OCTOP  3352  --LGNNLMSSLLK-AIRYGI   KEAQQRFCHLLQL--LETHA-ESIDI    FINESKK---IP-CWMFLNWINQMISL  3414
Query_POMAC  3420  --FPATAVKCLLD-AMRAGS   AEALERFPRLLHI--IEHYP-DTTDM    FLKKSET---IP-SWMFILWISQMMAL  3482
Query_WHIPW  3299  ---------ALEE-ELWLPG[2]--VYSLFKPIPSI---EQYPDCRSTF    -VRESQQ---VP-TWLFLGWLDQLLAV  3353
Query_TRICH  3303  -----VMVRSILK-SMKNGS[2]--ARQLFPRLLSV--MEKYPESELEF    -QAEIDS---VP-SWMFLSWMNQMLSA  3362
Query_AASTA  3098  ------MVETVLF-GMALAD[1]-TCAAYFPRVLGL--VRDHPADLLPI[ 1]-TSKMQP---VP-LWTCLHWSAQIMAM  3157
Query_AEUTE  3024  ------FVTNVLW-GMSMAD[1]-GCSAYFPRVLGL--LREHPT-LINM[ 1]-REKLKQ---IP-LWTCLHWSAQVMAL  3082
Query_PLASM  3283  kvLIKLLAENVLR-GMAAGD[1]-ECANYFPGLCDV--VSPYPDIVAEC[ 1]-QCVLTD---VP-LWTCLQWSAQLMAL  3348
Query_BREMI  3266  ktLVKLLVDNVLG-GIAAGD[1]-ECAHYFPQLCDV--ISPYQEIVTDF[ 1]-QFVMKN---VP-LWTCLQWSAQLMAL  3331
Query_OLIGA  3322  rtMLESLTENVLS-GMVAGD[1]-ECSNYLPQLCEL--IPPFPVIVAKF[ 1]-KRMLSE---VP-MWTCLRWTAQLLAL  3387
Query_LIGNO  3706  --LRAQAVDSVTS-AMRCGS   QEASIRFPRLLCL--LGESADDEASQ[10]FSRQMAS---VP-VWLSLPWLPQLVSL  3779
Query_CESTO  3485  s-YTEIFTRSVLT-AMRLGS   EGGRIRFGRVLQL-eVARCSASPSTY[ 4]FRDLTKDls--P--CLFSQWFDRLMTA  3554
Query_OVI2_  3915  saLCYSFIEATLC-SMQLNH   RAGHLRFPRALQLasIDNDPGSSIQS[ 1]FISGTAK---LP-TWMFTQWLDLLVTG  3983
Query_PRIAP  3459  --YAETVITSLLR-AMALGS   SEARDSFPRLLQL--VDLYP-DTLPL    LTSKVES---VP-CWMFIGWISQMVAL  3521
Query_ROTIF  3381  --FPEIVTEQLFY-SIRSGS   LEARTRFPRLLQI--IELYPDECLEK    FVSSSED---IP-SWMFLSWLSQMTAL  3444
Query_SPONG  3465  --FPSIIINHTLT-AMSLGS   SKARELFPRLLQL--VEAHP-DTMDL    MINKSST---IP-CWMFIGWINQMIAL  3527
Query_TARDI  3418  -------TASLLK-AIKLGS[1]-EARDLFCRLLVW--LREHP-NIEDA    FQVGCAD---IP-AWMFIQWANQVLAL  3475
Query_TRICH  3434  --FPEVIIRYTLE-AMAYNS   MDARQRFPRLLQI--VERNPN-TIDG    FMRKSSA---IP-SWMFIGWISQMMAI  3496
Query_OSTRE  3261  s-LSQVFIRHTLK-SLEMYG[5]-PARHLLPRLLVL--LRDHMSPQTAQ[ 1]YSQTLLR---VP-SWLFLDWIPQLLSM  3330
Query_NITEN  3593  --YPPLVVKHVLR-ALSLDA[2]-KPHQLVPRVLAL--VGQNK-ETLA-[ 1]FESGVES---VP-SWVFISWVSQILSV  3656
Query_CLUBM  3221  --FREAMVTNMLR-TLSMDA[2]-GPQLAVARLISM--AG-HQ-EIQA-[ 1]FSARIDK---VP-SWIFLGWLPQLLAL  3283
Query_MOSS_  3377  --LSLLVVRHMLK-ALALDS[2]-QAQHAVARLLSL--VGFSS-ATQN-[ 1]FERFLDN---VS-SWVFVAWIPQMLAN  3440
Query_LIVER  3431  --YPSVIVKHMME-AISLDA[2]-RAQHGLARVLSL--VGQYP-ETHT-[ 1]FADVVKQ---VP-CWVFISWIPQMLAV  3494
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3466  LDKP-E   SSAIHEILLNIADIYPNALVFPFKLSNHNYKFTS-DAV    SSKS   KSVCDRLHQ-MLS--kNPL--I  3528
Query_LINGU  3476  LDKP-E   APAVHNILIRIAREYPQALVYPFKVSKECYTFER-SSE    GKEN   QAVVERLDF-SLS--kIPL--V  3538
Query_LIMUL  3533  LDKP-E   APPLHGIVERMAKEYPNALVYPFKLSSKGYQFDN-SNL    GRKN   KAFVEKIES-LMGV--VPL--V  3595
Query_SPIDE  3465  LDKP-E   AKAVQCIIEKIALTYPNAIIYAFHLSSGSYKFSN-SKS    DSSN   KKFVDKIKA-ILEK--ETL--I  3527
Query_SCORP  3488  LDKP-E   ATAVQYIIKEIGTNFPNALVYPFKLSYEMYKFPD-TLQ    GKEN   KEFVMKLKS-SLSQ--LKL--V  3550
Query_TICK_  3322  LDGA-V   GPHLFGLVDSIASDYPNALVYAFRVSNEAYSFDI-P--    --ET   KDFILRLKQ-KMDK----LplV  3380
Query_MONOS  3573  GATD-Q[3]--AVKQILGLVAKAYPQAVAYPAILLQDQLEERIvSND[ 8]--KN   -----RISG---QD--------  3632
Query_ROSET  3484  LDAP-E   APVVQQVLKTIASSYPQALCFPVMLSEEKLKFSGaASS    TQRT   --FFTNLVR-MVTT-----dhL  3544
Query_BELCH  3540  LDKP-E   AVAVHGILTRIAQEYPQALVYPLKISSENFEFED-SVT    GRRN   QEAVTSLKK-TLSV-iYSL--V  3603
Query_CIONA  3506  LDRD-V   APAVQPILTRVMRDYPQAIVFPFKAAFEGYKFEE-NEA    GKKN   KQFVKKVKH-HLKQ-eVPL--A  3569
Query_CAECI  3525  LDKK-E   AVAVQHIIEDIADAYPEALVYPFMISSENYSFED-TAT    GHKN   REFMLRIKS-KLDK--GGM--V  3587
Query_HUMAN  3506  LDKD-Q   AVAVQHSVEEITDNYPQAIVYPFIISSESYSFKD-TST    GHKN   KEFVARIKS-KLDQ--GGV--I  3568
Query_MOUSE  3506  LDKE-E   AIAVQHTVEEIADNYPQAIIYPFIISSESYSFKN-TSS    GHNN   KAFVERIKS-KLDH--GEV--I  3568
Query_XENOP  3523  LDKK-E   SIAVQHIIEEIAENYPQALVYPFMVSGESYNFED-TVV    GHKN   REYVNRIKS-KLDK--DNV--A  3585
Query_PLATY  3513  LDKE-E   AVAVQHTIEEIADNYPQAIVYPFMISSENYSFKD-TSA    GHKN   KEFVTRIKS-KLDK--GGV--V  3575
Query_TASDE  3518  LDKE-E   AVAVQHTVAEIADNYPQAIVYPFMISSENYSFKD-TSA    GHKK   KEFVLRIKS-KLDQ--GGV--V  3580
Query_CANAR  3514  LDKD-E   AVAVQHTVEEIANTYPQALVYPFMISSESYCFKD-TAA    GCKN   KEFVERIKN-KLDR--GGV--V  3576
Query_OSTRI  3484  LDKD-E   AIAVQHTVEEIANTYPQAIIYPFMISSESYSFKD-TVT    GCKN   KEFVERIKK-KLDR--GGV--V  3546
Query_TURTL  3523  LDKE-E   AVAVQHTVEEIANSYPQAIVYPFLISSENYSFKD-TAT    GHKN   KEFVARIKS-KLDR--GGV--I  3585
Query_GECKO  3524  LDKE-E   ALAVQYTVEEIANSYPQAIVYPFMISSESFSFQD-TII    GHKN   KEFVARVKS-KLDR--EGV--V  3586
Query_DANIO  3501  LDKP-Q   ATAVQHVIEEIAECYPQALIYPYMISSENYTFEE-SAS    GQRN   REFVEKLES-LLDK--GGV--I  3563
Query_SNAKE  3509  LDKN-E   APAVQHTVEEIASSYPQAIVYPFMISSESFSFPE-TAI    GHKN   KEFVKRVKS-KLYE--GGV--I  3571
Query_STERL  3528  LDKE-E   AVAVQHVIKEIAEYYPQAILYPFMISSENYVFED-SAT    GHRN   KEFVARLKS-QLDK--GGV--V  3590
Query_MILII  3524  LDKK-E   AVVVQHIIEEIADSYPQAIVYPLMISSESYTFEQ-TAR    GNTN   KEFVERLKN-KLDK--GGM--V  3586
Query_BAMBO  3514  LDKK-E   ACVIQPIIEKIADSYPQAIVYPFKISSENYTFEP-SAN    GNKN   KEFVERLKT-KLDK--GGV--V  3576
Query_ALLIG  3481  LDKD-E   AVAVHHTVEEIAHNYPQAIIFPFMISSENYSFKD-TAT    GHKN   KEFVTRIKS-KLDR--GGV--I  3543
Query_GAR__  3525  LDKK-E   AVAVQHVIQGLAESYPQAVVYPFMISSETYQFED-SAT    GHRN   KEFVTRLKD-QLDK--GGV--V  3587
Query_TETRA  3713  LKKQ-N[4]---FEPLIKKLAKNYPEALIYQFNTIFSNQTTCE-TAL    ---S   ----QEIYE-ALHK---EMphY  3770
Query_PARAM  3353  INDNvR[3]---FKRIFDDFIKKYPQQFVYQFNVTYQTRKEFKpDSE    ASKM   ---FDKLHQ-CIPK---NYa--  3413
Query_SLEMN  3588  VNRP-E[1]-SVIEEKVIQIAKKYPQALYYPFKVVESNVKVNVlDSD    VQKT[3]QKLS-KYYEkTFRN-------L  3651
Query_STENT  3021  FDKP-E[1]-NVFAKIMERLVIKYPQIVYYPYKTFCS---IDSlSGF    TKVK[3]KSLTWKILEeKFQD--FGF--L  3085
Query_NEMAT  3496  MDKK-E   ARAVHGIVTEIAKNYPQALCYAFKVSSEQFVFDS-SPE    GNTN   RKAVERIRK-LLES--P-L--M  3557
Query_PISTI  3535  IDKP-E   GKAVHGIVSEIAKRYPQALWYPFKISSEHFVFEN-SED    GKAN   EEAVKRLRS-SLEH--P-L--L  3596
Query_CORAL  3520  LDKR-E   SHAVQGILQDLAKHYPQALLYPLKISAEQFKFGG-SIE    EQKN   KKFVESLNE-SLSF--P-A--S  3581
Query_HYDRA  3352  IDMP-I   VVHLKPIIDGIVKSYKQALSYPLKVSSKQFTFST-VDS    --QN   KEYLKWLLR-CVN---IPV--I  3411
Query_LATIM  3522  LDKK-E   AVAVQHIVEKIAESYPQAIVYNFMISSESYSFEE-SPK    GKKN   KEFVARLKS-ELDK--GAV--V  3584
Query_SHRIM  3453  VDKV-P   GPALQPVVERLAAEYPQAVIYPFGISKQQYNFSS--HE    GKLA   KAMCDKVEA-LL-S-sHKL--L  3514
Query_MCRAB  3446  VDKL-S   GEALQPIVEKLALEYPQAIVYPFRISMKQYDFTG--SV    GKAA   QSMCQRVES-SL-C-kNSL--F  3507
Query_AMPHI  3490  LEKT-P   GRVLRPIVVQLAEEYPQAVFYPFSVSKTCYKFDG--PH    GSDV   RATCVKVEE-ILDT-kCRL--L  3552
Query_BARNA  3269  GASA--[2]GGTVRRLLLRLAEQYPQALAYPLAASRADRGAVR--PA    ----   -----EMSA-CLEQ-aHQL---  3322
Query_COPEP  3290  LHND-E[1]SKYFTPVVSKMSEDYPAAVLYPLRSAASNPDLSDiAMP    ----   --IFYKLTAkLVVS---SVh--  3348
Query_DAPHN  3105  FTSS--   -SAIPKLLIDIAKQHPVHVKIPFGLTRSSLGVEVlAQS    -ETR[1]------LESkLLTD----Lt-W  3160
Query_STARF  3551  LDKP-E   GPAVHGILTDIAYEYPQALCFPFKISSESFTFDE-SNT    GRKN   ASAARKIGA-ILES-i-PL--L  3613
Query_STRON  3542  LDKP-E   SLAVQGILKAIAKDYPQALVYPFKISSEDFKLGD-KSK    EKQR   KDAVKEIED-ILST-v-PL--V  3604
Query_SEACU  3516  LDKP-E   ADGVAGIVKAIAVEYPQALVYPYKTSSENCKFSK-DTD    GKKR   QQAMIELKQ-LL-N-f-PV--I  3577
Query_FUNGI  3374  LGSDsQ   GDTVLPILVSLAEKYPQSIYYSYKISSEDDSL--qSPA    MQKR   ---LEPLNK-ILKD---EL--I  3432
Query_DPURP  3544  LDLP-Q   GPLILPLLLDIAKYYPQAIYFPYKISSEQFG-----PM    AKKL   ---SVQLEK-ELSD---PL--L  3598
Query_HALBU  3537  LDTP-Q   GPYILPILQEIASNYPQAIYFPFKISSEQFG-----PT    AKKL   ---TAKLEQ-TLKN---PL--I  3591
Query_GLOBO  3272  LDKV-T   GLTLLPLVQKISTKYPNALRFPFSISCEQYQF---DAF    NAPN   AKLIERLKS-SLQS---AM--H  3331
Query_PALUS  3333  LDKP-S   AQAVYPVIEKIGRDYPSALYFPLSISSEQYTF-GnDVG    SQRN   KKRIEALKT-IISS---PL--L  3394
Query_CALIF  3433  LDKD-C   GYTLIPLLKKIALSFPNALSYPLNISCEQYNF--eTEI    SSKN[4]EEEVKKLKE-IIYD---EQ--N  3497
Query_GLOMU  3451  LDQS-I   AHCVFPILFKLAEKFPGSLYYPLTISSEYYKFDEnNPL    TREN   KAKIQELKQ-MIKC---DV--T  3513
Query_LOBOS  3426  LDKP-I   GACVMPILKSIALQYPNALYYPLTISAENFAF-EqGPA    GEKC   MQDVAFLKK-TVFS---PL--K  3487
Query_RACEM  3414  LDTS-L   AHCVFPILHKVAKSYPGALYYPFRISEEHYSA-IgSRL    EKKN   REEIEKIKQ-TIYS---PL--M  3475
Query_PVIOL  3536  LELP-Q   GPLVLPILLEIAKYYPQAIYFPYKISSEQFT-----PV    ALKL   ---SLPLSN-LLQH---PL--L  3590
Query_BOMBU  3251  LGT---[2]APLIIPVLKRIVETYPNAVIYTFRLTMETNPALL----    NE--   -TSTYEIRQ-ILYN--RPE--I  3307
Query_AEDES  3392  VDFS--[1]QSFLEALLTRIATSYPMALYYPAKMSLQRHQLRP----    DN-V[1]SPFANRLMQ-LL-D--IPK--L  3450
Query_ANT__  3305  LETPfI[7]AIAIIPIVHRLAETYPDAVVYTYYLTIERNPSIL----    QN--   -EEIQRIRK-LLGD--KAEe-Y  3370
Query_WHITE  3382  LYSP--[1]VAAVSDLLLRLAEKYPNALIYPFRLSKEMQAKVS----    S--S[1]QTIVDSLGT-LLDC--DYG--T  3440
Query_BEETL  3207  IDSS--[1]VTVIKPLVLKIATTYPQAIMYAYRLSKENYKCEQ----    R--E[1]KELIETLDE-LLLT--DEG--A  3265
Query_THRIP  3293  LDSP--[1]AKLVAPLLLKIAETYPNAVCYPYNISKEKNSSST----    VKLD[1]IKKLDRLLEdSSFQ--IPL--V  3354
Query_TERMI  3382  LGTP--[1]GPVLHDIILRLAETYPQSLIYAFQLSREKYNLES----    GHVS[1]KALVDQLAS-LLFT--DPV--I  3442
Query_FLEA_  3310  LTSN--[1]SETIGPIVIRVAESYPDALHYPFRLSILNYKCLA----    GS-Y[1]EYCVAKLKQ-VLMK--NSL--I  3369
Query_MIDGE  3409  YDFE--[1]DCYLDELLMRIATKYPNALYFPFELSFNNHNKIT----    TK-M[1]ENVF-KIRE-TI-S--NPT--I  3466
Query_LAODE  3357  IDTE--[1]GPTLEPVLLRLACEYPSALKYPLNIRSTNFTSPT---A    --KN   --TFEKLKS---YApfDNLt--  3413
Query_LOTTI  3453  IDKT-E   SPAVQPILLQIGTEYPQAIAYPFRISCEGFEFED-SAQ    GKKN   KEAVKKLAS-ILTEdkVPL--V  3517
Query_APLYS  3544  LDKP-E   ADIVEPVLTKIAAEYPQALVYAFRMSFEGFKFGS-TPQ    EKKR   LEKVEKLAE-LVN---IPL--V  3605
Query_SCALL  3520  MDKP-E   APAVQPVLLKIATMYPQAVVYHFKLSQEGFQFGD-KPH    DKKN   KAMVERLAN-LLNEenVPL--V  3584
Query_CRASS  3478  LDKP-E   GSTVHHILTELAETYPQAVVYPFKLSSEGYKF-----T    NKKD   EAFVTKLGS-TLGDeqLPL--V  3538
Query_OCTOP  3415  LDKP-K   AYAIHDIVYNIACQYPQAIIYSFKISNEGFRFEG-NDS    GTQN   RKFVEKLSC-RLNS--KTLpqV  3479
Query_POMAC  3483  LDKH-Q   AHAVRPIILRIATEYPQAIVYPFHISQESFTFGS-TAE    DRRN   KQGIEELDA-LLSEerVPL--V  3547
Query_WHIPW  3354  FDSS--[1]GCCFTEVLNRIADDYPQALIHPLSLAVSLKGS-----S    TLCY   RKLMERLEA-NLHG--SEV--V  3412
Query_TRICH  3363  LKRS--[1]K--FYSILERIADDYPQALIYPLLMLKSATGDVG---E    TDFY   R-----LQL-KLRK--TKI--V  3416
Query_AASTA  3158  LDVHaT[2]NDWILSLLETMAKEYPKALYYDFRLSSASIPDMA----    ANAR   ---FHRLAA-LLQD---HT--L  3216
Query_AEUTE  3083  MSEKpS[2]NDFILTLLEMMAAEYPRALYYDFRLSQESIIDIE----    TNPR   ---FHRLSA-ALRD---VH--L  3141
Query_PLASM  3349  LN---G[2]GMTIVAILEKLAERYPVALFYDFMVTCRSSRDI-----    FKVD   ---LRRLEV-LLAN---PV--M  3403
Query_BREMI  3332  LN---G[2]RRTVLVILEKMAEKYPIALFYDFMVTCKSSLNK-----    FDVD   ---LHRLEV-LLTN---SV--M  3386
Query_OLIGA  3388  VN---G[2]GKSIVRILEKMATEYPAALFYDFKVTCAGET-------    -KTD   ---IRRLTT-LLSN---PV--M  3439
Query_LIGNO  3780  LDKPdQ[1]GSAVLPLLESVCDIYPAALVYPFRLARANYRFET---G    DSGA   RQAAERLWQ-RLRR----LplV  3842
Query_CESTO  3555  PLEG--[3]--LVDDIIRRIAAEFPQGLSLPFRVLLTSFCDWNlVGQ[13]GKGD[1]ASFYSSIRS-VFSK----LvnL  3632
Query_OVI2_  3984  LFATaT[3]-VLVNQIIDRVAVDYPQIMLPPFLVKTSALLDDHwSST[14]---N[3]RPTVHRLFR-LLSE--SQVy--  4064
Query_PRIAP  3522  LDKR-E   ARAVQGILLSIAESYPQALIYPLNLSREHYTFTV-DAA    GAEA   EAYLHRLLA-TLST-cMDMsqL  3587
Query_ROTIF  3445  LDKP-Q   AKAIQKIIENISIEYPQALVYPFRMSTENFKFSN-EL-    -NEQ   SDFVQKISR-HLNT---QLplV  3506
Query_SPONG  3528  VDKP-E   SRAVHDILIRIADTYPQSLYYAFRMSSSGYKFGH-N--    NSEG   KELFEKLKE-KLQS---QL--M  3587
Query_TARDI  3476  LDKP-E   GNAVHDILGRMASEYPQAIIYGFKISKENYVFDEkDSR    GRRN   KKFFDELSE-KLDI---PL--I  3538
Query_TRICH  3497  LDKP-E   ARAVQEILFRIADEYPQAIIYAFKVSNEQYTFEE-TSV    GRQN   KTAVSKLSS-KLQL--P-L--I  3558
Query_OSTRE  3331  VSDT-Q[2]-AVTSPLTRRLLSKYPAAVRPHFNLMKETINRDI----    ---S   ----ERISF-ALAS---SLh--  3383
Query_NITEN  3657  MEQP-E   GRVLARVLERVAEEFPQALYYPFNVSQPAFG-----DV    ARER   ---TSRVAR-LLHN---PL--L  3711
Query_CLUBM  3284  MDSR-E   GGLFIGILESIAQHFPQVLYFPFHVSKDQYG-----TV    GLDR   ---TRKLAA-TLKS---PL--F  3338
Query_MOSS_  3441  IEED-S   GRVLIKVLESIAKSYPQALYFPFRVSETDFG-----AI    GRQR   ---TQHLKI-LLRN---PL--L  3495
Query_LIVER  3495  MDQD-E   GNILVPVLENVAKLYPQALYFPFHVSVKDFG-----PQ    GRRK   ---TRRLAT-LLKS---VS--L  3549
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3529  DRLIFSLEMMHEPAVIIKDWIM-DM--KNAKGPggSSQ[1]LS-MAYV---KERYQELYK-ELLCIESED[10]----[2]  3600
Query_LINGU  3539  ETFINALEQFGQPNALFKDWIG-EM--AKLL----HVK[1]RN-TDAV---QRCYTQMYS-SLMMTQRQA[ 9]-TEP[5]  3611
Query_LIMUL  3596  DDFIKALGHLGVPYGVFKETCR-NI-----TSL-kKFK   EN-SEKI---QKEFQEMWQ-DIFDAGNQD    NSGP[2]  3656
Query_SPIDE  3528  TDFVTALDHLSVPSVIFKDYVN-EL-----VQS-vSAK   TS-ADKV---KEIFTRIND-KLFNPNAAE    KNRR[3]  3589
Query_SCORP  3551  DEFVNALDSLSVPELLYKEYFK-EI-----LQM-kKDK   GD-KNEI---VKLYKKLCH-ELFNYNPRD    KMK-     3608
Query_TICK_  3381  SEFIEALELLQFPDVAFKNWCE-SM--REALN---NQK[1]VD----V---RKLYADGVR-QLFGGQGRD[ 3]SSGG     3441
Query_MONOS  3633  ---------------------E-----KSALLI-----   ---EKRL----RFGHKVFGsQAMKRRATD    ----     3661
Query_ROSET  3545  RRFVTELKCLTNPVMAYMDWLE-EQ--DQAIKK-----   RDqDALL----TAFKHLKS-ELLVTSSTT    --AS[4]  3604
Query_BELCH  3604  DKLIGGLEQLSHPDQLFKDWGD-DAdlKRLL----KVD[6]ED-KVTI---KRLYDEMYD-SLLDFRSAS[12]SSDL[7]  3689
Query_CIONA  3570  DEFVRQLRRICHPSIAFKDWKE-DF--TNLFCS--KET   -S---RM---KHIVENYQN-MFNDFFQCI[13]SPTS[3]  3643
Query_CAECI  3588  QDFIQSLEQLSNPDMIFKDWIE-DM--KNELA---KSQ   RN-KNNI---KKLYEEMYM-NLGNLK---    APGI     3644
Query_HUMAN  3569  QDFINALDQLSNPELLFKDWSN-DV--RAELA---KTP   VN-KKNI---EKMYERMYA-ALGDPK---    APGL     3625
Query_MOUSE  3569  HSFINALDQLSNPDLLFKDWVS-DT--KDELG---KNP   VN-KKNI---EKLYERMYA-ALGDLR---    APGL     3625
Query_XENOP  3586  QDFIRALEQLSNPPMIFQDWWE-DV--SNELS---KPN   VN-KNKI---KELYKEMYT-NLGNPK---    DHFM     3642
Query_PLATY  3576  QDFIQALEQLSNPEMLFKDWVE-DI--KIELT---KTP   IN-KNNL---KRMYEGMYK-TLGNLK---    APGL     3632
Query_TASDE  3581  QDFINALEQLSNPELLFKDWVD-DV--RGILM---QTP   VN-KKKL---EKMYEEMYA-ILGTLE---    VPGI     3637
Query_CANAR  3577  QDFVHSLEQLSNPVMIFKDWAE-DV--RNELS---KPQ   RN-KKKL---KEMYERMYK-NLGNLE---    SPGL     3633
Query_OSTRI  3547  QDFVHSLEQLSNPMMLFKDWTE-DV--RNELV---KTQ   RN-KNKL---QDMYERMYR-NLGNLE---    APGL     3603
Query_TURTL  3586  QDFVHALEQLSNPVMLFKDWVE-DV--KNELG---KTQ   RN-KNNL---KEMYEGMYR-NLGDLQ---    APGL     3642
Query_GECKO  3587  QEFVQALEQLSNPVMLFKDWTD-DM--KNELG---KTK   RN-KENI---LKLYDGMYQ-NLGNVH---    ASGL     3643
Query_DANIO  3564  QGFVDALQQLSNPEMLFKDWWD-EV--KNQLD---KPN   LD-KKKM---KLQY--MTE-LLGDAK---    SPRF     3618
Query_SNAKE  3572  QDFAHSLEQLSNPAMLFKDWIG-DV--QNELG---KTK   KN-TNNI---QQLYDGMYQ-NLGNLE---    APGL     3628
Query_STERL  3591  KNFIDALQQLTNPEMLFKDWVE-DV--KKELE---KPK   TD-KKLL---KNMYEEMYS-ALGNPR---    CQGI     3647
Query_MILII  3587  EVFINSLEQLSNPDMLFKDWVD-NF--RSKLE---K-K   EN-KTQL---KNLYSEMYS-NLANAG---    APGI     3642
Query_BAMBO  3577  ESFIHSLEQFSNPDLLFKDWAN-DF--KTELE---K-K[1]KD-KKLL---KNMYDEMYS-NLASLQ---    FPGI     3633
Query_ALLIG  3544  QDFVQALEQLSNPEMLFKDWVD-AV--RNELG---KAQ   KN-KAKL---KEMYEGMYG-NLGNLQ---    ASGL     3600
Query_GAR__  3588  KDFVEALQQLSNPEMLFKDWLE-DA--RNTLS---KPK   TD-KKDA---KKMYEEIIS-VLGNAK---    AAGI     3644
Query_TETRA  3771  YNFMYHLNQITHPEHRLIYYLE-YL--KSQFLSfqKIH[6]KD-YVKLnkaVEEYQKIIDrILKETEEDL[ 1]QHNI[5]  3849
Query_PARAM  3414  --FIEALNQLTHPEHRLQAYLN-QI----------KDN[1]EDiDTVI---DKIQEDFFS------KRED    ----[2]  3461
Query_SLEMN  3652  NAWAEALDCLVYPEHRFKYWLQ-II--NDIVSD-----   -DnMRGIqkaK--LKDIII--MMAYDLTC[ 6]--G-[3]  3715
Query_STENT  3086  DNFIEALDILTYPEHRLKSFLD-LI--KDSIMQ-----   -NsSSEY---------------INYLFNE[ 6]--GI[3]  3139
Query_NEMAT  3558  DSFIAALEQLHNPEMAFKDWCDgEI--KQLLQ---TKN[1]K----AI---KEAFRSIYK-QLFDSSVNR[ 6]PAAE[2]  3624
Query_PISTI  3597  DDFIAALEMLTNPEMVFRDWCEgPM--KQLLE---DQI[1]RNdSQAV---KQCFGAMFQ-LLMDCR-KR[ 6]TRGT[2]  3666
Query_CORAL  3582  DDLIQALEQLTNPEMVFKDWMD-EV--KVLLS---TKG[1]A---VKI---RESYGEMSK-KLLNVK---[ 5]SDGL[2]  3644
Query_HYDRA  3412  DEFAKALYQLSDPVSLVKDFLE-QI--KGYVSN--GSK   -D---EI---RKCYKS-----INDLLLSA[13]----     3474
Query_LATIM  3585  EDFIHALEQLTNSDMIFKDWVD-DV--RNELD---KPQ   RN-KANL---KRMYEEMYA-NLGNLR---    APGL     3641
Query_SHRIM  3515  HQFVSAMSLMVVPSIACKDAIQ-NV-----LQE--KDK   QKmEEGL---SEIVSEFL--KVQNRSSKG[ 1]AEEK[7]  3580
Query_MCRAB  3508  HHFVTAMSLIVVPENTLKKSFK-KI-----LLI--KNK   EDiEVAL---NNLVADHL--KVNNRSTRG[ 1]AGEK[7]  3573
Query_AMPHI  3553  QPFLEAITLVTEPFGDVKVDVR-NA-----ERL--PSS   QSkANKL---KYIKEQFFA-NVTSLALQQ[ 1]ETQR[8]  3620
Query_BARNA  3323  ------MANCRLPERLVAELGR--------LGV-----[1]----SRL---RSLCRHLQV---GQLEKAA    --GA[4]  3367
Query_COPEP  3349  DSILQGLGLLVFPSVAINDLLK-SC---KELQS----N[4]--wMQKA---KEEYKDFKV-RYLSKDPDV    ----     3405
Query_DAPHN  3161  EMFLKSIDYLKPPEKAARQLLNsETy----------AK[3]------I------FQSKYL-NLPPPTARS    ----[2]  3209
Query_STARF  3614  DDFICALEQLNNPEFVFKDWAD-DV-mKPLL----NVA[1]RD-KNKI---RQACHEIYM-NLINHRSTV[12]SSPS[3]  3689
Query_STRON  3605  ENLTLALQRLTHPGFIFKDWIS--V-iNSKM----NSK[1]RD-HPAI---KQEYQQLYQ-QLFESGGR-[10]VEP-     3672
Query_SEACU  3578  DQFTRALDQLNHPGVLFKDWAD-DE-lKPLL----QAE[1]RD-NVKI---KKAYSQMED-KLLNF----[ 6]GHDL[4]  3644
Query_FUNGI  3433  DDFISSLQKLTHPEHRFKDWGE-IL--KSLLKA---SK   RS-QQAI---KSTFAEIYA-DLLDSTDET[ 1]----     3489
Query_DPURP  3599  DRLVIEFARLTHPEHRFKDFME-EM--KSLIKS---PT   KD-ISAI---SILNSEIYD-DSFNP--KT[ 1]S---     3654
Query_HALBU  3592  DTLISEFERLTHPEHRFKDYME-TV--KTMMKS---TP   IN-TSEI---QKLGVEIYN-DCFNP--HT[ 1]R---     3647
Query_GLOBO  3332  DKLAVELRRLEDPAHIFKDLCE-RF--ESLISS---SL   PA-KQAS--iKKAYEEFKL-ICLET----    -KKN     3387
Query_PALUS  3395  EKFKFELQRLTEPMHIFKDWLE-RT--DSLIRS---GR   PD-KEAA--iVAAFEEMAA-YCLTQ----    -NSD[1]  3451
Query_CALIF  3498  ENFIMELQRLSEPIHIYKDYAE-KI--AEVLSS---KE   KS-SKKIenlKEYSNVIHE-YLIEP----    -PSN[2]  3557
Query_GLOMU  3514  DTFVTELHRLTDSEIHFCNWIN-PV--YNLVKR---NY   SG-VDKTaeiKASFQKMSS-FLLKP----    --NP[2]  3572
Query_LOBOS  3488  EDFIFELRRLTNPDHLIKDWLE-QT--VALFQN---KT   RN-VE---qiMALYQDLHR-LVLDV----    -GNP[2]  3544
Query_RACEM  3476  EDFCTELLRLTNPDQLLKDFVD-FL--RSVAMN---TE   TD-KEY---iQAMYRQFSE-LLLSP----    -TSS[3]  3533
Query_PVIOL  3591  DSLIQEFLRLTHPEHRFKDYME-EM--KALLKM---AP   IN-SQLI---CKLGTEIYN-DCFNQ--KT[ 1]K---     3646
Query_BOMBU  3308  DRFLTALHYVVKPELYLRHYLI-EL--RKNLSL--GTA   -TaIDIF--lKKVYPNLRE----N---KH[ 2]KP--     3363
Query_AEDES  3451  DRFVSELSQVVIPSMKISKIIT-DL--RMA-RC--DSN   -EeYRKL--vHDLANEAFP----EISPD-    ----[1]  3504
Query_ANT__  3371  ERFLQAIQYVVQPKLYLKHYLD-NA--IKDLSQ--GKM   -TaIKSL--lEKVYPNSST----HAAEEK[ 2]QP--     3429
Query_WHITE  3441  EKFLKAMACLCLPDVKLKSFLG-RL--KTSLEA--NDG   -KcAVCIgavLEIFDEAEI----------    ----[1]  3492
Query_BEETL  3266  NKFLKALSFIGIPAAILIYYVR-KI--VKLIAS--SEL   -GqAKI---vHDFVLKEMF----IKADSR[ 2]DLLQ[8]  3333
Query_THRIP  3355  ----KALSCLTYPSIILKDYTN-KL--TEAMKS--API   -NwL----slKEICDEFSY----IHKESG[ 2]SQCA[8]  3417
Query_TERMI  3443  DIFLRAFSCLGQPRKMLYHFVE-KL--EQLLNA--KRP   -NcD----slHHMYDTMM-------KE--    ----[6]  3496
Query_FLEA_  3370  DRFIEAISCICLPEILLINALG-KL--TDKNYT--DMK   -NlKRML--fQ---------------QNL    ----[4]  3417
Query_MIDGE  3467  DMFMKCLQCLVVPEKMIQAHFK-DF--EFEIES--KLP   -EmNTKM--fQDAVKVFYK----NVFENN    ----[4]  3525
Query_LAODE  3414  ANFVQAIKYVILPETLLFNCLI-DI-qKSILCK----E   SDvETKWnelERLTNNLFS----DTAPNT    ----     3470
Query_LOTTI  3518  NKFISALDQFGQPEMLFKDWTD-DI--KKML----QSK[1]RN-KSKI---MEKYKEIYN-MLIEYKETQ[10]TQTT[4]  3591
Query_APLYS  3606  NLFISALEQFEEPDIVFKDWGK-EM--RKLL----NAQ[1]RD-KGEI---KKVYLELYQ-SLLEQPASG[11]SSTV[5]  3681
Query_SCALL  3585  SKFICALEQFGQPDQIFKDWGN-DM--LRML----KKK[1]QD-KVKI---KELYQEMYR-QLLEPSQ--[14]TSSQ[5]  3661
Query_CRASS  3539  PKFISALEHMSQPDQIFKDWFD-DT--RLKF----MKN[1]PD-KKAM---KEEYQIMYK-KVLEFTETV[10]EETL[5]  3613
Query_OCTOP  3480  FTFISSLESFAEPLNVFKDWFK-NA-----LQLysKTN   PK-KQEV---IELYKDMCS-SLFEEN---    ----[1]  3533
Query_POMAC  3548  RTFISALEQFSQPVQMFQDWSR-DI--TKIL------E[1]RP-AAEI---KKKYKEIYD-MLLEMPHKG[ 8]QTSF[5]  3618
Query_WHIPW  3413  HEFTKSLVLLSHPDVMLRDAFE-EL-eCVTLGHahEMP[4]LG-REMI---AKISARLLL---SDYKPDA    ----[1]  3475
Query_TRICH  3417  KEFIDALNLLSHPDI---HAID-EL-eSIILNIsnNSN[3]---KEAI---AKVYSFVTF---ADIRRDF    ----[1]  3473
Query_AASTA  3217  NQFVAALEGLHHPEIRLKEALR-YL--SDVVDD--KKP[1]AQlASIV-------STTLS-AVFDDRTRV    ---L[4]  3276
Query_AEUTE  3142  DQFVAALHGLHHPEIRLKEALR-YL--SDVV-E--SKP[1]KEaSKII-------AATIS-HVLDESSKI    ---L[4]  3200
Query_PLASM  3404  EKFVAALGLIHHPELRLKEGLR-KI--ANLL-D--ENR[1]SDaRQTV-------ELLWR-DCFSPQRQL    ---L[4]  3462
Query_BREMI  3387  EKFVAALRLLHHPELRLKEGLK-EI--SELV-E--DNR[1]QDaRHKV-------TRLWK-DCFSLDRPL    ---L[4]  3445
Query_OLIGA  3440  EKFVAALRLIHHPELRFKEGLR-EV--SKLL-E--SGK[1]SEaKAKM-------AHVWQ-DCFASEHPL    ---L[4]  3498
Query_LIGNO  3843  EEFIAQMGEVCLPACTALDWLS-SA--KRSIGQ---TE   -----RL---RELHAEAIDqLFFDPDGPQ    -AGS[1]  3899
Query_CESTO  3633  NNFIDQLEYFDEPYLVFKDWVFnHA--EKVIHS---ST[3]ADlANAC-------NALLKgGFLRIAPET[ 1]SEGG[3]  3698
Query_OVI2_  4065  ARFVEELSFLEEPKLVLKDWLAsRA--KRTLSN---SE[1]-NsAARL----ALVRDYTNlLLRSFRRST[13]ASGS[3]  4142
Query_PRIAP  3588  QQLVASLEQLSSPHMLFSDFVA-DA--RPLL----MTR   ---RADF---PARFRKIYA-SLFDTTRR-[ 2]EEPG[2]  3647
Query_ROTIF  3507  EQFITALEHLNTPKLFFNDLIM-DI--NNNLK-----K   -N--DVI---MKLFKDLNN-NLIDLEGKI    ----[3]  3561
Query_SPONG  3588  DEFIAALEQLNSPELAWKDWYDnEL--KPLVGSvtVPK   RD-KNAI---KRVWEQMYS-QLLETRSGL[ 8]SPSL[2]  3661
Query_TARDI  3539  DDLIRHLTRLCEPHTFFQDFTK-RA--DYLLKY--KKK   -D---QL---KQEYAEINQ----------[ 3]RQGS[4]  3594
Query_TRICH  3559  ESFVRALEQLSQPDLVFKDWVD-EV--KGLKK---KSG[1]I--AESI---RICYQKLRDgLLVDASNES[ 6]TSGE[4]  3629
Query_OSTRE  3384  DGFTRAIELLHFPANRLSFWRQ-QL----YLREcqHDK[3]--------------NDLIEvMLLDLAVVD    ----[4]  3438
Query_NITEN  3712  DQFVTALSDLSHPNLKLKDWLS-EI--SMRLQR---R-   -D-NQGA---QEAFREMYR-DCLDA--RA[ 1]AAAE[3]  3771
Query_CLUBM  3339  EGLVRALEDLTFPEKRLFDGLC-EV--KSCLMS---G-   -N-RTKA---RETFRAVFA-DCLDV--EA[ 1]RNET[3]  3398
Query_MOSS_  3496  EGLVRALEDLTFPQQRLKDGLL-QI--LNFLFT---N-   -D-SERA---QLVFNEIYQ-DCLDT--QS[ 1]KDPV[3]  3555
Query_LIVER  3550  EGMVRGLEDLTFPEQRFRDGLS-QM--KNCLIA---G-   -D-KERA---LAVFAEIYE-DCLNV--NV[ 1]RKSL[3]  3609
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT   3601   GPYRSDFVKQFIS---QFDEYFG   KMGSKIPGLTF--MKVLSW---LSDKSEKVKK--AFSN------[2]PRKLKEY   3660
Query_LINGU   3612   GQYRQKFAKAYCK---DLDSAFG   KDGKKIMTMSE--EAADGI---FKKMDLASKG--TKG----HPN[3]SEKLGDY   3674
Query_LIMUL   3657   GLVFINFEKKYKK---KVVDLCG   QNGNRILSMST--KEFSEKS---NRLVEEMKKtcLNVG-----S   -VRLKDI   3716
Query_SPIDE   3590   GPLHKKFSEEFKD---KFLAICG   KDGCKLMKEKE--SKVALKAl--SDLVAEAKKkdKKLS-----S   ---LKDA   3648
Query_SCORP   3609   -------ERKLHL---KVAKVCG[8]KDGNKLISMPL--KEASVKY---ETLLKKITTmsEMNK-----E   -AKLRDF   3669
Query_TICK_   3442   GIVHRHFSDVVKK---KVTDAFG[1]-EGSKLVKMDA--KKFGSEC-------QVILK--------SYKS[2]PTALKDL   3497
Query_MONOS   3662   ----PKIISTLHA---KLEQWLS[2]---DKLKSMTMq----------KSMVKELGKcaEGLAdAAKDS   -STLSTY   3715
Query_ROSET   3605   GSAWKAAAKKASR---NFMDQFG[1]-DGSKLLRMDA--KAYAKVKk---VLKAKLREy------DTKKP   PRTLKEY   3664
Query_BELCH   3690   GPFRRRFAQQFAK---ECEKMFG   KDGSRILDMKE--KEFMKLKsgLQNMMGDFYR--KAFP-REKAN[3]PGLLKDY   3758
Query_CIONA   3644   GLIWTKFSKNYKA---KVERTFG[1]QTGDKLSRMNL--KQCRGE---INNLYSDIEKf-TLGT------   -SSLNDY   3702
Query_CAECI   3645   GEFRRRFVQAHAK---DLDTHFG   KGGSKLVDMKE--DAFAKT---AYSL---FNN---M-RgS-QPE   PGNLKEY   3702
Query_HUMAN   3626   GAFRRKFIQTFGK---EFDKHFG   KGGSKLLRMKL--SDFNDI---TNML---LLK---M-N-KDSKP   PGNLKEC   3683
Query_MOUSE   3626   GPFRRRFIQAFGK---EFVKSFG   NGGSKLLTMKV--DDFCKI---TGSL---LVR---M-K-KDSKL   PGNLKEY   3683
Query_XENOP   3643   GAFRRRFCEKYTK---DFDKAFG   PEGSKLLNIKC--DGFNKT---VGPL---ITK---M-KeQ-QKE   PGNLKEY   3700
Query_PLATY   3633   GAFRKRFIQAFGN---EFDHHFG   KGGSKLLQMKL--SDFTSI---TNSL---NSK---M-K-ESSKP   PGNLKEC   3690
Query_TASDE   3638   GAFRKKFIQTYGK---EFDRHFG   KRGSKLLVMKL--SDFNNI---TTTL---LNK---M-R-EDSKP   PGNLKEC   3695
Query_CANAR   3634   GMMRKRFIQAFGK---DLDHHFG   RGGSELLNMKT--SDFDAI---ANSL---HSK---MIK-T-HKE   PGNLKEC   3691
Query_OSTRI   3604   GLLRRRFIQAFGK---DFDHQFG   KRGSKLLNMKA--SDFDAI---TSAL---LSK---MTK-S-HKE   PGNLKEC   3661
Query_TURTL   3643   GLFRKRFVQTFGK---DFSSHFG   KGGSKLLDMKA--DDFNRI---AFSL---LSK---M-K-D-QKE   PGNLKEC   3699
Query_GECKO   3644   GLFRKRFIEKFGK---EFDSHFG   KRGSKLLDMKA--SGFNTI---ADSL---LLK---M-K-ECHQE   PGNLKEC   3701
Query_DANIO   3619   GSYRRKFIQKFSKeveKL---LG   AGGSKLYERRK--D---------KDFLQQVDR---MVQ-S-MRG[4]PGNMKEY   3677
Query_SNAKE   3629   GWYRKQFIKEFGK---ELDIHFG   KGGSKLLGMNA--SVFIKV---ANSL---SEK---M-K-KCEKE   PGNLKEC   3686
Query_STERL   3648   GAFRKKFAQHFSK---EFDKTFG   KGGAKLFDMKH--DAFRKA---SDSLWGSISE---EQK-K-EKE   PGNLKEY   3708
Query_MILII   3643   GAFRRRFVTGFKK---KIDQHFG   KDGLLLLNMNL--KAFNVA---VGQI---MET---QK----PEP   PGNLKEY   3698
Query_BAMBO   3634   GAFRTRYVQVFKK---AFDQYYG   KEGSRLLDMNI--KIFDEA---TKKI---LETvcKKQ----QAP   PGNLKEY   3692
Query_ALLIG   3601   GLFRRRFIQAYGK---EFDTQFG   KGGSKLLDMKD--NEFIKI---VNAL---LAK---M-R-DGQKE   PGNLKEY   3658
Query_GAR__   3645   GRFRKKFVETYAKrieEIVKELG   KDGSKLWDKRK--D---------EKFWQKLTT---SIR-S-MDS[3]PGNLKEY   3705
Query_TETRA   3850   GSYNIKFSKDHISefkKLVKKFT[3]-KGDDQDKKEAy------------DFIKKFRDkvNTKDlKSGYD   --KLSNF   3911
Query_PARAM   3462   GSYNLAFIQKKKNql--------   ---NQILKNKAd-------ShtINQFAQEISEeaKQIQvKNQQM   ---IHDM   3514
Query_SLEMN   3716   GSYNRKFAELWDK---VFAKEFG[1]-NLNRVDKMSN--SEIIKSV---QNLNNKLNQs--IGNlGQGYE   --KLSTF   3777
Query_STENT   3140   GVYNKKFAVDWEP---TLSGFFG[1]-GFTNINRMST--TELLQAI---ETIAKKLANp----SlPFGRE   --KLSTF   3199
Query_NEMAT   3625   GPFRKKFSKEHAS---RFENIFG   KDGEKLVNMSL--KTAK--------LKQQMV------S----KP   PGNLKEY   3675
Query_PISTI   3667   GPFRTKFAQEFAN---EVQRIFG   KEGEKLVGVKM--TDFSKW---CRDIQQKMVKf-RNRT----PP   PGNLKEY   3727
Query_CORAL   3645   GSIRRKFAQEFSG---NVVKVCG   EDGSKIIDMSI--KAFESA---MKTTNEKMKA-----K----KF   VGNLKDY   3701
Query_HYDRA   3475   --VFIEFAKKFSS---YVESLFG   KDGEKVLKMSL--NEFQEA---TQQINEKLKEy-KSIN------   -ASLSDY   3530
Query_LATIM   3642   GSFRSNFVKAYGK---EFERVYG   KRGEKLLDMKS--SDFFKV---SDQL---LTS---MRK-F-QRE   PGNLKEY   3699
Query_SHRIM   3581   DRAKKQMAEAFTK-------AFG[1]---------NL--KKAKSLS--IRDIQRELGSmgRVLQ-LSPEE[4]PRQLKAY   3638
Query_MCRAB   3574   GKLKKDVNVALTK-------EFG[1]---------KM--EEFKRMP--VEEIKKRLNSirTVVN-QRRKE[2]PNQLKAY   3629
Query_AMPHI   3621   DKLK----LVFKK-------YFG[1]---------KG--SEMKSLK--CTEIDRCITEllSCWS-SSLND   ---VKEL   3667
Query_BARNA   3368   GGLAAVLSDPPSEadrQLLRLCG[1]-NGERITTRSAp--------kaVDLIKKRLAEr------PASRR   ---LGEF   3424
Query_COPEP   3406   -------ESAFKT----LEKTFG[1]----KLDKAFAg-----------SKLLQDLKEiyKEADaKFRYP[3]PSQLKTF   3457
Query_DAPHN   3210   GHVHYSFSQAFTG---KFQETLD[2]-------RMKS-------------LLNSLEFf-KKKTfPGQKS   ---LRDY   3257
Query_STARF   3690   GNFRRRFAQEWQK---KFEAAFG   KDGTKVIDMQL--KTFSEH-------YGKLYK--EMAA-KRRDP   PGNLKEY   3748
Query_STRON   3673   GPYWKKFGREWQK---RFDSSFG   QDGSKLAAMDY--KTWSKL-------RQDICA--KMTP--GKDD   PGNLKEY   3730
Query_SEACU   3645   ANFRTSFARKWLK---KVEKVFG   KGGEKLLKAIT--NDFTTV-------VANMYA--EMNP-RGAKD   PGNLKEY   3703
Query_FUNGI   3490   GSYNKKFATTWSR---KVTSAFG   KDGANLCKMTA--EDLQSHL---SEFLEPMNR--DLNP--KKTA   KMKLVDF   3551
Query_DPURP   3655   GDYNIKFAKEWED---KYLKYYG   KDATILSRMDQ--KKFLEHF---AEMSSEMNK--NMRP--TSTA   TMKLKDF   3716
Query_HALBU   3648   GDYNLKFAKEWDH---HYSQHFS   KDGSKLAKMDV--KKLVELV---ADMSGKMAK--TMKP--SSTA   TMKLKDF   3709
Query_GLOBO   3388   GDIGRKFAEKHTA---KIANICA   -NGEKIDNNKL-----K-EL---VSYRDKLGS--EAES--KP-G   KNLLKSF   3443
Query_PALUS   3452   GEVGRAFASKHAA---KLFKILG   KDGSKLTNIKL--SEWK-EI---QKYYHQFIH--GKEA--LPQG   MVALKNY   3512
Query_CALIF   3558   GTIYMKFSKKYKN---ELKKVFA   L--SNLVKAKP--KTYMNIF---KKQLTTFRS--DKEFeRNETY   KTQLKNY   3619
Query_GLOMU   3573   GIFAKEFVNRHSE---YLFDICG   RDGSILARMSP--NDFKSKI---YDYYIKNIK--RQDR-KAQEG   VNLLKSY   3635
Query_LOBOS   3545   GSIAKAFAAKYGS---KMESLCG   KTGQKLATISN--RDFNSHI---LKFCRNEIL--SKDAsKKQSD   ADLLKAY   3608
Query_RACEM   3534   GVIPKAFATKHAS---QLRGLFG   KDGSKLLNLTQ--RDLK-RI---GDYQRNVKA--ESGR-----E   AGLLKSY   3591
Query_PVIOL   3647   GEYNLKFAKEWEA---KFIQHFG   ADGSKTLKMDV--KKLVEVV---GEMSANMNK--AMKP--TSTA   SMKLKDF   3708
Query_BOMBU   3364   GIIFNK-IASYKN---MIKELEN[1]----KTEEIKL------CVDriIERIKISLIKk-DNNK------   -NRLQDY   3416
Query_AEDES   3505   GREHQK-LIPLKP---EWLKLLD[1]----DFNQTEEihQHLNNLKekVTRL----------VPrQT---   TLELNKY   3558
Query_ANT__   3430   GDIFGKEIASYEE---KIRALAH[1]----DLDAARK------EILklKETLNESLWKh-MNKNvIYTNK   LSKLKDY   3490
Query_WHITE   3493   GNIYS----LIG----DLKKDVI[1]----LKKYVDK--GKLKEALkqTENI-MRAITp-CYHKqPSRDA   RISLKSY   3551
Query_BEETL   3334   KYEKD-----LK----NIGRLLT[1]----PSETP--------NVLkqLQKINTNLGNt-LQDRrENQRN   SRLLKEY   3386
Query_THRIP   3418   HKEHDI-PALFD----ELCAAIS[1]----RNTTVAS--EIINKVKktLEKV-KELDSk-------NIKH   YSKLKDF   3473
Query_TERMI   3497   GKIHG---ALFK----DI-QEYR[1]----RLEKAES--NIKCGNKkvLKDI-CFIKEk-LKYIiQDIKP   PVQLKDY   3555
Query_FLEA_   3418   GKAFEA-MTKFE-------NLFS[1]----YMEKPSP--ELLMEIKmlLKSL----------SNhRPKT-   ---LRDL   3464
Query_MIDGE   3526   GTSFSI-IHEFKA---KIREVFN[1]----KWNQKTK--SFLHELRfqIENL----------GKhRRGIE   TLEVHKL   3580
Query_LAODE   3471   GNFKGSLFIAFKDknqRMANQLR[2]-----LVKRPL------NVDndSYPLFRKTLEdsKLLEsERKEH   ---LDAF   3532
Query_LOTTI   3592   GEYRKKFADTFKK---DFDIAFG   KNGEKLASMSI--KKGRG------------------------K[3]PLSLHEY   3638
Query_APLYS   3682   GEVRRGFATSCKA---EFDKAFG   KNGDKLTSMNV--KEFQSM---VRKINEVINK--HKN-----KR[3]AANLKDY   3743
Query_SCALL   3662   GQYLRKFSARFKK---DIDKYFG   EDGRKLVGMKL--KDFFGI---FRKLGEDMDH--EVRGiPKNPR[3]PASLKEY   3728
Query_CRASS   3614   GKYPKKFAERFKK---DVDSCFG   KNGEKILDMSF--SKLAGA---YKKVYKDMDF--EV----KEKN[3]PANLKDY   3676
Query_OCTOP   3534   SPYYQKFTENFHS---LFVNSFG   KDGSKINKMNK--QELQKQL----LVVKEKLEkqKLKE-----P   -KALEEF   3592
Query_POMAC   3619   GEFHRRFAEVFKN---EFDNHFG   VDGSKLVTMTA--KEFSGI---LRNLQAQFEK--HKA------D[3]PTNIGKY   3679
Query_WHIPW   3476   GTVWK-----YKP---TLEQLFN[1]----QITSLDA--TWLEVTKrsMTEVIEEAKRmfDISR------[1]SDALCDY   3531
Query_TRICH   3474   GSFWKCYFKKYLP---KMQEIFS[1]----DLCAFDN--KAAENYLrlCSSFLSELRTfyIPEC------[1]PNNLNEY   3534
Query_AASTA   3277   GAYNRQWMAQHRK---HIEAMLG   KD---VTKAAL------QSAr--GWLSQTFQVm--PGKyGIDRH[1]KAKLADF   3335
Query_AEUTE   3201   GVYNRQWITKNRK---QIESMLG   KHGELANKSTL------QTAk--GWLSQSFQVm--PGKyGIDRQ[1]KAKLSDF   3262
Query_PLASM   3463   GRYNREWSRVAKR---DVEKIMG   KDGLKMTAKTV------SAAr--EWINNHFSVt--PGRyGITKD[1]KAHLGDF   3524
Query_BREMI   3446   GRYNREWSRKAKR---DVEKIMG   KDGSKMTAKSI------KYAr--EWIMKHFGVt--PGTyGITRD[1]KAHLGDF   3507
Query_OLIGA   3499   GHYNRDWAKRAKR---DVEKMLG   RDGSNVTAKSV------NAAr--EWIMNHFSVm--PGKfGITKD[1]KARIADF   3560
Query_LIGNO   3900   GRLRREFAARHRR---EFLAALP[1]--ADRLVGIRS--EAELQRA--VGKLWKRVRDerDSSS------[2]TARLADF   3961
Query_CESTO   3699   -TERQKSVETLSNl---IHANFG[3]KGLSTITPTEFd--------igVKRILSPYENs--------NPD[1]-SKLLNY   3755
Query_OVI2_   4143   GSHRTMLSERLMEac---NQVFG[1]-GGSQFVFASV------------NDIEQSIKQleLFLEtQQTPS[1]VNRLADY   4202
Query_PRIAP   3648   GRYWSVFTRTWAR---EVEEKCG   RDGSKLLAMAA--ADANQV-------LSALGK--EMKT-KAADS   PGNLKEY   3706
Query_ROTIF   3562   GRLWTDFSNFIKPl---FMKQFG   EQFRLLPQLSE--LEIKNK---LQNICSQVNEf--CQR----KP[1]SGNLGDY   3622
Query_SPONG   3662   GARQKTFAQAYAK---NIEDKFG   KGGAKIISMSG--TQLASA---FNEVYNKMRE--SLNKfSSSAL   ---LKDY   3722
Query_TARDI   3595   AATWKEFVEKFKQ---DFVYLFG   ENGQKIANLSS--SQWAKTK---EKWENDFKEygKKAR-----E[4]RGRLADL   3659
Query_TRICH   3630   GSHKRRFIQHFGK---VIEQYFG   KDGSKLQSMKA--SDFEKI---ASRLIQEISG---KIK----NR   PENLKEY   3688
Query_OSTRE   3439   GPINKRFVEIASP-------ILN[1]-----VVKSAL--KEKMSVSetLKKVEVQVLDewRKEFhGIENS[1]RLRLSDF   3500
Query_NITEN   3772   GDDLIAYARNFES---KIVSALG   KDGVKLARMSA--TEYKNAI---SSVSHVLDR--E--D--LAHG   KMQLSRF   3831
Query_CLUBM   3399   GEYNVKFARDYSK---PVVKALG   EDGSKLLQMDE--KKFATAI---AGVLTSLQK--AIKD--LPGG   KLSLASF   3460
Query_MOSS_   3556   GEYNLKFARDWAS---LIVKELG   LNGSKLRSMDA--KQFSSVI---SGLFEKLSK--NDIS--AEK-   -ISLGSL   3615
Query_LIVER   3610   GEYNLKFGRDWSK---RVIKALG   ENGSKLVKMDE--KKFMSAV---GDLLANLQK--AMLN--LQTG   VMSLGHL   3671
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3661  SQWLAEFQA-G    NHEASLEIPG-QYT   GEVC-PLPEYH-AKISSFDSQVTV    LRSLRMPKSIVIRGNDEKEY  3725
Query_LINGU  3675  SKWMSGLAA-Q    DMDLSLEIPG-QYT   GLSK-PMPEYH-VKIAGFDSRVLV    LSSMRKPKRITINGTDMKDH  3739
Query_LIMUL  3717  SLWLSDFHP-F    KYQDELEIPG-QYT   GTAK-PIPEYH-VKVSGFDQGVRI    LNSLTRPRTLTIRGNDEKEY  3781
Query_SPIDE  3649  SPWLNEFQQ-W    KYSSQIEIPG-QYT   GRSK-PLPEYH-VKIAGFDGKVLV    LSSLTAPRRITIRGNDEKEY  3713
Query_SCORP  3670  SDWLSEFHS-I    KYSEDIEIPG-QYT   GKRK-PIPEHH-VKICNFDQKVDI    LSSLTKPRCLTIRGNDQKDY  3734
Query_TICK_  3498  SPWLSRFSA-L    DRDNKIEIPG-QYT   GRSK-PMPEYH-VNIFGFEESVLV    LKSMQRPCRITIRGDDEKEH  3562
Query_MONOS  3716  APQLAESMFtG[ 3]-----LEVPG-QYG[3]--SM-PDVQSH-VMLCGFDSRVLV    LKSIRRPKRVTMLGSDEREY  3780
Query_ROSET  3665  SPWLSKYRSpG[ 3]----DIEIPG-QYG   GTEM-PIPEEH-AKIEGFDEKLLI    MASIRRPMRLTFRGSDEREY  3729
Query_BELCH  3759  CPWLSDFNP-L    DYDRELEIPG-QYT   GKVK-PLMQYH-AKIAGFDQRVKV    MSSLRKPKRVTIRGDDERDH  3823
Query_CIONA  3703  SPWLAKFQSlQ    RSSKELEVPG-QYD   GKSR-PLLEYH-ATITGFDDRVTM    MTSIRKPCKIIMRGSNEKDF  3768
Query_CAECI  3703  SPWMSNFKP-E    FLRHELEIPG-QYN   GRSK-PLPEYH-VKISGFNERVKV    MSSIRKPKRIIILGNDEKEY  3767
Query_HUMAN  3684  SPWMSDFKV-E    FLRNELEIPG-QYD   GRGK-PLPEYH-VRIAGFDERVTV    MASLRRPKRIIIRGHDEREH  3748
Query_MOUSE  3684  SPWMSEFKA-Q    FLKNELEIPG-QYD   GKSK-PLPEYH-VRISGFDERVKV    MLSLRKPKRIVIRGHDEKEY  3748
Query_XENOP  3701  SPWMSEFKP-E    FLRNELEIPG-QYS   GRSK-PMPEYH-VKISGFDERVSV    MASIRKPKRIIVRGNDEREY  3765
Query_PLATY  3691  SPWLSDFRV-E    FLRSELEIPG-QYD   GKGK-PLPEYH-AKITGFDERVKV    MASIRKPKRIIIRGHDEREY  3755
Query_TASDE  3696  SPWMSNFKV-E    FLRNELEVPG-QYD   GRGK-PLPEYH-AKITGFDERVKV    MASIRKPKRITIRGNDEKEY  3760
Query_CANAR  3692  SPWMSEFKA-E    FLRSELEVPG-QYD   GKGK-PLPEYH-AKISGFDERISV    MESLRKPKRITIRGSDEREY  3756
Query_OSTRI  3662  SPWMSEFKA-E    FLRNEIEVPG-QYD   GKGK-PLPEYH-VRISGFDERIKV    MESMRKPKRITIRGSDEQEY  3726
Query_TURTL  3700  SPWMSDFKA-E    FMRNELEIPG-QYD   GKSK-PLPEYH-AKISGFDERIKV    MESIRKPKRINIRGSDEREH  3764
Query_GECKO  3702  SPWMSEFKP-E    FLKHELEIPG-QYD   GKGK-PLPEYH-AKIAGFDERIKV    MLSIRKPKRIIIRGSDEREY  3766
Query_DANIO  3678  SPWLSSFKA-E    TLKNELEVPG-QYD   GKSK-PLPEYH-AKITGFDERVKV    MTSIRRPKRIIIRGDDERDY  3742
Query_SNAKE  3687  SPWMSEFKP-E    FLRTELEIPG-QYD   GKGK-PLPEYH-AKISGFDERIKV    MQSMRKPKRIVIRGNDEREY  3751
Query_STERL  3709  SPWMSNFKP-E    VLRNELEIPG-QYD   GKCK-PLPEYH-ARITGFDERIKV    MSSMRKPKRLIMRGNDEREY  3773
Query_MILII  3699  SPWMSNFKP-Q    FLGNELEIPG-QYD   GKCK-PLPEYH-AKITGFDERVKV    MSSIRKPKRLIIRGNDEKEY  3763
Query_BAMBO  3693  SPWMSEFKP-E    YLRNDLEIPG-QYD   GKSK-PMPEYH-AQITGFDERIKV    MSSIRKPKRLIIRGNDEKEY  3757
Query_ALLIG  3659  SPWLSEFKA-E    FWRNELEIPG-QYD   GKGK-PLPEYH-AKISGFDERIKV    MESLRKPKRIIIRGSDEREY  3723
Query_GAR__  3706  SSWLSNFRP-E    TLRNELEVPG-QYD   GRSK-PLPEYH-VKITGFDERVQS    LRSIRKPKRLIIRGSDERDY  3770
Query_TETRA  3912  SSFLQNFQAdH[ 1]-----IEIPG-QYT[5]--TE-PYKESH-IKIQSFDSQILT    LQSIRRPKRLCIYGSDEKCY  3976
Query_PARAM  3515  SRWFQNYNEeS[ 6]--QTEILIPI-QLS   GMNEfDIQMIP--KIASFDQGMLV    LSSIRKPKRIKIYGTDEKQY  3584
Query_SLEMN  3778  SEWLDSLDFnD[ 4]--SEYIELPG-QYE[2]-KSE-PLIEKN-VKIASVKKTLLI    LGSIRRPKRITVHGSNEKDY  3846
Query_STENT  3200  SEWLADYGL-D[ 4]--NS-IEVPG-IFT[1]-DQE-PFKRNS-VSITSFDQSILV    LSSAKKPKRIGIHGSNEKDY  3265
Query_NEMAT  3676  SPWLTNFQA-D    SHAQKLEVPG-QYT   GQSK-PMPEYH-VNIAGFDEKVLV    LASIRKPKRVTIRGDDERDH  3740
Query_PISTI  3728  SPWLSRFQS-N    NFTHTLEIPG-QYK   GKTK-PLPEYH-VKIVGFDEKVQV    LPSIRKPKCLIIRGDDEKDH  3792
Query_CORAL  3702  SPWLANFQT-Q    SETIQLEIPG-QYS   GQSK-PLPEYH-VKVAGFGEKVKA    MSSLRAPKRITILGNDEREH  3766
Query_HYDRA  3531  SPWLSSFQP-D    SFLDRLEIPG-QYT   GESK-PIVEYH-VKINGFDEKVLV    MNSLRKPKKITIRGDDEKDY  3595
Query_LATIM  3700  SPWMSQFKP-E    FLRNELEIPG-QYD   GKSK-PMPEYH-AKITGFDERIKV    MASIRKPKRLIIQGNDEREH  3764
Query_SHRIM  3639  SPWLANFQA-S    KYSEVLELPG-QYT   GMSK-PLPEYH-VKISSFDENVLL    MSSMRVPMRIIIRGDDEKEY  3703
Query_MCRAB  3630  SPWLANFQA-S    KHSDVLEIPG-QYS   GVSK-PLPEYH-VNISSFDENVMP    MQSLRVPLRIVIRGSDEKDH  3694
Query_AMPHI  3668  SPWLSSFQG-S    QYSDRIEIPG-LYS   GFSK-PNVEHH-PCISSVDNNLVV    LKSLRSPVVLKMHADDEKTY  3732
Query_BARNA  3425  SGWLASFQDvN[ 1]--QERLELPG-QYT   GRRR-PQPDYH-VRVANLSDEALV    LNSIERPVRITVRGDDAREY  3489
Query_COPEP  3458  SAFLANFLAsD[ 2]---QTVEVPG-QYT   GFSR-PDTTRH-ITISSFEPGLNV    FPSLRKPMKISMLGSDGNRY  3522
Query_DAPHN  3258  SPWLANYRIdD[ 3]---NLLELPG----[1]-KGK-------tLKVETFRDEIQY    LQSKQLPVRLTIVVNGAVEY  3315
Query_STARF  3749  STWFSDFQQ-L    SYDRELEIPG-QYG   GKSK-PLPEYH-VKIAGFDEQVKV    LSSMRKPKRIVIRGNDEHDY  3813
Query_STRON  3731  SPWFTAFQT-L    RGD-ELEIPG-QYD   GSKK-PLPAYH-VKIAGFDETVLT    LSSLRKPKCITIRGNDEMDY  3794
Query_SEACU  3704  SPWFLQYQ--L    TEQAQFEIPG-QYT   GVSK-PLPEYH-VKIAGFDEQVRT    LQSLRKPKTIKIRGHDQKDY  3767
Query_FUNGI  3552  SSWLQNFEQlD[ 1]GNGSSIEIPG-QYS   GTRI-PVPSEH-VKITGFDASVLV    MGSLRKPKRVKMLGDNEKEY  3618
Query_DPURP  3717  SYWLTEFDRsN[ 1]PTSKQLEIPG-QYN   GFGK-PQPETH-VTISSFDSNILV    MGSLRKPKRIKIHGNDERDY  3783
Query_HALBU  3710  SNWLVEFDRsN    -QLTEIELPG-QYH   GTSK-PQPEIH-IKISSFDPNLLV    MGSLRKPKRIKIHGSDEQDY  3774
Query_GLOBO  3444  SLWLSEYQIgN[ 1]DPTEYLEVPG-QYT   GTSE-PNISQH-VLISNFDPNVLV    MSSMRRPKKIIMIGSDEREY  3510
Query_PALUS  3513  SPWLTNFNArD[ 1]N--EELEIPG-QYE   GLAM-PRPQQH-VKVISFHSEVLV    MSSMRRPKRLIILGSDEKEH  3577
Query_CALIF  3620  SNWLSEYQGkN[ 1]D--REIEIPG-QYD   GKYN-PRLKKH-KMISIFNSNILV    LNSLRKPKRLEIIGDDESQH  3684
Query_GLOMU  3636  SAWLAEYSSsN[ 1]E--KEIEIPG-QYD   GLLSiPDPSKH-IKIASFDPKILV    LRSLRKPKRITILGTDGNEY  3701
Query_LOBOS  3609  SPWLHAFQAsD[ 1]D--GVIDIPG-QFS   GLTP-PNSQQT-ATIVRFDQRVLV    MSSIRKPKRISILGSDEKEH  3673
Query_RACEM  3592  SPWLSTFQAaN[ 1]N--ESIEVPG-QYS   GLSR-PYPEYH-ARIASFDDRMLV    MSSLRKPKRLKIYGSDGRGY  3656
Query_PVIOL  3709  SFRLAEFDKsN[ 1]SSSQEMEIPG-QYS   GVGK-PQPDTH-IKISSFDTNILV    MGSLRKPKRIKIHGNDEKDY  3775
Query_BOMBU  3417  SPWLCNF---P    ERDV--EIPG-QYT   GNRK-PMPQYH-AKILKFDSNIRV    MWSMRKPIRITMIGNDAKDY  3477
Query_AEDES  3559  SPWLAGYHFsE    REEM-LELPG-QYN   IDHK-PNVCNH-VKIVKVCTELEL    FKTLRNPIRIRINGSDGKSY  3623
Query_ANT__  3491  SPFLHKYVGsS    SDGI--EIPG-QYT   GEKE-PEPRYH-VRIVRFEPQVEV    MQSLRKPIRISIVGDNGKEY  3554
Query_WHITE  3552  SPWLASFIStA    YSST-IEMPG-QYS   GKMR-PDLKNH-IMIAGFKTEVSV    MQSLRRPIKLTMIGNDGNEY  3616
Query_BEETL  3387  CPWLANFSAsK    NNLE-LEIPG-QYD   GERI-PLVQHH-VKISGFCPNVSI    MSSLRKPIKITILGNDTKEY  3451
Query_THRIP  3474  SPWLAEFSSsK    KNIC-IEIPG-QYT   GDQK-PNPERH-IMISGFNDTVLV    MPSIRRPVKLTVIGSDAQSH  3538
Query_TERMI  3556  SPWLSEFQGgK    HCFLELEIPG-QYT   GESK-PLIQHH-VKIAGFKQTVVP    MKTLRLPLQVTIFGSDAKEH  3621
Query_FLEA_  3465  CPWLADFT-pQ    DSVA-LELPG-QYQ   GHKK-PLPHYH-AKIARFDQKITV    LLSQRLPVKIGIYSSDGKEH  3528
Query_MIDGE  3581  CNWFAEYKWcG    ENDF-IEIPG-QYV   GDRK-PFIEQH-VKIVRFETRLKV    FRSKQQPIEIKMYGSDGKSY  3645
Query_LAODE  3533  STWFNSYQSlP    ---HRIEIPG-QYS   GDNM-PVAADKiVKISKFSDRFTV    MSSLRKPVKISVVGNDGIEY  3596
Query_LOTTI  3639  CQWLADFNP-N    RDSRELEIPG-QYD   GLSK-PLPEYH-VKVAGFDERVLV    MSSLRRPKRITIRGNDEKDY  3703
Query_APLYS  3744  CPWMAEFNP-S    HNGRELEIPG-QYR   GLDK-PLPEYH-VRVSGFDERVLV    MKSLRKPKRITVRGNDERDY  3808
Query_SCALL  3729  CKWMVDFNP-N    TEGNNLEIPG-QYT   GLTK-PLPEYH-VKIAGFDQRVLV    MSSMRKPKRVVIRGNDEKEY  3793
Query_CRASS  3677  CQWLANFNP-S    IQGKDLEIPG-QYD   GVKK-PMPEYH-IKVVGFDQRVKV    MTSIRKPKRITIRGNDEREY  3741
Query_OCTOP  3593  SSWMANFDP-L    L--DTLEIPG-QYH   GFSK-PLLEYH-IKISGFDQQVQT    FRSLQKPRKIIIRGNNQKEY  3655
Query_POMAC  3680  CSWMANFNP-N    HESRELEIPG-QYD   GLRK-PMPEYH-VKVAGFDERVLV    MSSLRRPKRITIRGNDEKDY  3744
Query_WHIPW  3532  SLWMFRYSS-T    DSGIHLEIPG-QYT   GDHK-PLPEYH-VTINSFGAKLQL    ISSLRKPKVISIYGNDGRKH  3596
Query_TRICH  3535  SPWLAGYRP-H    PMDVFLEIPG-QYS   GDRR-PFPERH-VHIIGFASDIEI    MRSLKKPKKICIQGDDGRDY  3599
Query_AASTA  3336  SDWLAQLDPvK[ 1]----RIELPG-QYT[3]--GK-PEPATH-TYILSCEPQLMV    LPSKQLPKRLVLHASDERTY  3399
Query_AEUTE  3263  SDWLMQLDPvK[ 1]----RIELPG-QYI[3]--AK-PDPSTH-TYILSCDPQLMV    LPSKQLPKRLKLYASDEKTY  3326
Query_PLASM  3525  AEWLQEFDHsS[ 1]----ILELPG-QYT[3]--GP-PDPATH-VRILSFDSMLGV    LASKQLPKKLTAHCSDEKNY  3588
Query_BREMI  3508  AEWLEEFDHsN[ 1]----SLELPG-QYT[3]--GP-PDSSKH-VKILSFDSVLGV    LASKQLPKKLTMHCSDEKEY  3571
Query_OLIGA  3561  ADWLDEFDHiK[ 1]----RLELPG-QYT[3]--GK-PDLSNH-TYILSVGAQLGV    LVSKQLPKQITFHCSNQKEY  3624
Query_LIGNO  3962  STWLAEFAAtE    HGDAYLEVPG-QYQ[7]SLAK-PQPELH-TRIAGFGDTVRV    MSSLRRPHRLTIRGNDQRDR  4034
Query_CESTO  3756  SHWLANYSSdG    --RDYIEMPG-QRL[4]STTK------RlITIERIDPNVKV    LPSLRRPKLIRLLGSDGRWR  3819
Query_OVI2_  4203  SPWLSSFSSnD[ 3]--AQFLSVPSaQYS[6]-LAQ----RVH---IFRFGFSLNL[24]LVSLRQPKLISLVGSNAKHY  4294
Query_PRIAP  3707  STWLTDFHS-S    PATCQIEIPG-QYR   GRSK-PMPEYH-VKIAGFDEKVLV    MKSLRLPKRITIRGHDENEY  3771
Query_ROTIF  3623  SAWLKGFK--R    NITKDIEIPG-QYN   GNEK-PMPEYH-VKIESFDERILE    MSSLRRPKCITIRGDDQREY  3686
Query_SPONG  3723  SPWLDTFRQ-E    DYLQVIEVPG-LYD   GRSK-PLPEYH-AKITRFDAKVLI    LSSLRKPKRLTIIGDDTREY  3787
Query_TARDI  3660  NPWLAAFTS-L    QYGRALEIPG-QYD   GRER-PRPETH-VKITSFGDQVLT    LASIRRPRRIIIRGDDEKDY  3724
Query_TRICH  3689  CPWLSNFQA-A    NFTYSIEVPG-QYT   GRSK-PLPEYH-LKIASFGERLLV    LKSIRKPLRLTVQGDDEKEY  3753
Query_OSTRE  3501  SSWFELMHSsS    --EQVLEVPG-QYD[4]-----PPDASKhVTVMGFAPDVQI    FASKQRPKKIIMRGSDGQDY  3565
Query_NITEN  3832  SKWLADYDQsR[22]SDAGGIEVPG-QYG   GFCR-PDPGSH-VTIVGFDQTLTC    FSSIRRPKVLIIRGSDEKEY  3919
Query_CLUBM  3461  SKWMANFDQpR[ 9]-----IEIPG-QYD   GLQC-PDVSSH-ITLVGFDETLIG    LSSKQRPKILTMRGSDEKEY  3530
Query_MOSS_  3616  SKWLADFDQtK[10]KNEMQIEIPG-QYN   GSSI-PNLSTH-VQLLGFDQVLIC    LKSKQRPKILTMRGSDENDY  3691
Query_LIVER  3672  SKWMADFDQqT[22]SSEGDIEIPG-QYA   GFRE-PSPDLH-VKLLGFDQTLIC    LSSKQRPKVLTMRGSDEEEY  3759
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT   3726   KFLVKGGEDLRQDQRVQQVFNLMNEIYANDA-NCHAR   RL--QLKTFQVIPMTPRLGMIEWLKNSCTLKDFLRQ---G   3796
Query_LINGU   3740   KFLVKGGEDLRQDQRLEQLFYIMNRVFSQDP-DCKPR   QY--KLRTYQVIPMTTKVGLIEWVNQTKTLKEVVYG---T   3810
Query_LIMUL   3782   KFLVKAQEDLRQDNRIEHLFSVMNGIYAGDP-ACSFR   HL--SLATYKVIPLTTRLGLIEWVDNTVVLNQFLGD---M   3852
Query_SPIDE   3714   KILVKSGEDLRQDARIQQMFSVMNEIYAKDV-NCAQR   HL--YLKTYKVVPLSTRLGLIEWVDNTTVLNDFLND---G   3784
Query_SCORP   3735   KFLVKSGEDLRQDSRIEQLFDVMNDIYQCNA-ICFQN   KL--KLETYQVIPLTTKIGLIEWVDNTNVLGSFLKD---A   3805
Query_TICK_   3563   RFLVKTGEDLRQDDRIERLFGAMNALFTADP-SCRAK   HL--QLVTYGVVPLTTRVGLIEWLDGTVVLKDFQRR---G   3633
Query_MONOS   3781   KFLVKAGEDLRTDDRIEQVFSVMNRILAHDR-ACQQR   RL--NLRTYAVIPMTTRLGLLEWMDNTTTLRDFVMQ----   3850
Query_ROSET   3730   RFLVKAGEDLRTDARMEQLFAVMNQALAADP-ACGER   GL--SLRTYNVVPMTTRIGMLEWVPNSCVFKEMINL---G   3800
Query_BELCH   3824   PFLVKGGEDLRQDQRIQQLFDIMNDILTLDP-ACSQR   QL--RIRTYGVIPMTSRVGLIEWMQNTTPLAGFLEN---A   3894
Query_CIONA   3769   GFLIKGGEDLRQDQRIQQIFVIFNQLLESDM-GCRKR   NL--SMRTYQVIPISSKVGMIEWVDNVVPYIEFISQ----   3838
Query_CAECI   3768   PFLVKGGEDLRQDQRIQQLFDVMNIILSQDA-ACSQR   NM--QIKTYQVIPMTTRLGLIEWLEATNTLKDFLYN---S   3838
Query_HUMAN   3749   PFLVKGGEDLRQDQRVEQLFQVMNGILAQDS-ACSQR   AL--QLRTYSVVPMTSRLGLIEWLENTVTLKDLLLN---T   3819
Query_MOUSE   3749   PFLVKGGEDLRQDQRIEQIFEVMNAILSQDA-ACSQR   NM--QLRTYRVVPMTSRLGLIEWIENTMTLKDLLLS---N   3819
Query_XENOP   3766   PFLVKGGEDLRQDQRIEQLFEIMNIILSQDA-ACSQR   HM--QLKTYQVIPMTTRIGLIEWLENTCTLKEFILN---T   3836
Query_PLATY   3756   PFLVKGGEDLRQDQRIEQLLDIMNIILSQDA-ACSQR   NM--QLKTYQVIPMTTRLGVIEWLENTCTLKELILK---S   3826
Query_TASDE   3761   PFLVKGGEDLRQDQRIEQLFEVMNIILSQDA-ACSQR   NM--QLKTYHVIPMTTRLGLIEWIENTCTLKELLLN---N   3831
Query_CANAR   3757   PFLVKGGEDLRQDQRIEQLFDVMNIILSQDA-TCSQR   NM--QLKTYQVIPMTTRLGLIKWLENTCTLKEFLKN---S   3827
Query_OSTRI   3727   PFLVKGGEDLRQDQRIEQLFDVMNIILSQDA-ACSQR   NM--QLKTYQVIPMTTRLGLIKWLENTCTLKEFLKN---S   3797
Query_TURTL   3765   PFLVKGGEDLRQDQRIEELFDVMNIILSQDA-ACSQR   NM--QIKTYKVIPMTTRLGLIEWLENTCTLKNFLME---S   3835
Query_GECKO   3767   PFLVKGGEDLRQDQRIEQLFDVMNTVLLQDA-ACSQR   SM--QLKTYQVIPMTTRLGLIEWLGNTCTLKEFLLQ---N   3837
Query_DANIO   3743   PFLVKGGEDLRQDQRIEQLFGVMNMILSQDT-ACSQR   SL--ALRTYQVIPITSRIGLIEWMENTCTLKDFLSS---R   3813
Query_SNAKE   3752   PFLVKGGEDLRQDQRIEQLFEVMNNILSRDA-ACSQR   NM--QIKTYQVIPMTTRLGLIEWLENTYTLKEFLLK---N   3822
Query_STERL   3774   PFLVKGGEDLRQDQRIEQLFDVMNIILSRDA-TCSQR   NL--QLKTYQVIPMTARIGLIEWLENTCTLKEFLFD---R   3844
Query_MILII   3764   PFLVKGGEDLRQDQRIEQLFEVMNIILSRDA-ACSQR   SL--QLKTYQVIPMTSRLGLIEWLDNTCTLKDFLSS---N   3834
Query_BAMBO   3758   PFLVKGGEDLRQDQRIEQLFDIMNIILSRDA-ACSQR   NL--QLKTYQVIPMTSRLGLIEWLDNTCVVKDFLLS---N   3828
Query_ALLIG   3724   PFLVKGGEDLRQDQRIEQLFDVMNIILSQDA-SCSQR   NM--QIKTYQVIPMTTRLGLIEWLPNTCTLKNFLHS---N   3794
Query_GAR__   3771   PFLVKGGEDLRQDQRIEQLFGVMNIILSQDA-ACSQR   NL--GLRTYQVIPMTTRIGLIEWLENTCTLKDFLDS---R   3841
Query_TETRA   3977   KFLVKGIEDLRLDQRIEQLFQVMNNIFEYDP-QCQSK   EL--NLSTFNIIPMTNILGLLEWVDKTSVLKEILNKef--   4048
Query_PARAM   3585   LFLVKGGEDLRLDQRIELLFDVMNKILPDSD------   -----LSTYGVFPMTKLFGMLQWVDNTTVIKEIIEK---E   3647
Query_SLEMN   3847   HLLIKGGEDLRLDQRVQQLFQIINSIFKDDA-QCQNR   DL--YIKTFKVVPITNRLGTLEWVDDTEPMKSLINR----   3916
Query_STENT   3266   LLLVKGGEDLRLDKRIQKVFMVMNRIFSNDP-SCAKL   EL--KLKTFTIIPVTKKVGLLQWVSNTETLNTFITN----   3335
Query_NEMAT   3741   MFLVKGGEDIRLDQRIEQLFCMMNEVMSEDP-ACSQR   NL--RLRTYQVIPMTQRVGLIEWLKNTKPLKELLHI---A   3811
Query_PISTI   3793   MFLVKGGEDLRLDQRIEQLFGLMNDIMANDS-ACSQR   GL--RLRTYQVIPMTPRVGLIEWMNNTKPLKGVLNS---A   3863
Query_CORAL   3767   TFLVKGGEDLRLDQRIQQLFGLMNDVMRDDA-ACHQR   NL--SLKTYKVIPMTSRIGVIEWLENTTVLKEFIRN---G   3837
Query_HYDRA   3596   YFLVKEGEDLRQDQRVQQVFVLMNKILQNDS-VCCNR   LL--KLKTYQVVPITSRLGMIEWVSNTIPYFDFLYN----   3665
Query_LATIM   3765   PFLVKGGEDLRQDQRIEQLFEVMNIILSQDA-ACSQR   NL--QLKTYQVIPMTTRIGLIEWLENTCTLKEFLSN---A   3835
Query_SHRIM   3704   KFLVKWGEDLRTDQRMEQMFTLMNSIYARSP-LCTKA   SSvpSLDTYQVIPLSLRVGILQWVDSTQPLKDFIND---S   3776
Query_MCRAB   3695   KYLVKWGEDLRTDQRMQQVFTLMNSVYSSSP-LCAHT   SFhpSLDTYQVIPLSSEVGMLKWVESTQPLLDFIKD---S   3767
Query_AMPHI   3733   KFLVKTGEDLRTDQRIELMFTLMNNIFAHSP-VTKRL   SVeaRVVTYSVVPLTVELGMLQWVDSTVTLKDFMEA---A   3805
Query_BARNA   3490   RYLIKAGDDLRTDQRVQQLFRLLNGALAAEP-AAADR   RL--ALTTYTVLPVHPTLGIIQWLDHTSTLHQFIGS---A   3560
Query_COPEP   3523   GFLVKAGEDLRQDQRIENLFEICNICLAHES-SCRSK   DL--NIKTYKIIPLTPTLGLIEFVPNTKPLKDLVLS----   3592
Query_DAPHN   3316   PILVKHGEDLRQDERIVQLLNSIDVALVRDV-ESRKR   SL--RIRTFQVIPLTTAFGILQWLPETLTLKAFLET----   3385
Query_STARF   3814   SFLVKGGEDLRLDQRVEQLFRIMNSILEQDS-TCSQQ   GL--QLVTYEVIPMTPRLGMIEWAQNTTTLKEFLHL---A   3884
Query_STRON   3795   PFLVKVGEDLRMDQRIEQLFCIMNSILAQDA-ACSQR   GL--TLTTYQVVPMTPRLGILEWVKRTTTLKDFITR---S   3865
Query_SEACU   3768   EFLVKGGEDLRLDQRIEQLFCIMNDIMAEDT-ACSSR   EL--NLTTYQVTPMTSRLGLIEFVPNTTTWGAFLKV---K   3838
Query_FUNGI   3619   PFLVKGGEDLRLDQRIQQLFSVTNRILERDP-ACMKR   KL--SMKTYSVVPMTSKVGIIEWIENTKPLKEIMEE---Q   3689
Query_DPURP   3784   PFLIKGGEDLRLDQRIQQLFMIMNEILKRDN-SCNKR   GL--KVTTYQVIPMTSKVGIIEWINDTRPLREILED---Q   3854
Query_HALBU   3775   PFLIKGGEDLRLDQRIQQLFFVMNEILKRDP-ATSKR   GL--HIKTYQVVPMTGKVGIIEWLNDTKPLKEIIED---Q   3845
Query_GLOBO   3511   PWLVKGGEDLRLDQRIEQMFSIMNDMLVKNP-FCSRN   RV--AVATYKVIPMSTSLGVIEWVDNTKPLKACMAD---S   3581
Query_PALUS   3578   PWLIKGGEDLRLDQRVQQMFSIMNDIMLHDR-NCLEG   GL--FLRTYKVIPMSTSLGILEWVENTKPLREVMMS---I   3648
Query_CALIF   3685   LFLVKGGEDLRLDQRIQQLFNIMNDLLIKEA-YCKKN   NI--AIKTYKVIPMTGSLGIIEWVPDTISIGSCIQE---M   3755
Query_GLOMU   3702   LFLVKGGEDLRLDQRVQQLFSLMNGILKKDS-FCSQR   NI--TQRTYKVVPITGHLGIIEWINNTTPMRYFIER---E   3772
Query_LOBOS   3674   LFLVKGGEDLRLDQRIQQLFSLMNDIMRKDP-QCSQQ   NF--CIGTYKVIPMSSSLGVLEWVDNTKPLRHCIEG---E   3744
Query_RACEM   3657   LFLVKGGEDLRLDQRIEQLFGVMNSTLQRNA-YCARQ   SI--SVATYKVVPMASNIGIIEWVENTKPLKSCLDA---E   3727
Query_PVIOL   3776   PFLIKGGEDLRLDQRIQQLFRVMNEILGRDT-AASKR   GL--RVTTYQVVPMTSKVGIIEWINDTKPLREILEE---Q   3846
Query_BOMBU   3478   HFLTKFGEDLRQDQRLQQLFTIMNKTLQIDA-ACRQR   QL--TIDTYQVIPLSKTVGLIQWVNNTRSLQELIN----F   3547
Query_AEDES   3624   DFLVKYGEDLRQDQRIQQLLGTISNQLALDR-HCKEH   QL--SVRTYEVIPIRVDFGIIGWLSNTSSIKNIAVRsmvR   3697
Query_ANT__   3555   KFLVKFGENLIVDRGLQQLYSTMNRTLRNDP-DCRQR   RL--AIDTYEVIPLSSSFGLIQWIDDTKSLDDLIR----F   3624
Query_WHITE   3617   SFLVKCGEDLRQDQRIQQLLTLMNSSLTSNS-GTAAR   KL--SIVTYAVVPLSSSLGLIQWVNDTVPLKEL----mfK   3686
Query_BEETL   3452   PFLVKFGEDVRQDQRIQQLFSLMNNIFQHDR-NLNGR   KF--HILTYQVIPLTTSLGIIQWI----------------   3509
Query_THRIP   3539   NFLVKYGEDLRQDERVQQLLGLMNGIFSNNH-LCRDR   HL--SVVTYKVMPLSTRLGMIEWVDNTSCLKDV----mmS   3608
Query_TERMI   3622   QYLVKFGEDLRQDQRIEQLFNLMNKILDQDP-TCRHR   NL--SITTYQVVPLSSHLGIIQWVDGTEALKGL----mlG   3691
Query_FLEA_   3529   VYLIKYGEDLRLDQRIQQIFSLMNEQLSLNT-ECNFR   RL--SIVTYNVTPLTLNLGMIQWIPKTKTLRD-----mmF   3597
Query_MIDGE   3646   SFIIKYGEDLRQDQRIQQALCTMSRQLTMDK-NCLKN   KL--KIQTYEVVPINSTCGMISVVEDATTMADFLNKttsN   3719
Query_LAODE   3597   GYLVKQGEDLRQDERIMQALRLMDSFLALDN--VDSP[2]KY--SLKTYNVIPFTSNLGLLSWVDESTCLKDFILSc---   3668
Query_LOTTI   3704   HYLVKGGEDLRQDARIEQLFVIMNKVFDKDP-VCRHR   KL--HLKTYQVIPMTPRVGLIEWMNNTTPLKDFVYG---A   3774
Query_APLYS   3809   PYLVKSGEDLRQDDRIESLFYIMNKVMTSDP-NCRQR   KL--HLKTYQVIPMTPKVGLIEWMSNTCPLKEFLQD---A   3879
Query_SCALL   3794   PYLVKGGEDLRQDQRIEQLFFLMNQVFERDP-ACRQR   KL--NLRTYQVIPMTPKVGLIEWVSDTMTMKDFLNS---S   3864
Query_CRASS   3742   HYLVKGGEDLRQDQRIEQLFFLMNQALESDP-ACKQR   NL--KIKTYQVIPMTSRVGLIEWMNDTIPLKEFLIC---N   3812
Query_OCTOP   3656   PFIVKCGEDLRQDQRIGQLSEVINEIFFKNS-ICQTK   RL--QLKTYQIIPMTTRVGLLECLENVLTLQDTLQK---V   3726
Query_POMAC   3745   HYLVKGGEDLRQDQRIEQLFVLMNKVLEKDP-VCKAR   RL--VLKTYQVIPMNTRVGLIEWMSNTIPLKEFLLN---A   3815
Query_WHIPW   3597   KYLIKNGEDLRQDQRIQQLFKLMNRIFETGP-TDTHY   KM--HLRTYDVVPLGTRVGMIEFLPNVVELKKFFLS---P   3667
Query_TRICH   3600   KYLVKAGEDLRQDERVQQLFDLMNGILQKQH-QCSRL   KA--RIRTYRIVPLSIKLGLIEFLPNVVPLQQFFME---E   3670
Query_AASTA   3400   MYLVKGGEDLRLDQRIEQLFGVMNGILHQHT-TCSRR   QL--STRTYNVVPMTTTIGLVEWLGNTTTLKTIVEE---E   3470
Query_AEUTE   3327   LFLVKGGEDLRLDQRIEQLFGAMNTILQQSN-QCSRR   HL--NTRTYNVIPMTTKIGLVEWLENTTTLKSIVEQ---E   3397
Query_PLASM   3589   TFLVKGGEDLRLDQRIEQLFDVMNQIFKADP-RCRDQ   KL--SLTTYNVIPMTQEIGLLEWIDGTSTLKGLVEA---Q   3659
Query_BREMI   3572   TFLVKGGEDLRLDQRIEQVFGVMNQIFKADP-RCRDQ   RL--CLTTYEVFPMTQDIGLLEWVDGTLTLKGVIEA---Q   3642
Query_OLIGA   3625   TFLVKGGEDLRLDQRIEQLFDVMNQILASHP-RCRDR   HL--NTKTYKVIPMTTEIGMIEWLHNTSTLKGVIEK---Q   3695
Query_LIGNO   4035   PYLVKAGEDLRQDQRIQQALGLFNSLLQRDI-QCRLR   RL--SLVTYQVVPLSPDLGILEWLERTQPYKDFLVH---Q   4105
Query_CESTO   3820   GWLVKGGEDLRQDSNVQRLFGFANHAISSGLvGAEPN[1]SV-tPLRTYVVVPVSSQRGIIQWLDGTATLSSFCE-----   3891
Query_OVI2_   4295   SWLVKGGEDLRQDTRIQHMFAITNQLLGSNRcGAASF[6]NVpvLVRTYSIVPISSHIGLVEWLESTTTLLAFYKA---A   4374
Query_PRIAP   3772   MFLVKGGEDLRQDQRLEQLFDIMNRSLDADA-ACRQR   AL--RLRTYAVVPMSTRLGLLEWVPRTKPLKQLLQE---A   3842
Query_ROTIF   3687   KFLVKCGEDQRQDQRVEILLDIINNLLKSDS-KCYQR   NL--SIKTYQVVPMTSKLALIEWLPETKPLKVILEN---D   3757
Query_SPONG   3788   MFLVKGGEDLRLDQRIEQLFETMNKVLSDDP-ICSQR   RL--NLRTYSVIPVTPRVGLIEWVRNTMPLKDFI-----Q   3856
Query_TARDI   3725   PFLVKGGEDLRLDQRIEQLFVLMNRVFDQNP-ECAQR   RL--QIRTYKVVPMTTQVGIIEWVQGAMPLKEFLDK---A   3795
Query_TRICH   3754   LFLVKCGEDLRLDQRIEQLFCVMNEILAQDS-ICSQQ   DL--RLRTYQVIPITQRLGLIEWMKNTITLKEFVTD---A   3824
Query_OSTRE   3566   AFIAKGGEDLRQDERIQQLFRAMNGMLASFT-ESRGR   RM--EVRTFHVAPLTPRTGLIEFLSNTSPLLRLLSK----   3635
Query_NITEN   3920   RFVVKGGEDLRLDQRIEQLFEVMNNALRVDP-MCARR   RL--GVRTYGAIPVSSRCGLLEFVSPTRTIHDVIAS---G   3990
Query_CLUBM   3531   KFIVKGGEDLRLDQRIEQLFAKMNSVLREDG-ACAKR   EL--SMRTYAVIPVSQQCGLLQFVDNTCVLDDVIKD---G   3601
Query_MOSS_   3692   KFVVKGGEDLRLDQRIEQLFEAMNNVLNRDA-ACSQR   KL--AIRTYSVIPVSKKCGLLQFVENTCVLQEVIRD---G   3762
Query_LIVER   3760   KFVVKGGEDLRLDQRIEQLFEVMNSILQRDS-QAARR   GL--TIRTFAVVPVSKSCGLLQFVENTCVLEDVIKD---G   3830
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3797  CSDVSWE-QIQN---K-NL    PNNPANLLCQ--WIGdSSNTRQ   MIEn-YRK   CYQK   YSRNETLNAFKKRE  3855
Query_LINGU  3811  MTEQEKA-QQKA---A---    --ILQRNFLS--RYGnPSDP-N   ICKk-YWG   MYER   CSRTEAVKSCQDKE  3864
Query_LIMUL  3853  FTKKDQErINQA-------[ 1]KKFHSYMKSL--SKS---PQE-   ---v-SRN   AYRK   LGRKETLKAFQEVV  3903
Query_SPIDE  3785  MSETEKKqFYKR-------[ 1]GNLAQSEILQ--WDR---KAE-[1]LASm-YHS   FYMK   KTREECIQKYEKIV  3839
Query_SCORP  3806  MTDKEKEsFRES-------[ 1]PKMTFYKFLL--SKT---NSRD[2]IPLl-YEN   IY-K   IADTEVNKKFTECI  3861
Query_TICK_  3634  LDESSQG-----------N    --------LA--RCS--RSFGL[1]-STedYAK   AYKE   PS--QAVSRFGRCV  3675
Query_MONOS  3851  -SDEDEKdIQRAketYLLF    ----ANKFAD--KRAhSSVADN[2]--------   ---R[4]GTRSDVKLAFEKAQ  3905
Query_ROSET  3801  MTDAERE---DAkkrKATQ[ 1]PQNHFSQFAN------------   ----rFSK[6]MYSS[4]AKRSDISRSFLERQ  3859
Query_BELCH  3895  MTTREQE-AFLH---QR-L    PLIDFDAYVH--RLGgTVKNDP   WGKv-YGS   LYQK   ASRTDVTLAFRQVQ  3953
Query_CIONA  3839  --QAKRM-NCNScitEPAR    ------TWIP--KNMvQKPPCE[1]-----YGV   LMQR   GSYAQVVEKFKRAE  3890
Query_CAECI  3839  MSEAEQE-NCSS---T-TK    ---HYKDWIQ--KM--GRC-DN   ISR--YEY   MYRR   ASYTETVVAFKARE  3890
Query_HUMAN  3820  MSQEEKA-AYLS---DPRA    PPCEYKDWLT--KMS-GKH-D-   VGA--YML   MYKG   ANRTETVTSFRKRE  3875
Query_MOUSE  3820  MSQEEKV-ANNS---DPKA    PIRDYKDWLM--KVS-GKS-D-   AGA--YVL   MYSR   ANRTETVVAFRRRE  3875
Query_XENOP  3837  MTEDEAK-IYNS---K-TT[ 2]PLYHYNAWLD--KK--EKV-GD   ARQ--HVT   SYTR   CDRTNTVASFRERE  3893
Query_PLATY  3827  MSEGEKA-AYES---NPKG    PVHEYSEWLY--KMS-KKQ-D-   FGA--YML   MYKH   ASRTEAVTAFRNRE  3882
Query_TASDE  3832  MSEEEKA-AYES---GPKA    PQLEYSTWLT--KIT-GKQ-D-   ISA--YML   MYRR   ANRTETITSFRNRE  3887
Query_CANAR  3828  MSEEEDI-NYYS---P-RG    PRATYSEWLS--RMG-GKA-QG   ISR--YHA   MYRN   ASRTEAVTSFKSRE  3883
Query_OSTRIC 3798  MSEEEDF-NYSS---K-KG    PRATYNEWLS--KMG-GKA-QG   VPR--YHV   MYRN   ANRTETVMSFKSRE  3853
Query_TURTLE 3836  MSEEEKD-DYHS---R-KG    PRIMYSDWLS--KMA-GKE-NG   IAR--YKT   MYSR   ANRTETVLSFRNRE  3891
Query_GECKO  3838  MSEQEKN-DYNS---S-KG    PRADYNEWLS--KMG-DRN-AG   LVN--YVS   MYKK   ASRIETVTSFQQRE  3893
Query_DANIO  3814  RTEQEQK-TI-------TR    PNEFYDEWIS--KVA-GKV-EG   IRR--YAE   LYKK   AKRVDTVNNFRRIE  3866
Query_SNAKE  3823  MSEQEKN-CYNS---S-QG    PFADYKDWLC--KMG-ERN-R-   IEC--YMT   MFKR   ASRTETVMSFQHRE  3877
Query_STERL  3845  MTEEERK-HC-------KR    PAESYMEWLS--KVA-GKE-KG   IAQ--YIT   MYKT   AKRTETVRAFRNVE  3897
Query_MILII  3835  ITEETQSrRQQEvlpK---    ----YREWLQ--KMA-GRK-DM   FLR--YPT   MYLK   ASRTETVTAFRTHE  3888
Query_BAMBO  3829  MTEDEHKrPSPKel-----    ----FSEWLQ--RIS-GSKSDV   TTL--YGQ   MYMK   ASRTETVKAFRQCE  3881
Query_ALLIG  3795  MSEEEKA-DYCS---E-KG    PLFAYNEWLS--KMAgGKI-ED   FAR--YMP   MYKR   ASRTETVMSFRNRE  3851
Query_GAR__  3842  RTEPEEK-SS-------PS    PQDMYSAWLE--KIS-RRS-KG   QEQ--YHE   VYKK   ASRTETVANFQSIE  3894
Query_TETRA  4049  ---QEKMgSKEDiaqDNEG    -LRMRREWLK-------QLDRK[2]LAEq-HIK   LQLE[2]---DKIEQAFNEQA  4104
Query_PARAM  3648  HQEEEGLeLFKNkalSKRD    ------EYAA--KAAkRQGQYN[2]--NknFEV[3]--LQ[2]-NATDVVSNFNEQK  3708
Query_SLEMN  3917  --EHMRNeAGKD--lNQSR[ 2]NERRK--WLK-sSVPgNAKKEG[2]------DQ[3]LLGC[1]-EKI-VIEGFNKHK  3976
Query_STENT  3336  --ELSKIyNISS--lIEA-    -NKQRYSWL--sSLPa-ITTNN[2]------AL[3]ALKK[1]-SNT-IISNFREHE  3391
Query_NEMAT  3812  ATESEKQ-AMVV----RRK    T----KNML--------DLTLR   ILF--DII   ISRE   ANRTDTESEFKKKQ  3859
Query_PISTI  3864  MTKAEND-HYIG----PNG    AGACYRNWID--KFK-GDPRNP   SRF--YGE   MFKK   ANRTETERSFTQKQ  3920
Query_CORAL  3838  LSETENE-TF-------DT    IGYKHNKWIV--EKF-PVNGNK   HIInnYYP   MFKK   ASREVIERKFRELQ  3893
Query_HYDRA  3666  --DDEKK-VLHGpsgPIK-    ------------KYKqFYGKFE[1]-----VSN   VYKS   ANRSETVKVFEDCV  3712
Query_LATIM  3836  MTEEERT-NYETlspK-KG    PKAQYNEWLT--KMA-GGS-VG   PQR--YGS   MYKR   ASRTETVMAFRSVE  3894
Query_SHRIM  3777  YRENETK-CYGE------A    -RALYGKNEN--YTVaKKAKKR[1]VIKa-Y--   ----   ----------EEVV  3817
Query_MCRAB  3768  FKEGEAK-CLDE------A    -ETFYRTEEP--WNFrKKTTDK[1]RVKk-Y--   ----   ----------EDAV  3808
Query_AMPHI  3806  LTDTEKR-PYSK------N    -LVKYGDGLS--TQDlLKQSKV[1]CVKa-Y--   ----   ----------SEVV  3846
Query_BARNA  3561  CPDLEQQ---------CAE    ARELFRKQMR--RHAtQSGAND[2]LQSe-YAR[1]APRD[3]ADLAAREALMRDER  3621
Query_COPEP  3593  -------------drKAET    ------EILE------------   -ARnaYKV[1]---K[2]-TSDDIRRKYEKAV  3626
Query_DAPHN  3386  ------DiNFRSslkSLKP[ 1]--------LE------RKLSAN[2]---lhFMT   ----   ATQLQLLNDYSNFA  3428
Query_STARF  3885  MTEQEGN-AYNS---RH-G    PRYQHSEWVD--KYA-KGNQSN   HGVr-YGA   VYRH   ANYTDTVKCFRSWE  3942
Query_STRON  3866  MTEAEHK-VYTS---RTTG    PVALFTTWLQ--KFC-Q-TDGN   ITKl-FTE   AYKK   ANETETIKAFRERE  3923
Query_SEACU  3839  SLDKSKL-E----------    ---AYGNWVQ--KFR-TKKDEK   IHLt-FAT   MYKQ   AGLQETTHNFRQCE  3887
Query_FUNGI  3690  LAKDTGKpKADTsilKTS-    AAAQQDAWVK--SFGsKVVGKK[3]PCDy-YYA   MFQH   ANRAETDKQFTKHA  3755
Query_DPURP  3855  IAIQTKTpRSNVsisKLE-    ATNIHNTWIQ--SFA-KYSK-I[2]PGPl-YQQ   MFIH   ANRDNAVKKFEKQI  3917
Query_HALBU  3846  MAVHQNVeRSNVslsRME-    AYKIHNEWIN--SFG-KYIKMA[3]AGPl-YQQ   MFMH   SDRADAIKKQEKQH  3910
Query_GLOBO  3582  PQFD----------kKMA-[ 1]SQAEYSAFVV--KFG-----KS[3]FGGa-YEP   FMMN   ATPRHVIENMQSIW  3633
Query_PALUS  3649  PGFK----------gEYS-[ 1]ADQKHAAFVE--SYK-----GH[3]LGDm-YGN   MFRK   ANRNRTMDHMSSL-  3699
Query_CALIF  3756  LDDK----------nLIT-[ 1]SRNIYCDWVN--SFY-----KR[3]--Nv-FHT   MFKK   ASREDTINNLIKVE  3805
Query_GLOMU  3773  PRTS----------kVFD-[ 1]AQSMHSRWVQ--QYK-----S-   -----YHN   MFME   VNRDAVVQHFKDLQ  3816
Query_LOBOS  3745  IPHK----------eTWR-[ 1]AQDQYNKFVA--SFK-----GD[3]-----YHN   LFIN   GTREKVVKHMESLY  3792
Query_RACEM  3728  VPSS----------aLIA-    ARKAYLAFVN--KHG-----QN[3]-----YVT   TYTK   ESRDSTVRAFEQSV  3774
Query_PVIOL  3847  LAIQTKMpRSQVsisRLE-    STKIHNDWIQ--EFS-KYLK-G[3]AGPl-YQQ   MFIH   ANRDNAVKKLDRQH  3910
Query_BOMBU  3548  TLSKQEIeRHNSiskLYQE[10]LHEKYKEAIV------------   --------   ----[1]YNASEVTNKMNEFI  3601
Query_AEDES  3698  FNTDG--dVTGAinaEYIQ[10]SPSTPLSKLY------------   --------   GKVA[2]CTPERITLKFNELR  3754
Query_ANT__  3625  TLSNKEKeCCKMirdEYAG[11]ITDRYKEAVS------------   --------   ----[1]YSQQMVMENMKKLI  3679
Query_WHITE  3687  FFLESDTeKYKKqvaEHYS[12]AGRIYGSAAQ------------   --------   ----[1]YSRQVVINTYNEHV  3742
Query_BEETL  3510  -------------------    -----GNTKSlhEFSnNENFDK[4]FTKwiYTS[7]AYGR[4]YSRDKTIVKFKTLV  3567
Query_THRIP  3609  SRQLSDQdRVSIr-tKHES[11]NINIYGIAHD------------   --------   ----[1]ISREASVTQFNKLV  3662
Query_TERMI  3692  VLDKNEKeSYKKsceEHCK[11]FYEKYGEAHM------------   --------   ----[1]YSPEETSAEFRKKV  3746
Query_FLEA_  3598  LKEPGIKnRVEEvrsSYNK[12]FVALYGNAVV------------   --------   HYHR[2]VISK-----YQELV  3653
Query_MIDGE  3720  FLQKSLYdFIPEirnEYKN[10]FSNAYTNALL------------   --------   RRGK[2]MAEE-----FASFE  3773
Query_LAODE  3669  LDSQEMKeYVNTe----KN    ----YQEWLP------SGASKN[2]-----YHN   LYCE   VSRPKTVENFRKNV  3718
Query_LOTTI  3775  LTEQERK-YILG---D-QG    PGKQHENWTY--KYFdAKKR-C   IKSm-YQK   VYQN   YSRAETIKDFRMKE  3832
Query_APLYS  3880  YTEEERK-FLEG---N-QS    PVIQHNQWMA--KLVdKKDSGN   WQMm-YDR   VYMK   YSHTETVKEFRLKQ  3938
Query_SCALL  3865  LTDQERN-FVHH---KTNG    PNGQHTAWLQ--K-LgKD-L-P   LAKm-YGL   MYQK   YNKTEAIRELRLKE  3921
Query_CRASS  3813  LTEQEAK-FV---------    SGNLHNSWIM--K-LsKDPR-D   VQQg-YMQ   MYLK   YNMTETRRELQKKE  3864
Query_OCTOP  3727  LEEKEN-------------    ---YIMETIE--RHR---EFKK[2]VASg-IMN   IFKK   CKKTDAELGYQSIC  3773
Query_POMAC  3816  MTEKERT-AYFG---E-NG    PSQLHYKWLP--K--lPKKT-D   WGLi-YGQ   VYQN   YSKTETMKEFRIKE  3871
Query_WHIPW  3668  ADRKNYQaMALKal--SEG[ 1]-----YAILN--TGFwGGLTPS   ----vYWK   SYAT   ATPTAVVRCFRHLL  3722
Query_TRICH  3671  SLRKQYQtIAMEmf--TEG[ 1]-----EKILL--KAAgEGSSPL[2]---ltYWR   SFQN   ITPETAARCFNEVL  3728
Query_AASTA  3471  W-------DDGTsskKKTN[ 5]PGVQYEAFWA-------KQRGK[1]-----YGQ   ---K[3]SATSAVVSSFVQAE  3524
Query_AEUTE  3398  -----------MsttKKSN[ 5]PGQLYENFWS-------KQRGK[1]-----YGQ   ---K[3]ASSSAVTLVYSQAQ  3447
Query_PLASM  3660  LQIDERCvGLKSnnsQKLD[ 5]AAKAYESFLL-------EQRGA[1]-----YSA   ---K[3]SRSEDVVDHFAKVH  3720
Query_BREMI  3643  MQVDERCiDLKSdksKKLD[ 5]AAKAYESFHR-------KQQGA[1]-----FSA   ---K[3]PRSKDVVDQFHRVQ  3703
Query_OLIGA  3696  LQVDLRCtELLSnkrAKLQ[ 5]PAKNYETYLM-------KQRGG[1]-----FAA   ---K[3]PSHKDVVANFEGAQ  3756
Query_LIGNO  4106  MSAAEQR------------    --------------ShSKQSLD[4]VSNd-HCK   LYTL   SDQQVVQRNFQPLL  4149
Query_CESTO  3892  -------------------    ------NAMTtkEVErFKSLDA[2]---srHQG[2]LWRS[3]STAALVVSRFTELE  3937
Query_OVI2_  4375  MCDTELQ----SfkyQVSS    ----FGELLA-----------G[2]-----YSG[7]SYPK[2]-RPEQIITRFRELE  4427
Query_PRIAP  3843  MTQQQYD-AYYS---YSMN    PVEHFNKWAA--TAL-TASESK[4]HVKl-YAR   TFVQ   CSKSDTVRSFRERE  3905
Query_ROTIF  3758  LVNEAQN-NYAEfi---YG    --------IN--KDGsKKTQSE[1]-----YGM   VYAK   YSRELICREFDKIQ  3806
Query_SPONG  3857  LSDPEKK-SLPA---ASKM[ 1]-----QKWLE--QKFgGSNYVN[1]-----YYE   MYKK   LKADEVKKEYRRIV  3909
Query_TARDI  3796  QSPQDKVsLETAsfaRYDL[ 2]KQLKISDKKA--IRTyAGGEKE[3]SEKdlNNR   AFLE   VKPADMRKIFQQVT  3865
Query_TRICH  3825  MEDHEKQ-HYLS---SQNS    PHQLYSKWLG--KF--GSAGGP   VYM--YQS   MYRK   AKRSEVENHFKFRE  3881
Query_OSTRE  3636  ---------------PVTG[ 4]PFKEHQRWIRehARGiGKRKRD[2]QGEmtTTH[1]DYLA[4]ASRKESETILSRLR  3698
Query_NITEN  3991  LAASLNLqTPDGaktLLS-[ 1]VQKGWQQWLT--SHT-----GA[3]LQEa-YQK   SYTS   LSRQSFEAQLGVAG  4052
Query_CLUBM  3602  VAQQLAAaKRKEk------[ 7]LRNQYHQWVE--KKG-----G-[3]VHDc-YKN   LYTK   VKAGEMEARMNALV  3663
Query_MOSS_  3763  LSSKLASrTHEGykpLQAT[ 7]IQHEYIEWVR--KKG-----G-[3]QTEc-YCN   LYRK   VGFEEVSLKLSALQ  3830
Query_LIVER  3831  LLCKMPNlRREGkdaPLA-[ 4]LRAQYHEWIM--KRG-----GP[3]PPEc-YRN   MYQK   ATSDEVTQKMSSIE  3895
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  3856  AMV-PWDMLRKSYQKLASSPE-SYFVLRGNCATSHALNCISAYIMGIGDRHLSNHMVDLINGKMIAIDFGHAFGSA--TQ  3931
Query_LINGU  3865  NSL-PWDLLRRAYLQMSTSPE-TFYTLRRHFAISHALLCICHYILGIGDRHLSNFMVDLETGGMIGIDFGHAFGSA--TQ  3940
Query_LIMUL  3904  DLV-PSDIMKRGFVKLSSCPE-AFLTLRSQFVISHAVLSISQWLLGIGDRHLGNFLIKTDTGKELGIDFGHAFGSA--TQ  3979
Query_SPIDE  3840  NFV-PKNLLKKSFLKLSSSPE-AFIALRTKFAKSHAVICISQWIIGIGDRHLGNFLIDKTSGCEIGIDFGHAFGSA--TQ  3915
Query_SCORP  3862  NMI-PWDLFRRAFLKFSSSLE-AFVILRYNFITSHAVLCISQWLLGIGDRHTGNFLVSTTTGIEVGIDFGHAFGTA--TQ  3937
Query_TICK_  3676  RQV-SRESLSKALLGLSSCPE-AFFALRSRFVLSHATLCVAHWVLGVGDRHLGNFLVSTNTGLEIGIDFGYAFGLT--AQ  3751
Query_MONOS  3906  SMV-RATLLRDALLALASTPE-GFFTLRQHYGRTLSTINICQYILGIGDRHLSNVMVDLRTGGLIGIDFGHHFGSA--TY  3981
Query_ROSET  3860  ALV-RPTLLRDALLHLSTTPE-AYLTTRQLFGRTLATMNICQYILGIGDRHLSNSMIDLRTGGVVGIDFGHNFGSA--TQ  3935
Query_BELCH  3954  SKV-TWDIMRRGYLQMSSCPE-AFLTLRSHFTTTHATLCIAQYVLGIGDRHLGNFMIDLESGGMVGIDFGHAFGSA--TQ  4029
Query_CIONA  3891  NML-PWDLFRQAYIRLCANPE-SFVALRKRFTSSYAVMCVTHWLLGIGDRHTSNFMLSLTTGEVVGIDFGHAFGSA--TE  3966
Query_CAECIL 3891  NLV-PGDLLRRAFTKMSTTPE-AFLTLRSHFASSHALICISHWILGIGDRHLSNFMINMETGGMVGIDFGHAFGSA--TQ  3966
Query_HUMAN  3876  SKV-PADLLKRAFVRMSTSPE-AFLALRSHFASSHALICISHWILGIGDRHLNNFMVAMETGGVIGIDFGHAFGSA--TQ  3951
Query_MOUSE  3876  SQV-PPDLLKRAFVKMSTSPE-AFLALRSHFASSHALLCISHWLLGIGDRHLNNFMVAMETGSVIGIDFGHAFGSA--TQ  3951
Query_XENOP  3894  ALV-PKDLLRRAFVKMSTTPE-AFLSLRSHFARSHALLCVSHWIVGIGDRHLSNFMINMETGGMIGIDFGHAFGTA--TQ  3969
Query_PLATY  3883  NRV-PGDLLRRAFVKMSTRPE-AFLALRSHFARSHALLCVSHWILGIGDRHLSNFMINLETGGMVGIDFGHAFGSA--TQ  3958
Query_TASDE  3888  SKV-PADLLRRAFVKMSTTPE-AFLALRSHFASSHALMCISHWILGIGDRHLSNFMINMETGGMIGIDFGHAFGSA--TQ  3963
Query_CANAR  3884  SSV-PEDLLRRAFVKMSTSPE-AFLALRSHFASSHALMCISHWILGIGDRHLSNFMINKETGGMVGIDFGYAFGAA--TQ  3959
Query_OSTRI  3854  SSV-PEDLLRRAFVKMSTAPE-AFLSLRSHFASSHALMCISHWILGIGDRHLSNFMINKETGGMVGIDFGHAFGSA--TQ  3929
Query_TURTL  3892  TSV-PGDLLRRAFVKMSTAPE-AFLSLRSHFASSHALMCISHWILGIGDRHLSNFMINMETGGMVGIDFGHAFGSA--TQ  3967
Query_GECKO  3894  NLV-PEDLLRRAFVKMSTTPE-AFLSLRSHFASSHALMCISHWILGIGDRHLSNFMVNMETGGMVGIDFGHAFGSA--TQ  3969
Query_DANIO  3867  QMV-PDDLLKRAFVRMSTTPE-AFLSLRSHFSSSHAVLCISHWILGIGDRHLSNFMINTETGGMIGIDFGHAFGSA--TQ  3942
Query_SNAKE  3878  NRV-PEDLLRKAFVKMSTTPE-ALLSLRSHFASTHALLCISHWILGIGDRHLSNFMISTETGGLVGIDFGYAFGSA--TQ  3953
Query_STERL  3898  HHV-PEDLLRRAFMKMSTTPE-AFLSLRSHFTCSHALMCVSHWILGIGDRHLSNFMINMETGGMVGIDFGHAFGSA--TQ  3973
Query_MILII  3889  QAI-PGDLLRRAFIKMSTTPE-AFLALRSHFASSHALMCISHWILGIGDRHLSNFMVNLETGGMVGIDFGHAFGSA--TQ  3964
Query_BAMBO  3882  RAV-PPDLLRRAFTKMSTTPE-AFLALRSHFASSHALMCVSHWILGIGDRHLSNFMINLETGGMVGIDFGHAFGSA--TQ  3957
Query_ALLIG  3852  NRV-PEDLLRRAFVKMSTAPE-AFLSLRSHFTSSHALICISHWILGIGDRHLSNFMINTETGGIIGIDFGYAFGSA--TQ  3927
Query_GAR__  3895  KLV-PQDLLRRAFIQLSTTPE-SFLSLRSHFASTHALICVSHWILGIGDRHLSNFMVNMETGGMVGIDFGHAFGSA--TQ  3970
Query_TETRA  4105  DLI-PRDVFRNALERLASTCE-TYLYLKERFIRNYAVVCVSGYILGIGDRHLENFLVNYSNGDIVSIDFGYSFGAG---L  4179
Query_PARAM  3709  KYV-KDTLLRNGLSKLSANLE-AFVFIRDKFLINYAGICASGYILGIGDRHLENILLNYSNGNLIAIDFGFSFGQG---L  3783
Query_SLEMN  3977  EAI-PWYLLRNGLENLCLTSS-AFLAIKNQFTKSLATFSIASYLIGIGDRHLENFLIDTSDGEILGIDFGIAFGSG---I  4051
Query_STENT  3392  SFI-PWDLIRQGLLSLSTTPE-SFMHLRKQFIHSLAAISICGYIIGLGDRHLDNFLLDKNDGGLLSIDFGVSFGHG---V  3466
Query_NEMAT  3860  AQV-PWDLLRRAFQQLAASPE-AYLTLRTHFARTHACICICQYVLGIGDRHLSNFLVDMRTGGLIGIDFGHSFGSA--TQ  3935
Query_PISTI  3921  ALV-PWDLLRRAFLQLAVSPE-AFFILRSHFASSHACLCICQYILGIGDRHLSNFLVDMETGGMIGIDFGHAFGSA--TQ  3996
Query_CORAL  3894  NLM-PHDALRRSFMRLASSPE-AFLTLRAYFARTYGCMTICHYILGIGDRHLSNSMVDLKQGRVIGIDFGHAFGSA--TQ  3969
Query_HYDRA  3713  STV-PWDSLRRAFEKLALSSE-AYFVLRKNFIQSHSCLCIAHYILGIGDRHLNNMLIDNKTGGVVGIDFGYAFGIA--TQ  3788
Query_LATIM  3895  SLV-SRDLLRRAFTKMSTSPE-AFLTLRSHFASSHALMCISHWILGIGDRHLSNFMINMETGGMVGIDFGHAFGSA--TQ  3970
Query_SHRIM  3818  NKI-PWDVLRRGLVRMSSSTE-GFFSLRSAFAVSYAALCISQWLLGIGDRHCSNSLVSLKTGRVIGIDFGHHFESA--VQ  3893
Query_MCRAB  3809  NRI-PWDILRRGLVRLSSSSE-GFFSLRSLFANSYATMCVSHWLLGIGDRHCGNTLVSLKTGRVVGIDFGHHFETA--VQ  3884
Query_AMPHI  3847  NSL-PWDLLRRSLSNLASSYE-SFFVIRTKFCSSHAVLSIAQWLLGIGDRHLCNFLVDYKTGTMVGIDFGHHFETA--TQ  3922
Query_BARNA  3622  GRA-CSTLLQQAVFRLSSGPE-ALLRLRTRMAESHALLCAAGWLLGVGDRHLGNFLVSLEDGAMIGIDFGHVFGSA--TS  3697
Query_COPEP  3627  ELV-EKYHLKNSILKMSSSMN-GFYYIRKNFITSYAVLSVMQWILGIGDRHLSNYLVCQKTGRVVTIDFGYAFGIS--TS  3702
Query_DAPHN  3429  DRVyPYG-LRGALKKLSQSPE-SFFYLRNRLATSHGTQCLALWILGVGDRHLSNFLVSTKNGDIIGIDYGMAFGMA--TS  3504
Query_STARF  3943  ALV-PWDLMRRAFIQLSASPE-AFLTLRSHFARSHAVICICQYILGIGDRHLSNFLISLKTGGMIGIDFGHAFGSA--TQ  4018
Query_STRON  3924  AKL-PWDLLRRGFMQLSASPE-AFLTLRAHFARSHAVLCICQYILGIGDRHLSNFLVSLETGGMVGIDFGHSFGSA--TQ  3999
Query_SEACU  3888  EGI-PWDLSRKAFLEMSASPE-AFLALRTHFAKSFATLCICQYILGIGDRHLSNFLVSLNSGRMIGIDFGHAFGSA--TL  3963
Query_FUNGI  3756  EGL-PSDLLKRGLFAASASME-TYLGLRSQFAKTLSVFSVASYLIGIGDRHLENFLFSFTDGSLIGIDFGHAFGTA--TQ  3831
Query_DPURP  3918  AKV-PENLLQNGIWALSSSPE-SYLFIRNSFARSLATFSVCSYVIGIGDRHLENFLISQRDGTLIGIDFGHAFGTA--TQ  3993
Query_HALBU  3911  DKV-PANLLQNGIWAISSSPE-SYLFIRNSFARSLATFSICSYIIGIGDRHLENFLISQKDGILIGIDFGHAFGTA--TQ  3986
Query_GLOBO  3634  NKN-RSSYLREFFLKLTISAE-AFFQIRGEFANSFAALSICSYLLGIGDRHLDNFLVNLKNGKIVGIDFGHAFGSA--TE  3709
Query_PALUS  3700  -TS-FEPYLRTWMHKLAASPE-AFLSIRAQFAKSLATLNICSYLLGIGDRHAENFLIDMTTGQIVGIDFGHAFGSA--TE  3774
Query_CALIF  3806  NQI-PKRLISKYLLKIIRTPD-IYIKIRNEYANSLAAISICSYILGIGDRHLENILVNKNSGQLIGIDFGHAFGTA--TE  3881
Query_GLOMU  3817  REI-SNDMLKKALYKLSTSPE-AHLSIRSEFIKCHAVINICGYLIGIGDRHLENFLVDTTKGSLISIDFGHAFGSA--TE  3892
Query_LOBOS  3793  NQF-REDYLRQSISRLAASPE-AFLMLRSEFTKSLAAINVCSYILGIGDRHLENFLLDMSSGCLIPIDFGHAFGSA--TE  3868
Query_RACEM  3775  SHL-EPDYLRHYLLKLASSPE-AFFFIRNGFANSLAVTSISGYVLGIGDRHLENLMVDLKSGRLVPIDFGHAFGSA--TE  3850
Query_PVIOL  3911  ARI-PADLLQKGIWALSSSPE-SYLFIRNSFARSLAAFSICSYIIGIGDRHLENFLISQKDGRLIGIDFGHAFGTA--TQ  3986
Query_BOMBU  3602  GKT-EWNSLRKTLTVLCPSIE-SFITMRRNFITSYATMCIAHWILGIGDRHLENTLIVIDSGRCLGIDFGLAFDAG--V-  3676
Query_AEDES  3755  YKI-KEDTFKRALFEMAVSAE-SFYSLRGNFAKSLMAMNVSCWILGIGDRHTSNVLIDRSNGKLAGVDFGIAFGAG--T-  3829
Query_ANT__  3680  SKT-RQTALRDAFETISSSPE-CFVTLRRNFVTSYATMCTAHWLAGIGDRHLQNTLVQVATGRCMGIDFGSAFGSG--I-  3754
Query_WHITE  3743  SKL-EKINLRKSFMSNSLNLE-AFISMRNEFLCSYAVMSICHWILGIGDRHLSNTLVNLFTGKCLGIDFGHAFGSA--TQ  3818
Query_BEETL  3568  SKI-SWDVFRSTFLKLSSNTE-GFFALRNNFIKSYAIMCASHWIVGVGDRHLSNSLVCLNSGKVLGIDFGHAFGTA--TQ  3643
Query_THRIP  3663  SMT-KPDLLRKGLQLWSSSAE-GFYWLRHSFITSYSTLCIAHWLLGVGDRHLQNCLVSQTNGQCIGIDFGHVFGSA--TL  3738
Query_TERMI  3747  NMI-PWDIMRRALQRLSVSPE-SYFILRQSFACSYAVLCIAHWLLGIGDRHLSNILVSVKDGRCVGIDFGSAFGTA--TQ  3822
Query_FLEA_  3654  NRI-PEDILKDSFKEMSANAE-SFFTLRHNFATSYGTMCIAHWLLGIGDRHLSNCLININSGFAVGIDFNCAFGSG--VM  3729
Query_MIDGE  3774  ALL-PQDIIRRSLMESAFSME-TYYILRQNFITSMTAMNVAHWLLGIGDRHMNNILIDMKTGKLIGIDFGVAFGAA--T-  3848
Query_LAODE  3719  NSL-PQDILRRGILELSQSAE-SYHYLRTQFGSSLSAMSVCHWLLGIGDRHLENALLCLGSGQLCGIDFGHAFGSA--TE  3794
Query_LOTTI  3833  NKV-PWDLFRRAFHQMSTSPE-AFHVLRCELVVTHATISICQYFLGIGDRHLSNFMINLKTGGMVGIDFGHNFGSA--TQ  3908
Query_APLYS  3939  GKI-PWNLSRRALQQLSSGPE-AFHVLRTSLVTSHATLCLCQYLLGIGDRHLSNFMVSLKTGHMIGIDFGHAFGTA--TQ  4014
Query_SCALL  3922  NKI-PWDLFRRAFHDISTSPE-AFHVLRCTLAVSHALVSICQYILGIGDRHLSNFMINLKTGQMIGIDFGHAFGSA--TQ  3997
Query_CRASS  3865  NII-PWDLSRRAFHKMSTSPE-AFHVLRCKCASTHALISICQYVLGIGDRHLSNFMINLKNGEMVGIDFGHAFGSA--TQ  3940
Query_OCTOP  3774  ARI-PQDLLRKSFKKMSTSPE-AFHVLRTGFTISHAVLCICHYILGVGDRHLSNLMIHQGTGQMIGIDFGHAFGSA--TQ  3849
Query_POMAC  3872  SKV-PWDLSRRALQHLSTSPE-AFHFLRCEMLTSHAVLCICHYLLGIGDRHLSNFMVNGRTGHMVGIDFGHAFGSA--TQ  3947
Query_WHIPW  3723  DQI-PRHLLRHKLINCCASPD-MFCCLRERFVASLAALSVGTYLLGIGDRHLDNILIDTSMGELIGIDFGYAFGTS--LQ  3798
Query_TRICH  3729  KKI-PDHLLRDQLLNCCVSPD-VFCFLRQKFTVSLASMCIANYLLEIGDRHLGNIVLDTKTGEVIGIDFGYAFGAS--LQ  3804
Query_AASTA  3525  AAV-PCDLLRRRLVHMANTSD-TFFQLRESFATSLAAFNGCSYVLGIGDRHLDNFLLDERTGAVVGIDFGISFGAG--AS  3600
Query_AEUTE  3448  EMV-PMDLIRRHLVNMAQTPA-AFFQLRETFSTSLAAFNGCSYILGIGDRHLDNFLLDYITGTVVGIDFGISFGSG--AS  3523
Query_PLASM  3721  AMI-PPDLLRRQLLKLGSTSD-VFFLVRDNFLKSLAVFSACSYVLGIGDRHLDNFLLDLGSGRVIGIDFGVSFGAG--AS  3796
Query_BREMI  3704  AMI-PADLLRRQLLALGLNFE-AFLLVRDHFLKSLAVFSACSYILGIGDRHLDNFLIDLSSGRVIGIDFGVSFGAG--AS  3779
Query_OLIGA  3757  ALI-PADLLRLQLLGLGSDFE-EFVSIRDQFASSLGVFSACSYILGIGDRHLDNFLLDHQTGRVIGIDFGVSFGAG--AS  3832
Query_LIGNO  4150  DCV-PDCLLARGLRQLSSSWE-AYYAMRDTLVRSHAAMCGAHWLLSVGDRHMSNYMLCLDSGKMIGIDFGYAFGTA--IT  4225
Query_CESTO  3938  SFVfSLRLIRRALFKSIATSAeHFTCLRYRLISSYAALSALCFILGVGDRHLGNFLLEKLSGSLIGIDFGYVFGVSl---  4014
Query_OVI2_  4428  NRIkSWRLLRRGLQQMSRDVT-HFMELRSMCIRTQASLSAVHYVLGIGDRHPLNFLLCRKTGAFIGIDFGYAFGLT--TL  4504
Query_PRIAP  3906  SMV-PPDLLRQAFLRLSASPE-AFLALRTHFIKSHATLSVCHYVVGVGDRHLSNFMVSQEDGGMVGIDFGHAFGHShpRC  3983
Query_ROTIF  3807  NSI-PWDLFRQFLRTLSSDSE-AYFVLRNEFIKSYAVASTCHYILGIGDRHLSNWMIDLKTGRAIGIDFGIAFGHA--TM  3882
Query_SPONG  3910  SAV-PWDLLRRSLISLASTSE-SFLLLRCTFIQSLSVINICQYILGIGDRHLSNFMVDMETGQLVGIDFGHAFHSA--TQ  3985
Query_TARDI  3866  RPL-DPLLLRRQLQRLTVSPD-AFIIVRNLFAQAFGVMSIGQYVLGIGDRHLSNCMVSTETGLIFPIDFGHAFGSG--VQ  3941
Query_TRICH  3882  NSV-PWDLLRRGFVALSASSE-AYLMLRSHFARTHATVSICHYLLGIGDRHLSNFMIDRLTGGMIGIDFGHAFGSA--TQ  3957
Query_OSTRE  3699  SSShSNNALRNILLECAGNAE-AFLMMRQAFAASLSSSSICGYVAGVGDRHLDNILLDLSTGQLVHIDFGYAFGTA--TS  3775
Query_NITEN  4053  GKL-PSDALRAGI-AQQAPGAeSFLALRGQFARSLAALSICGYVAGVGDRHLQNFLLDVADGTLVPIDFGYSFGTA--TQ  4128
Query_CLUBM  3664  AQV-PWDALKTGL-YKRTSSAeSFLAMRSQFARSLSVMSICGYIAGVGDRHLANTLIDMQTGSLVPIDFGYSFGTG--VI  3739
Query_MOSS_  3831  GQL-PWDTLRMGF-SRLVTSAeSYLALRSQFARSLSVTSICGYVAGVGDRHLGNTLVDMRTGALVPIDFGYSFGTN--VL  3906
Query_LIVER  3896  AQM-PWDTFRAGV-SRWATSAeSFLALRSQFAKSLSVISICGYVAGVGDRHLANTLVDMRTGALVPIDFGYSFGTG--VI  3971
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT   3932   FLD-IPELIPFRMTRQLRNLLLPLR--EKGVIEGTMVHSLRALRQHKEQLISTMDVFIKEPSL-DWQNAA    KKEIQA  4003
Query_LINGU   3941   FLP-VPELMPIRLTRQILNVMMPLR--EKGLIESTMVHALRALRANYDLLLNTMDVFVKDLSV-DWLAFA    EKQLQT  4012
Query_LIMUL   3980   FLP-VPELIPFRLTPQYLNLMLPLK--EKGLYEATMIHTLNALRKNHDLLLNFMDVFIKEPTL-DWKNWA    QKQRNE  4051
Query_SPIDE   3916   FLP-VPELVPFRLTPQYLQLMAPLK--VKGIYAATMIHALKAISTKRDLLLNVMDIFIKEPTL-DWMNLA    NKQKQT  3987
Query_SCORP   3938   FLP-IPELVPFRLTNQYLNLMLPLK--EKGVIEATMISALSALRDNADILLNTMDIFIKEPTV-NWQENA    LKQKRI  4009
Query_TICK_   3752   FLP-IPELMPFRLSPQYLGLLEPLG--KEGPLASAMQDALRALRAGAESILDVLDVFVQEPTV-DWMRLA    KRQQAG  3823
Query_MONOS   3982   LLP-IPELMPFRLTRIMEQALQPHT--RNGLIKHVMCHVLTALKARRSDLLAVIQIFATEPTL-DWQQNA    KKVIKQ  4053
Query_ROSET   3936   ELP-IPEMIPFRFTKQLEAVFLPHK--KEGLVQEVMVHVLRALKANQRLLLSTMDVFIKEPSL-EWQQQA    -----R  4002
Query_BELCH   4030   HLP-VPELMPFRLTRQMLNLLLPLK--ESGLLQSTMVHVLRALRSSPNVLINTMDVFIKEPHL-DWKNF-    ---AAK  4097
Query_CIONA   3967   VLP-VPELVPFRLTNQIINLLSPHS-aIHGELRHVMVRCLSALRSRTDTILPLLDVFVKDPSV-DWAKFG    KRSRDQ  4039
Query_CAECI   3967   FLP-VPELMPFRLTRQFVQLMLPMK--ESGLIESVMVHALRAYRTDPSLLITTMDVFVKEPCL-DWNNFM    LKQLKK  4038
Query_HUMAN   3952   FLP-VPELMPFRLTRQFINLMLPMK--ETGLMYSIMVHALRAFRSDPGLLTNTMDVFVKEPSF-DWKNFE    QKMLKK  4023
Query_MOUSE   3952   FLP-VPELMPFRLTRQFVSLMLPMK--ETGLMCTVMVHALRAFRSCAGLLTDTMEIFVKEPSF-DWKSFE    QTMLRK  4023
Query_XENOP   3970   FLP-VPELMPFRLTRQIVNLMLPMK--DSGLFDSVMVHSLRAYRSDPGLLVTTMDVFIKEPSL-DWKNLE    LKQMKK  4041
Query_PLATY   3959   FLP-VPELMPFRLTRQLINLMSPMK--ESGLVYSVMVHALHAFRTDPDLLLNTMDVFVKEPSL-DWKNFE    TKMLKR  4030
Query_TASDE   3964   FLP-VPELMPFRLTRQFINLMLPMK--ETGLIYSVMVHALRAFRLDPGLLISTMDVFVKEPSL-DWKNFE    QKMLKK  4035
Query_CANAR   3960   FLS-VPELMPFRLTRQFVNLMMPVK--EWGLIYSVMVHALRAYRSDPDLLISTMDVFVKEPSL-DWKNFE    QRQLKK  4031
Query_OSTRI   3930   FLP-VPELMPFRLTRQFVNLMMPVK--EWGLIYSVMVHALRAYRADPDLLISTMDVFVKEPSL-DWKNFE    QRQLKK  4001
Query_TURTL   3968   FLQ-VPELMPFRLTRQFINLMLPMK--ESGLIYSVMVHSLRAYRTEPDLLVSTMDVFVKEPSL-DWQNFE    LKQLKK  4039
Query_GECKO   3970   FLP-VPELMPFRLTRQFINLMLPMK--ETGLIYSVMVHSLKAYRVYPDLLINTMDIFVKEPSL-DWKNFE    LKQLKK  4041
Query_DANIO   3943   FLP-VPELMPFRLTRQFINLMRPLA--ESGLIQSVMVHSLRAFRAEPDLLLNTMDVFVKEPSL-DWKNFE    LKQLKK  4014
Query_SNAKE   3954   FLQ-LPELMPFRLTRQFVNLMLPMK--ESGLMYSVMVHSLRAYRINQDVLINTMDIFVKEPSL-DWKNFE    LKQLRK  4025
Query_STERL   3974   FLP-VPELMPFRLTRQFLNLMLPMR--ESGQIYSVMVHSLRAYRADPDLLINTMDVFVKEPSL-DWKNFE    LKQMKK  4045
Query_MILII   3965   FLP-VPELMPFRLTRQILNLMLPMK--ESGLIRSIMVHGLRAFRLDPDLLLSTMDIFVKEPSL-DWKNLE    LKQMKK  4036
Query_BAMBO   3958   FLP-VPELMPFRLTHQFLNLMLPMK--ESGLIYSTMVHGLRAFRLDPDLLLNTMDVFVKEPSL-DWKNFE    LKQLKK  4029
Query_ALLIG   3928   MLP-VPELMPFRLTRQFINLMLPMK--ESGLIYSIMTHALRAYRKDPDLLINTMDVFVKEPSL-DWKNFE    LRLLKK  3999
Query_GAR__   3971   FLT-VPELMPFRFTRQFLNLMLPMG--VSGLISSVMAHALRAYRAEPDLLLNTMDVFVKEPSL-DWKNFE    LKQLKK  4042
Query_TETRA   4180   GLA-VPELMPFRLTRCFTNLMSPIG--INGIFRQSMISSMMALKKKRNILLEFCEVFINDPLV-DWVKLT[21]TYNAEQ  4272
Query_PARAM   3784   ALP-IPELMPFRLTQVFEGLAKPIG--LGGLYKSQFVKIMSALRKKRHILLDFCEVFINDPLI-EQIKIG[ 9]RQRLEE  3864
Query_SLEMN   4052   QLG-IPELMPFRLTQQIEGVIAPHP--LEGVYKQTMVHALNALKHRKSLIMDTCEIFVKEPLL-DWVKDA[14]VDSLIS  4137
Query_STENT   3467   GLN-LPELMPFRMTRQFLSLMSPEG--PQGELRQSMIICLKALRKHKNIILDCCEVFRNEPLL-EWIKPL[ 1]-----T  3534
Query_NEMAT   3936   FLP-VPELVPFRLTRQFTNLLLPLK--ESGLVRNTMIHVMRALRGNHELLLNTMDVFVKEPSL-DWQMFA    RKQAED  4007
Query_PISTI   3997   FQP-LPELMPFRLTQQFVNLMLPLK--KSGTLQSVMIHTLRALRSNHDLLLNTMDVFIQEPSL-DWQVYA    RKQAKT  4068
Query_CORAL   3970   FLP-LPELMPFRLTRQICNMMEPLR--VSGQLQSTMVYTMRALRNNHEILLNTMDVFVKEPSL-DWQNFA    RKQAEK  4041
Query_HYDRA   3789   LLK-YPELVPFRLTQQMLNLLLPLK-dV-SPYKQSMIVVLQALQQHSKMLISTMDIFINEPLI-DWQNNA    KKYE--  3858
Query_LATIM   3971   FLP-VPELMPFRLTSQFLNLMLPMK--EQGLLYSVMVHALRAYRANPDLLLNTMDVFVKEPSL-DWKNFE    LKQLKK  4042
Query_SHRIM   3894   FLP-VPELMPFRLSPQIVNVFQPIG--QVGMLKEIMVSALGALQESRHVLTAVLEAFVKEPTK-DWL---    --DFVQ  3960
Query_MCRAB   3885   VLP-FPELMPFRLTPQIVNVFQPLG--PVGMIKDVMVAVLGTLQESRHVILSVLEAFVKEPTE-DWK---    --EFVR  3951
Query_AMPHI   3923   FLP-LPELMPFRMTPQIVNVAQPFT--TKGIIRESMVSCLQALVSNKKVILAAVEAFVKEPTQ-NWL---    --EFVK  3989
Query_BARNA   3698   LLP-VPELVPFRLTAQLAGLTAPLG--PAGLTRHTLLRAVRALRSSRHLLAAALQGFVREPAA-QWTEHQ    ------  3763
Query_COPEP   3703   FLQ-IPELVQVRLTPQILAVMEPLG--THGLFRETMVQVLRCLREHPEMLISALNTFLQEPTL-DWIRSA[ 6]-KKIKK  3779
Query_DAPHN   3505   QLP-VPEVVPCRLTAQFQFLIPGLS--LNGSVRDSMVHTLRVAKAESGSIIAALSVFVTEPTL-NWLKEA    EKMTEN  3576
Query_STARF   4019   FLP-VPELMPFRLTRQIVNLMLPFK--KEGILQSSMVHTLRALRQNHELLINTMDVFIKEPSL-DWKTNA    SKQAKA  4090
Query_STRON   4000   FLP-IPELIPFRVTRQIINLMLPMK--IDGVLQSTMVHTLRALRQNHELLVNTMDVFVKEPSL-DWKCFA    VKQADY  4071
Query_SEACU   3964   ILP-IPELMPFRLTRQIVNLMMPMK--IDGLLKNTMVHCLRALRQNHDLLLNTMDVFIKEPSV-DWKQFA    DKTASK  4035
Query_FUNGI   3832   FLP-IPELMPFRLTRQFIGLLSPLD--ADGFLKFNMTYTLEALQNGKEKLLNTMDVFIREPLI-DW----    -E----  3894
Query_DPURP   3994   FLP-IPELMPFRLTRQFTSFLRPLD--SVGLLNHNMTYSLSALQSHKDILLTTMDVFVKEPLL-DW----    -S----  4056
Query_HALBU   3987   FLP-IPELMPFRLTRQFTSFLMPLD--SVGLLSHNMTHTLTAIQNNKDFLLSTMDVFVKEPLI-DW----    -I----  4049
Query_GLOBO   3710   VLP-IPELVPFRLTNQIEKFLLPLG--VQNLMVHPMTNTLAAIQEEKDGLINALNIFVNEPLI-EWRKFA[ 1]LQMQRQ  3782
Query_PALUS   3775   VLP-IPELVPFRLTRQMELALEPLG--IPVLLEYPMINVLTALRNRKDVLLNALNIFIKEPLM-EWRRFA[ 1]SQAQKQ  3847
Query_CALIF   3882   VLA-IPEMVPFRLTSQMVDCLDPLG--ISVLLEIPMIYIQTVLRENSEILLNAMSIFINEPLL-DWRNFA[ 7]KIKPKE  3960
Query_GLOMU   3893   TLN-IPELIPFRLTKQIETLMEPLG--SKGLLEYPMIKILQIMQANKDILLNSMDIFVKEPLL-DW----[ 7]AKQQKD  3967
Query_LOBOS   3869   VLP-VPELAPFRLTRQLVAFLNPLG--TKGLLEHPMVCIMKALQAKKDVILNTMDVFVKEPLL-DWRKYA[ 7]KKRGAD  3947
Query_RACEM   3851   QLP-VPELVPFRLTRQLVGALAPLG--VTGILETPMTHIMKAFRSEKSLLLNTMDVFIKEPLL-DWRKRA[ 7]KRDNSQ  3929
Query_PVIOL   3987   FLP-IPELMPFRLTRQFTSFLAPLE--SVGLLNHSMTYTLGALESNKDILLSTMDVFVKEPLL-DW----    -S----  4049
Query_BOMBU   3677   DQR-IPELMPFRLTPQILGLLKPFT--EKDLLKTIMIHTLQAVRNDQGPILSCMDVFVHEP-L-NWTEHL[ 3]LRESE-  3749
Query_AEDES   3830   RDQgIPEMVPFRLTPQFVNVMEPMR--TSGMMSKCMVYTLRCLRDSRKLLRSCLQVFVREPTV-DW----[ 2]------  3894
Query_ANT__   3755   RAP-IPELVPFRLTPQILELLQPFT--ERDFLATIMTHAMRALRNDQGPILACMDIFVHKP-I-QRSFSI[ 3]EIMKD-  3827
Query_WHITE   3819   ILP-VPELIPFRLTPHILDLAEPFK--EFGLLRESMVSALKCQKDNQHVLLATMQVFIQEPSM-DWFENA[ 2]LAQKQS  3892
Query_BEETL   3644   ILP-VPELVPFRLTPHIVNLMEPLG--ATGLFREIMVHCLRSLRSNSASLLSTMNVFIEEPSL-DWLEHA[ 2]------  3711
Query_THRIP   3739   NLP-VPELIPFRLTPHIVHVTQPHG--TKGLIQETMVNALRALRADPEILLATMQVFIQEPSL-DWL---[ 1]LAKKH-  3807
Query_TERMI   3823   FLP-VPELMPFRLTPHIVSLLQPLE--ETGLFRETMIHCLRTLRNNYDMLLATMEVFVQEPSI-DWLEFA[ 5]LNKKS-  3898
Query_FLEA_   3730   TLP-IPELLPIRLTPHIMNLMEPTA--AFGLIPRVMVHALTALRDARSVLYATMNIFMHEPLI-NWLEPI[ 3]------  3798
Query_MIDGE   3849   TLS-IPELVPFRLTSHFVNVVEPLG--VNGMIRKNMNHVMRCFRHFSQTILICLEMFVREPTL-DW----[ 2]------  3912
Query_LAODE   3795   VLA-IPELLPFRLTPQMLNLFEPFR--VTGLIKESMFRCLGILRKRKLLIMASIEVFIHEPLM-DWAKQA    ------  3860
Query_LOTTI   3909   FLP-VPELMPFRLTRQIINLTLPLG--IKGLLENTMIHTTRALRQDYDLLVNTMDVFVKEPSV-DWLNFA    EKQRNE  3980
Query_APLYS   4015   FLP-VPELVPFRLTRQLVNLTMPLP--VKGQMESSMCHVLRALRRDSDLLLSTMDVFVKEPSL-DWLDLA    ERQMNE  4086
Query_SCALL   3998   FLL-IPELMPFRLTKQLRNLMLPLQ--VKGLMESTMVHALRALREDYDLLLATMDIFVKEPSV-DWLVNA    EKQIQS  4069
Query_CRASS   3941   FLP-IPELIPFRLTRQLRNLMMPLQ--VHGLMESTMIHTLRALRNNCDLLLNTMDIFVKEPSV-DWLVNA    EKQMNE  4012
Query_OCTOP   3850   YLL-IPELIPFRLTRQLQNVCLPHG--TKGTFERTMVHCLESLRSDKDLLIHTLEIFVREPSF-DWLKNA    RNNMHR  3921
Query_POMAC   3948   FIP-VPELMVFRLTRQIVNLGLPFK--VKGLLENTMIHVLHALRQNHDLLLSTMDVFVKEVSL-DWLQFA    ERQMRG  4019
Query_WHIPW   3799   LLP-IPELMPFRLTNQLCGVCDPVG--VNGLLRSSMVNFLRSLRDNGQLLLNAMDIFIKDPSL-DWTREA    LK----  3866
Query_TRICH   3805   YLP-IPELMPFRLTKQFVGVVDPVG--MHGLLESTMKYCLKAFRDSSYILLNTMDTFVKEPSL-NWIVEA    RRQIGE  3876
Query_AASTA   3601   LLP-VPELVPFRYTKQLRGVLQPHD--ASLLLQQDLAAVLDALRGQQQRIDSVMRVFLNEPLL-EWQTQS    ------  3666
Query_AEUTE   3524   MLP-VPELVPFRFTRQLQGVLQPHN--AKLLFQQDLAAVLEALRGQHQRIESVMSVFLNEPLL-EWQVTV    ------  3589
Query_PLASM   3797   VLP-VPELIPFRYTRQMNFVFQPYD--GPSLLAREMQTVFDALRSKRQVVESVMNVFLHDPLL-DWQQSI[22]NAVMKD  3890
Query_BREMI   3780   VLP-IPELIPFRYTRQMEFVFQPYD--GANLVCQDMQAVFDALRSKRQVIESVMNVFLHDPLL-DWQQST[22]DAEMED  3873
Query_OLIGA   3833   VLP-VPELIPFRFTRQMEAVLQPYD--GSNLLIQDMQAVFEALREKKQVIESVMNVFLHEPLL-DWQQST[18]EDVAME  3922
Query_LIGNO   4226   FLP-TPELVPFRLTRQICGAMGPLP-pSQGAFVRCLGHCLAALRRERELVLSVLQVFINDVSA-DWKQLA[14]DTSEHR  4312
Query_CESTO   4015   SLP-LPEYVPIRLTSSLRELLEPSG--PAGLFGSILSRTLASLRYNSGLFLSSIQTFMKDKSTtDWSVFS    ----ER  4083
Query_OVI2_   4505   ILP-VPELVPFRLTACQRELLEPSG--PAGMFGYTLSRTLAAARMARPLLTSLLQVFVKDSSC--WALYS[ 2]TQQPQT  4577
Query_PRIAP   3984   CLP-IPELMPVRLTRQFLQLLLPHEgqRRGLLYQTMVHTLRALRDNHRLLLNTLDVFIKEPSI-DWKAT-    ---AER  4053
Query_ROTIF   3883   NIP-VPELVPIRLTRQILKLIAPLE--QRGLFESSMVHCLKALRENNDLLVCILDVFIKEPSI-DWIGSA    VKIAKK  3954
Query_SPONG   3986   FLP-LPELMPFRLTPQFVNLLLPHK--VSGQLRSCMIHTLRALRHSPTLLLNTMDVFIKEPSL-EWESNA    KKQA--  4055
Query_TARDI   3942   SLP-IAELAPIRLTRQFFNLMAPQC--VKGLLQQTMVHAMRALRAERQILLATMDVFIQEPSM-DWLQAA    RKQSPT  4013
Query_TRICH   3958   FLP-LPELMPFRLTRQMLNLMLPLR--ENGQLKCSMVHTLRALRNKPDLLVNTMDVFIKEPSL-DWQNFA    NRQNAK  4029
Query_OSTRE   3776   ALP-IPELVPFRATPALLDVLAPMS--ARTWLETDMARTVKALQNGSTLLKGVMDIFLREPLI-DWEREA[ 3]SRSSKS  3850
Query_NITEN   4129   MLP-VPELVPFRLTRQFAQLLRPLD--TLDLLKIDMTRVMTVLQENKDVLVAVMDVFVKEPLV-DWRNES[ 8]AK----  4204
Query_CLUBM   3740   LLP-VPELIPFRLTPQLTNVLLPLD--SLGLLKNDMVRVLSALHKSRQMICAVMDVFVREPLV-DW----    -K----  3802
Query_MOSS_   3907   LLP-VPELVPFRLTAQLKNFLLPLT--AVGLLRSDMVRIMTALRGSRGLISAVMEVFVKEPLV-DW----    -RQEAV  3973
Query_LIVER   3972   LLP-VPELIPFRLTNQLLNFLLPLD--ALGLLRTDMVHTLAALHASREVISAVMDVFVKEPLV-DW----    -K----  4034
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  4004  M   SNRKEDTS[2]T    LDSQWFPKE   KVNIVRQKLEGSNPSH   VMRAELALG--Q-LGS---YLKYFVI--V  4061
Query_LINGU  4013  T   MSSQEDK-[1]-    -DAKWYPKQ   KVAFAKRKLKGANPWC   ITRDELELG--H-KRR-gKTYQAFIA--I  4068
Query_LIMUL  4052  -   GKA-----   -[ 6]-DTQWYPKQ   RVQFAQQKLQGINPCY   ITMLRV-LFR-S-DK--------MKL--V  4100
Query_SPIDE  3988  M   GM------   -[ 5]DD--WFPRQ   RILIARKKLEEFNPCY   ITRKELELGH-S-MRP---CFQNMVS--V  4040
Query_SCORP  4010  I   GKVDENVD[8]N[10]CDFNWYPKQ   RILFAKRKLEGINPCY   ITKEELRLGH-S-NKP---IFQRLTD--V  4084
Query_TICK_  3824  S   SPQASKRQ[2]A[ 2]VSLDWFPRQ   KVSIVRQKLEGGHPSY   ILRDELKLR----KTN-qPDFASLEL--V  3884
Query_MONOS  4054  E[2]----EELQ[3]-    -----YATR   NLERLRLKLSGCHPAA   IMAEEVADSR-H-----iAAYRSKVQe-V  4105
Query_ROSET  4003  G   --------   -    -DVALFAEQ   KIAVAKRKLAGDNPAH   ILISELESNR-HvTSD-rSWHRKAKD--I  4052
Query_BELCH  4098  Q   MDKGMVGE[3]L    DDISWYPRE   KIKFAAKKLQGVNPTH   ITKAELQLG--H-KNL--PWFKSFCL--V  4157
Query_CIONA  4040  D   QRYSDSDV   -[ 2]SDRVWYARR   KMALTKKKLEGENPSS   ITCDELNHNStI-TKF-k---KSYCC--I  4097
Query_CAECI  4039  Q   GTWTKKVN   T    AEKNWYAVQ   KVNCARRKLAGTNPAI   ITCDELELG--H-EKS--VAYKEMKA--V  4095
Query_HUMAN  4024  G   GSWIQEIN   V    AEKNWYPRQ   KICYAKRKLAGANPAV   ITCDELLLG--H-EKA--PAFRDYVA--V  4080
Query_MOUSE  4024  G   GSWIQEIN   V    TEKNWYPQH   KIRYAKRKLAGANPAV   ITCDELYLG--H-EAS--SAFRSYTA--V  4080
Query_XENOP  4042  K   GEWKKAVD   V    TSHNWHPQQ   KIHCAKRKLDGANPCE   ITCEELRLG--H-ESA--PEYKDFIA--V  4098
Query_PLATY  4031  G   GTWTRAVN   T    ADTNWYPLQ   KISYAKRKLTGTNPAI   ITCDELRLG--H-EKV--PAFGDYIS--V  4087
Query_TASDE  4036  G   GSWTREVN   M    TEMNWYPLQ   KINYAKRKLAGANPAA   ITCDELHLG--H-EKV--PAFECYVS--V  4092
Query_CANAR  4032  G   GTWIKEIN   T    SEVNWYPLQ   KVNYVRRKLTGANPAI   ITCDELRLG--H-ERS--PAYNDFAA--V  4088
Query_OSTRI  4002  G   GTWIQEIN   T    SEVNWYPLQ   KVSYVKRKLTGCNPGR   ITCDELRLG--H-EKS--PFFSDFVA--V  4058
Query_TURTL  4040  G   GTWSKEVN   T    AEINWYPLQ   KVNCAKRKLAGANPAV   LTCDELRLG--H-EKS--PAYKEYVA--V  4096
Query_GECKO  4042  G   GTWTKDIN   M    DEVNWYPLQ   KVNFARRKLAGANPAV   ITSDELRLG--H-EKS--SAYRYYVS--V  4098
Query_DANIO  4015  G   GTWTESVN   T    KEINWFPLQ   KVNFARRKLEGTNPSV   ITSEELCLG--F-EKM--PEYKGLLA--V  4071
Query_SNAKE  4026  G   GTWTKEIN   T    DEVNWYPLQ   KVKIARRKLAGANPAV   ITCDELRLG--H-EKS--EAFREYVS--V  4082
Query_STERL  4046  A   GSWSEAVN   T    NEINWYPMQ   KVNFARRKLEGANPAV   ITSEELKLG--F-EKS--PAYQSIIA--V  4102
Query_MILII  4037  G   GTWNKDVN   T    REIHWYPMQ   KVRYARRKLEGANPAA   ITRDELQLG--H-ETT--EAYKSYVA--V  4093
Query_BAMBO  4030  E   GSWSKDVN   I    NEINWYPRQ   KVNYAKRKLEGANPSA   ITRDELQLG--H-EKS--KAYASYVA--V  4086
Query_ALLIG  4000  G   GTWLKEVN   T    CEVNWYPLQ   KVNCVKRKLAGSNPAA   ITCEELRLG--H-EKS--LAYRDYVA--V  4056
Query_GAR__  4043  G   GTWKETVN   T    KEIDWYPRQ   KVQFAQRKLEGANPAA   ITSDELKLG--F-EKE--SWCRAALA--M  4099
Query_TETRA  4273  S[6]GLSTMEAS[4]L[ 3]-------MQ   RIDIVKLKLIGMNPSH   LFTEELSSTR-H--KT-gNYFQHLKS--A  4336
Query_PARAM  3865  Q   ---SKIEE[8]S[10]IKETWYPRK   KIDIVHKKLIGVHCSL   ILLEEFDESR-HqSAR---YADNLRK--A  3937
Query_SLEMN  4138  -   -EYEEELQ[2]-    ---AWYPRK   KIEVMKDKLNGLNPVR   IMAKELQDSK-H-EKK--PYFKVILN--A  4191
Query_STENT  3535  -   -SYNVS-E[2]-    ---ADIPSF   KLEVVKKKLTGINSAL   IMLEELEKTI-H-NNS--NIQDRIAE--A  3587
Query_NEMAT  4008  Q   HMEIGMYY   -    EDLTWYPRE   KVLTAQRKLKGANPAF   VTRQELELG--H-RRR--SWFRNLQS--A  4063
Query_PISTI  4069  Q   G--INEEE   -    QKADWYPKQ   KVQSAWRKLEGFNPAY   VMRDDLQLG--H-SKQ--SWYKQMES--V  4122
Query_CORAL  4042  Q   KLNLNDM-   -    TDLAWYPKE   KINSAKRKLKGENPAE   IMKQDLESG--H-KKA--GYYEEMLS--V  4096
Query_HYDRA  3859  -   ---SLALL   -[ 2]SSPDWYPKQ   KINICKQKLNGVNPAL   ITSEELRNSFlH-TRD-qLLLESYIS--V  3915
Query_LATIM  4043  G   GSWTKEVN   T    KEVNWYPVQ   KVKCGRQKLEGVNPIV   ITCNELQLG--H-EKL--SAYQSYLA--V  4099
Query_SHRIM  3961  K   QEGDTQDS   -    -KVELFSSK   RIELLNDKLSDINPAY[4]AVSQNKFVA-----KQ--KAYENLKQ--A  4017
Query_MCRAB  3952  E   LEGNLDES   -    -KIEMFSEE   RMRCLKEKLTGFNPAY[4]ALSKNKAVR-----RD-pTAFEELKK--V  4009
Query_AMPHI  3990  L   QEDVVDQE   -    -GIQSYSSM   RMKIITKKLDGVNPCI[4]AMNKNRRIP-----QN-vKSV--LSK--I  4045
Query_BARNA  3764  -   --------   -    QQPAEHARR[1]-VRQTRGRLCGSHPAA[3]----QLLAGAgRlSAQ---EVEAWSA--A  3811
Query_COPEP  3780  K   ----EEDH[3]-    -----FPKQ   RVGMTRRKLEGVHPSI   ITCEELRAGV---DKS---DLNKLVS--A  3828
Query_DAPHN  3577  R[2]ATFTKEST[1]-    ----WSPTQ[1]-VGRTDNLLNGFHPSV   ITCQDLQENKiFlTADgqKLLPQLIA--V  3636
Query_STARF  4091  Q   KDAGRVLE[8]S    LDESWYPRE   KVNFAQKKLEGANPSD   ITQRELALG--H-SRN--PAFSAIKE--V  4155
Query_STRON  4072  Q   KLSEDSRE   -    -DISWYPKK   KIKFAKMKLQGVNSTY   ITKAELEMGP-H-SKD--ARDAPFLK--A  4127
Query_SEACU  4036  G   QGLEEI--   S    SDALWYPKE   RLLIAKAKLEGANPAY   ITERELEKSV-H-GKS--TSLQKFIA--V  4091
Query_FUNGI  3895  -   ----RLAK[5]Q[ 5]SQA-WFPKK   KIEIAKKKLQGENPAH   ITALELKESV-H-AN--aPYLPHLQQ--I  3956
Query_DPURP  4057  -   ----KLAN[5]Q[ 6]TKNVWFPKQ   KIHIAKKKLDLYNPAY   ITLEELSASV-H-SN--sPYEKNLQE--I  4120
Query_HALBU  4050  -   ----KLAS[5]Q[ 6]VES-WFPKE   KVDIARRKLERWNPAY   ITSEELANSV-H-KG--qIYEKSIQN--V  4112
Query_GLOBO  3783  G   -KGTVAIS[5]-[ 1]AAPEWYPRQ   KLEIARRKLDRENPIF   ITADELKNGH-E-K---kKWYKNMRV--T  3843
Query_PALUS  3848  G   -KTSQDTS[5]I[ 1]NPPEWYPQQ   KLDIARRKLEGDNPPY   ITTQELEWGH-G-E---kPWFQAVRD--V  3909
Query_CALIF  3961  -   -------E[5]T[ 1]EPPKWYPKQ   KLNIAERKLKGENPSN   ILITELENGH-S-N---tTFLRYLKE--I  4015
Query_GLOMU  3968  -   ----RNSN[5]D[ 1]QTVEWYPRQ   KLDIARRKLELENPAY   IVCSELEAGH-S-N---kPWFNSIIK--I  4025
Query_LOBOS  3948  -   ----MDSF[5]S[ 1]APPAWYLQQ   KMDIARKKLEGHNPVY   LTVQELNMGH-A-N---kPFLAAITK--I  4005
Query_RACEM  3930  G   -SPASDIS[5]F[ 1]AEIAWYPYQ   KIDIAKRKLAGENPSY   IVATEVRNND-S-PHV-kAAIDRIEE--V  3993
Query_PVIOL  4050  -   ----KLAN[5]Q[ 6]TKNEWFPKQ   KVALAKKKLEMYNPAY   ITLEELAGSV-H-NG--lAYDSALQL--I  4113
Query_BOMBU  3750  -   ------ED[3]V    K---WVPIR   KIEAVTKKLNGIKPSL   ITLDQLKEQHhD------KYFDRYYA--I  3798
Query_AEDES  3895  -[3]RRFVQEEN[3]L    R---WDPQA   RINMAVRKLNGANPKV   LIAEELRLGQvG-CNR-eILEGYL----K  3954
Query_ANT__  3828  -   ------DN[3]L    I---WSSKR   NIEIVAKKLNGIHPSI   ITLEHLKEAHnD------EYFARYCA--I  3876
Query_WHITE  3893  G[6]SRSNYFKD[3]S    -----FSLR   NIELAWKKLEGGHPTE   IMIKELEAG----HKT-sPYLDAFKK--I  3953
Query_BEETL  3712  -   -----FST[3]T    SNCDWYPAQ   KIEQARKKLAGANSVN   IMIEDLLAGHq---KN-qAYMKAYIA--L  3766
Query_THRIP  3808  -   -----WSE[3]N    ANPAWLPVV   KIHQAKKKLKGTNPVH   LMREDLESKIhTvFKD-g-----YLK--T  3860
Query_TERMI  3899  -   -----HEE[2]-    ---TWYPQQ   KLDLAKLKLEGANPVT   ITTEELKAG----FKD-sKYLAKFLE--C  3947
Query_FLEA_  3799  -[3]SEDNEALN[3]S    SISSVENEQ[5]KLRMARLKLQGSNPVH   ILQDELKHSFvS-TIN-eIYVNHYIE--K  3868
Query_MIDGE  3913  -[3]MKKNIENA[3]M    GISDWNPEQ   RIAIVERKLKGENPVR   IVKDEMSVSEiA-NNK-qLYDAYC----N  3975
Query_LAODE  3861  -   -KYSAILN[2]-[ 4]TDYNSRPRE   KIERARKKLDGANPVK   VMTEELESGH----KN-kPYLEYLKS--M  3920
Query_LOTTI  3981  M   LMTADDE-   -    -ESTWYPKQ   KLQYALRKFKGDNPCH   ITKDELYLG--H-SRS--PAIKHFEK--C  4034
Query_APLYS  4087  G   KESEDAA-[1]-    -DDNWYPKQ   KIQFARRKLRGDHPCY   IMKEELALG--H-SKR--PAFKSLVN--V  4141
Query_SCALL  4070  L   NLGDSLEV[1]-    -NITWYPKE   KVHYVKRKLKGENPSY   ITRDELRLG--H-SKN--DFLRNFES--V  4125
Query_CRASS  4013  M   KVEVSPE-[1]-    -EVQWYPKE   KIEYVKRKLRGDNPVC   ISREELKLG--H-SKK--PAFPALER--V  4067
Query_OCTOP  3922  K   AQDFSE--   -[ 1]EDLSWYPKE   KIDFVKRKLDGEHPAY   ITRDELNLNS-H-VRK-sDALESITS--I  3978
Query_POMAC  4020  G   MGPKDAD-[1]-    -DMRWYPKQ   KIETAARKLKGDNPSY   ITRDELRIG--H-LRN--QALKDFEK--V  4074
Query_WHIPW  3867  -   QTKLSDSS   S[ 2]GRLSWYPKQ   KITVARRKLDGDNPVS   IMEAEL-SARqH-TDN--FPLQQILC--C  3925
Query_TRICH  3877  G   RTLWNDEA   N[ 2]DSFKWYPEE   KISAAARKLRGDNPVI   IMQDELMNSRrH-LSS--FPLHQIIN--C  3937
Query_AASTA  3667  -   ------KT[8]T[ 7]AAATWMPQV   KMDLARRKLRGEHVVH   ILCDELALNP-NvAAV-lKPLQATLP--P  3734
Query_AEUTE  3590  -   ------KK[6]-[ 6]ASSSWLPEL   KMQLAQRKLRGEHVVH   ILREELQLNP-HiASV-ySHFDQIIP--V  3653
Query_PLASM  3891  V[6]RESQSRKK[8]N[ 1]-TTAWLPNV   KIAIARRKLEGVSPAL   LLKQELSQNP-HiKQH-lNKYYSLVN--A  3964
Query_BREMI  3874  V[6]R---SNKN[8]I[ 1]-NAAWLPDI   KIAIARRKLDGISPAL   LLKEELSQNS-HlTHH-lSKFYALVD--A  3944
Query_OLIGA  3923  V[6]SKRQKRST[8]S[ 7]TTLAWLPDV   KVAIARRKLEGFSPRA   LLKEELAQNP-HlGKQ-lKLFQALVD--T  4003
Query_LIGNO  4313  A   GSKSASAS[8]S[ 5]CPERWFSTE   KLQVVRQKLSGAHPCQ   TLKSELLLDR-H-RGK--PYTSGMLA--V  4383
Query_CESTO  4084  Y   --------   -[ 1]QSQEEYQRS   RLSLIRRKLIGHCPAE   VLIDDLRGRFgS--QE---WFNRFEM--I  4132
Query_OVI2_  4578  E   --------   -    -----FTRG   RLELALKKLSGHPPAT   LLCEEVAIRYgS--HAwfPEFRRLINqlV  4625
Query_PRIAP  4054  L   ADKDIHSN[5]E    QVQEWYPKQ   TIGVARSKLIGANPAH   ITKAELALG--H-RKH--EALKQMTN--V  4115
Query_ROTIF  3955  N   LNEDMEIT[4]-    -----YAKD   RVNSVKQKLFGMHPSV   IMRNDLCSGI-H-KDS--KYLQYFVD--V  4010
Query_SPONG  4056  -   --LAQKIN   T    DDNTWYPKE   KVNICRKKLLGFNSSH   LMLIDLEHG--F-KNK--PEFKYLKG--I  4109
Query_TARDI  4014  -   -----DAD   S[ 1]EELLWYPRE   KIDTAERKLRGGNPAF   ITRTEVRSNTfL-KSD-qRLLAKVED--A  4068
Query_TRICH  4030  Q   G--LKTES   -    DDFSWYPRE   KVTFAKRKLKGANPAF   VTRDELRLC--H-H------LSILES--F  4079
Query_OSTRE  3851  S   --------   -    -----HIQS   RITHAWGKLELDNPAK   LVLEQCASR--HqDRS---YWTNMCE--A  3893
Query_NITEN  4205  -[1]RKGAVAAL[5]D[ 4]VLEQEHVSL   KIENAARKLARWNPAE   VTLAELQQSM-V-NG--kPFFRSLEG--I  4271
Query_CLUBM  3803  -   KEAARTAS[5]-[ 2]SHEQQHVEL   KIENAQRKLDLWNPAH   VTLGELQSSV-H-AG--tPWMHAIEA--A  3865
Query_MOSS_  3974  K[6]RDVSKSTA[5]G[ 7]FSEEEHVQL   KVETAQCKLDLWNPAE   ITISELGFSV-H-AN--kPYTRALEE--I  4049
Query_LIVER  4035  -   NEAMKTSS[5]K[ 7]TFEQQHVEL   KVENAQRKLELWNPAA   ITIAELRSSV-H-AR--kPYSKALES--I  4103
```

Supplementary Figure 7: DNA-PKcs: All phyla - expanded list: COBALT

```
Query_CAPIT  4062   LMGDK    QADIRARL   PAD---GLSSEEQVAALLNLATDPNILGRTFAGWEPWV  4109
Query_LINGU  4069   TQGQ-    -DNTRTKF   AEV---DLTPEQQVACLIDQATDPNILGRAYEGWEPWV  4114
Query_LIMUL  4101   CLGDK    EDNIRAKL   PNA---GLTVEQQVACLIDQATDFNILARIYWGWKQWL  4148
Query_SPIDE  4041   CIGDA    SHNLRAQL   KKS---GLSVEEKIESLIDQATDPYLLCLTYFGWNPWM  4088
Query_SCORP  4085   CLGDR    K-SLRQNL   PKN---GLSVKQQVKSLIDQATDSGILSRIYFGWDPWF  4131
Query_TICK_  3885   CLGAP[ 2]ESRMRRRF   ARGdk--LSPRDQVDCLLEQATDPAILGLTWTGWEPWL  3935
Query_MONOS  4106   VWGSR[ 1]DQDLRATL[2]--Eva--LTVEQQVDALLDQATDPNLLGRSWVGWDPYC  4155
Query_ROSET  4053   IAGDA    QSNIRATM[3]------GLSVEDQVRCLLDQATDPCILGVTWAGWEPWK  4100
Query_BELCH  4158   AMGDG    GKDVRAQK   PAE---GLSVEDQVACLIDQATDPNILGRTWQGWQSFM  4205
Query_CIONA  4098   ASGDT    K-NIRFLM   -PK-s-GLSVSEQTDCLLDLATDPNVLGRMWVGWKPWI  4144
Query_CAECIL 4096   AMGDK    THNVRARE   PKD---GLTEEIQVKCLLDQATDPNILGRVWEGWEPWM  4143
Query_HUMAN  4081   ARGSK    DHNIRAQE   PES---GLSEETQVKCLMDQATDPNILGRTWEGWEPWM  4128
Query_MOUSE  4081   ARGNR    DYNIRAQE   PES---GLSEETQVKCLVDQATDPNILGRTWEGWEPWM  4128
Query_XENOP  4099   ARGDK    KHNRRTNE   PPD---GLTEETQVQCLIDQATDPNILGRVWKGWEPWI  4146
Query_PLATY  4088   ARGSK    DHNVRARQ   PEE---GLSEEQQVQCLIDQATDPNILGRTWIGWEPWM  4135
Query_TASDE  4093   AQGNR    DHNFRAQQ   PDD---GLSEETQVKCLMDQATDPNVLGRTWIGWEPWM  4140
Query_CANAR  4089   ARGNS    NHDIRAKE   PED---GLSEETQVRCLIDQATDPNILGRVWEGWEPWM  4136
Query_OSTRI  4059   ARGNA    AHNTRAKE   PED---GLSEETQVKCLIDQATDPNILGRAWEGWEPWM  4106
Query_TURTL  4097   AQGSR    DHNIRAKE   PED---GLSEEIQVRCLLDQATDPNLLGRVWEGWEPWM  4144
Query_GECKO  4099   ARGSK    EHNIRAKE   AED---GLSEEIQVKCLIDQATDSNILGRVWEGWEPWM  4146
Query_DANIO  4072   ARGEE    QHNIRARL   ADK---DLTVEDQVDCLLDQATDPNILGRVWIGWEPWI  4119
Query_SNAKE  4083   AQGSK    QFNIRARQ   PQD---GLAEEIQVKCLIDQATDPNILGRTWEGWEPWM  4130
Query_STERL  4103   AQGEK    EHNVRARE   APQ---GLSVETQVECLIDQATDPNILGRVWVGWESWV  4150
Query_MILII  4094   AMGDK    EHNIRARV   PED---GLSVEIQVDCLIDQATDPNILGRVWEGWEPWM  4141
Query_BAMBO  4087   AMGDK    DHNVRARE   SEE---GLSVECQVDCLIDQATDPNILGRVWIGWEAWM  4134
Query_ALLIG  4057   ARGTR    DQNIRAKE   PED---GLTEETQVKCLIDQATDPNILGRVWAGWESWM  4104
Query_GAR__  4100   AQGEQ    GVNVRASQ   AAE---GLAVEIQVECLLDQATDPNILGRVYMGWEPWI  4147
Query_TETRA  4337   LKGKP    -KSIRNQY[1]-EKqf--LEVHEQVDCLIDLAVDPNIIGRVWIGWAPII  4384
Query_PARAM  3938   IQGIN    QNRLRHQY[3]-----qTLSVIQQVDCLIDLATDPNILGRAWKGWAPFI  3986
Query_SLEMN  4192   IKG-Q    ESSIRYKL[3]-NRry--LEVDEFIDCLIEQSRDPNILGRTWIGWSPYV  4241
Query_STENT  3588   VLG-P    KDSLRNN-   -NSnv--LSVDLQVDILLEHATDPNILGRTWIGWSPTV  3633
Query_NEMAT  4064   CLGDR    RHNVRAQK   EEN---ELTSEAQVDCLIDQATDVNILGRAWEGWEAWV  4111
Query_PISTI  4123   CLGDR    NANVRARE   PES---GLSVESQVSCLIDQATDANILGRTWQGWEPWV  4170
Query_CORAL  4097   LLGDK    QHNQRARH   YVN-----TVENQVACLIDQATDPNILGRTYHGWEPWV  4142
Query_HYDRA  3916   VEGSS    QHNIRKKL[6]PEK-e-FLTVEEQVDCLVDQATDPNLLGRAWIGWFPWI  3970
Query_LATIM  4100   VKGDP    DHNVRARE   VKE---DLPVETQVACLIDQATDPNILGRVWEGWEPWM  4147
Query_SHRIM  4018   VLG-D[ 1]GESVRASV   GER---GLSTHQQVDILLEQATDPNILGRSWIGWDPFL  4065
Query_MCRAB  4010   VLG-Q[ 1]CESARANM   DRT---GLTIHQQVDILLEQATDPNILGRTWVGWSPFI  4057
Query_AMPHI  4046   IMGQD[ 1]GRALMA--   -KE---GLSVPQQVDVLLELASDPNILGRTYRGWSPVT  4091
Query_BARNA  3812   ARGRR    GHEPRAEL[3]------GLTAEQQVDCLLELATDRYVLGLAYSGWAAHV  3859
Query_COPEP  3829   VEGER[ 1]STSGKRKD[3]-DDet-GLSPSDQVDILIHQSTDKNLLGRMYVGWDPFA  3881
Query_DAPHN  3637   VRGNR[ 6]-ETIKRKV[3]SSRra---SMPQQIECLIEQSTDPACLSRMYWGWEPWI  3692
Query_STARF  4156   VKGDR    SQNIRASS   PK--s-GLSVETQVECLIDQATDPNILGRVYAGWEAWM  4203
Query_STRON  4128   LKSTC    GGNTESKR   AQSkskGLSVEDQVACLIDQATDPRILGKTYQGWEPWV  4178
Query_SEACU  4092   VRGDH    RVNIRAQV   AREps-GLSVEDQVACLIEQATDPNILGRTYQGWQPFI  4141
Query_FUNGI  3957   VMGSK    THNIRARV   -GTkc--SSAKEQVDCLVDQATDSNILGRSWFGWAPWM  4004
Query_DPURP  4121   IKGDP    KYNIRSKV   -NKic--SSVKEQVDCLIDQSTDPNILSRAWIGWNGAL  4168
Query_HALBU  4113   VKGDP    KHNIRAKA   -CQic--STAKEQVDCLIDQSTDPNILARAWAGWSSWI  4160
Query_GLOBO  3844   LMGDP    VINIRAK-   AGNic--QTPKEQVECLIDLAMDPNVLGRAWLGWSPFM  3891
Query_PALUS  3910   VMGEK    GTNRRSD-   VGRkc--SSVKEQVECLIDQATDPNILGRAWRGWGAWC  3957
Query_CALIF  4016   VKGDP    KYNIRAQP[7]EGSnv--LSPIQQVKCLIDLATDKNILGRAWIGWSPFV  4071
Query_GLOMU  4026   AKGDP    NYNIRAR-   ARQkc--FSVKEQVECLIDLSTDPYVLGSMWVGWQSWI  4073
Query_LOBOS  4006   AMGDS    LHNVRAR-   HGRvc--ASVQAQIECLIDMATDLDILGRAWVGWMSWV  4053
Query_RACEM  3994   IMGDE    KHNVRAR-   IGPvc--SDTRQQVQALIDLATDPNILARIWGGWQPY-  4040
Query_PVIOL  4114   TKGDI    KYNIRAKV   -NKvc--SSVKEQVDCLIDQASDPNILSRTWVGWNGAL  4161
Query_BOMBU  3799   VTG-D[ 1]-DIKRTRA[2]-NG---FLTPEEQIECLLDQATDLNILGRTHVGWKPWL  3846
Query_AEDES  3955   LLQVS[ 1]-PLLESG-   --N----LTVEQQVQCLLKTATSGAVLGITYAGWFPWF  3998
Query_ANT__  3877   VTG-K[ 1]-ELKQARS[5]-ND---HLTPAQQVDCLLDQAKDLNILGRMYANWQSWL  3927
Query_WHITE  3954   LQGTN[ 1]--SLRFKF[2]-DQ---ELNANQQVDILLDQAKDKHILGITYAGWEPWA  4001
Query_BEETL  3767   VEGNS[ 1]-FNVRARL[2]-NG----LSEEDQAICLIDHATDYNLLGRMFEGWAAWV  3814
Query_THRIP  3861   LNGDA[ 1]-RNIRAKL[2]-ND----LTPRQQVECLIDLATDPNVLVKTYGGWEPWL  3908
Query_TERMI  3948   VKGHN[ 1]---VRSQM[2]-YQ----LLPEEQVDSLLDQATDYNILGKTWQGWEPWM  3993
Query_FLEA_  3869   GMGVE[ 1]-HNIRRRI[2]-KD----LTAKEQVDCLIDLATDRSILGCTWHGWDPWM  3916
Query_MIDGE  3976   LAEGD[ 1]-NNFRSQI[2]-KD----LSVEDQVQCLIEMATDKAILANMYLGWDPWI  4023
Query_LAODE  3921   LKGEK[ 1]--NVRSSK   PQY---GLTVREQVECLIDLATDEHILGKSYVGLRAWV  3967
Query_LOTTI  4035   VLGDA    KDNVRATL   PST---GLTVEQQVEALIDQATDPNILGRVWQGWEAWV  4082
Query_APLYS  4142   LMG--    KDSVRAGL   PAH---GLSVEQQVAALIDQTTDPNILGRTWGGWEPWV  4187
Query_SCALL  4126   VMGE-    RDDTRAQL   PQD---SLSVEDQVAALIDQATDPNILGRVYAGWEPWM  4172
Query_CRASS  4068   LQGDK    RENVRAQL   PAS---GLTVEQQVAALIDQATDPNIVGRVWGGWEPWM  4115
Query_OCTOP  3979   VLGTE    G-NVRYTL   PQK---HLTVEQQITALIDLATDPNILIRTFEGWNPLL  4025
Query_POMAC  4075   ALGDV    QDNIRARL   KEK---GLTEEQQVAALIDQATDPNILGRAWGGWEPWL  4122
Query_WHIPW  3926   IYGAD[ 6]ETRRNAPR[2]SSR---SISVDDQVSCLIELAVDPLVLARTWHGWEPWI  3981
Query_TRICH  3938   IYGKK[ 6]EEDQSAPR[2]PLQ---ALTVEQQTSCLIEMATDSVILSRTWFGWEAWI  3993
Query_AASTA  3735   PA-AP    --------   --Khs-VLSPVDQAKALIHVATAPNVLGRMFHGWSPWA  3773
Query_AEUTE  3654   PSDVP    --------   --Khg-FLSPLDQAKALIDIATDANVLGRMFHGWSSWA  3693
Query_PLASM  3965   ASTG-[ 3]--------   ADEia-ALSSLEQAQQLLAIATAPDLLGRTFQGWMPWL  4008
Query_BREMI  3945   ASFGD[ 4]--------   DNEvt-TLSSLGQAQQLLAIATAPDLLGRTFQGWMPWV  3990
Query_OLIGA  4004   ANAQA[ 4]--------   STTsg-AMSSLAQAQDLLALASAPDILGRTFQGWMPWL  4049
Query_LIGNO  4384   IDRRR[ 1]--SVGGGD[1]AER----LVPMQQAELLVDMATDPALLARTYGGWQPDL  4430
Query_CESTO  4133   VRSSV[ 5]--NLKGAL[3]-----------EQTRRLVCLSTCPELLARMHRGWKAAM  4178
Query_OVI2_  4626   LVGPQ    --NVEVNS[3]-------LTPEEQARQLTRLATSADLLARAHIGWHSFL  4670
Query_PRIAP  4116   VLGDP    RSDLRAQL   PPE---GLTVEQQVQCIIEQATEPSLLGRTWVGWEPWM  4163
Query_ROTIF  4011   LMGNK[11]RSKILAEN   GEKyr--LTIEEQVKCIIDQSTDTNILGRTWIGWQPFC  4070
Query_SPONG  4110   LLGDS    STNLRARV   GKRcd---SVEDQVDCLIDHATDSNVLGRTYGGWEAWV  4157
Query_TARDI  4069   CIGDA    -DGVRAKH[1]--Drh--LEVEAQIECLIDQSTDENILGRLFSGWNPWV  4115
Query_TRICH  4080   ALGDK    QYNERARM   PAD---NLSVEDQVTCLIDQATDPNILGRTWQGWEPWM  4127
Query_OSTRE  3894   LVGSN    --------[2]PTSkc--ATVEEQVRALIDLATNPELLAVAWSGWRPWL  3936
Query_NITEN  4272   VRGDP    RVNARAKV[1]-GKic--ATVQQQVECLIDQATDPNLVGRMWIGWAPWV  4320
Query_CLUBM  3866   VLGDP    VKNTRRRI[1]-RKic--ASVEEQVDCLVDQATDPNLLGRLWVGWQSWM  3914
Query_MOSS_  4050   VRGDP    SRNIRARV[1]-GNvc--VSVQEQIDCLLDQATDPNLLGRIFAGWQPWV  4098
Query_LIVER  4104   VRGSP    DRNMRART[1]-GNic--QSIREQVDCLIDQATDPNLLGRIWQGWQPWV  4152
```