

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

**Supplementary Figure 5: COBALT alignment of representative DNA-PKcs sequences from all major phyla/short list.**  
 Conservation setting 2 bits. See Supplementary Table 1 for more information on sequences and organisms.

Sequence number	Sequence name	Descriptor	Scientific name
Icl Query_10001	CAPITELLA	Annelid worm	<i>Capitella teleta</i>
Icl Query_10002	STARFISH	Crown of thorns starfish	<i>Acanthaster planci</i>
Icl Query_10003	LANCELET	Belcher's lancelet	<i>Branchiostoma belcheri</i>
Icl Query_10004	HUMAN		<i>Homo sapiens</i>
Icl Query_10005	ESHARK	Elephant shark	<i>Callorhinchus milii</i>
Icl Query_10006	LIMPET	Owl limpet	<i>Lottia gigantea</i>
Icl Query_10007	LINGULA	Brachiopod	<i>Lingula anatina</i>
Icl Query_10008	TRICHINELLA	Parasitic roundworm	<i>Trichinella pseudospiralis</i>
Icl Query_10009	AMOEBA	Slime mould	<i>Heterostelium album</i>
Icl Query_10010	GLOMUS	Fungus	<i>Glomus cerebriforme</i>
Icl Query_10011	MOSS	Spreading earthmoss	<i>Physcomitrella patens</i>
Icl Query_10012	TERMITE		<i>Cryptotermes secundus</i>
Icl Query_10013	SALPINGOECA	Choanoflagellate	<i>Salpingoeca rosetta</i>
Icl Query_10014	ROTIFER		<i>Brachionus plicatilis</i>
Icl Query_10015	SHRIMP		<i>Penaeus vannamei</i>
Icl Query_10016	SCORPION		<i>Centruroides sculpturatus</i>
Icl Query_10017	SPONGE		<i>Amphimedoa queenslandia</i>
Icl Query_10018	TRICHOPLAX		<i>Placazoa</i>
Icl Query_10019	STONYCORAL		<i>Stylophora pistillata</i>
Icl Query_10020	OOMYCETE	Plant mildew	<i>Plasmopara halstedii</i>
Icl Query_10021	FLATWORM		<i>Macrostomum lignano</i>
Icl Query_10022	STYLONCHIA	Ciliate	<i>Stylonychia lemnae</i>

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	1	MA---RA	LQEQLAGLVSDTR---GQPG-vVNARISHES--	-IGNIGAIc-LNDV-NASDvDLAY-SVIFN	57
Query_10002	1	MA--tAD	LPVKLSENLGLIS---TKSG-lHVAEKAQD---	LVVdLGHLC-LRQL-SENELAYCC-SQLFN	58
Query_10003	1	MAG-sGS	LEARLSELHALLA---ESDH-gHGAveAHN---	LVADIAQIC-LRHt-AERDvAYCC-SVLFQ	59
Query_10004	1	MA---G-	--SGAGVRCsLLR---LQET--LSAADRCGAAL[	4]LIRGLQbEC-VLSS-SPAVLALQT-SLVFS	60
Query_10005	1	MA---GP	RGAAADTLsAFSK---LHHL--LSLSDVKsAAA[	4]AVGSISRDFrCSSS-SDSQLALQS-SMVFS	64
Query_10006	1	MA--sGL	LEESLRDLHANLK---D-----GGQSELDQID[	1]IVPTLSQIC-LHEI-TEKDidYCS-SVLFd	57
Query_10007	1	MA---GN	LQDLLCTLHsCVG---GEEEdaRGGRKTFD---	LVRDIANLC-LKET-SDRETDYCS-SLLFN	58
Query_10008	1	MAY---S	--STVDRIRDALN---SLKD-fIEGADSDQAE[	1]-VQQIFNHfKLNg-SERNFGFCfnSLLS	59
Query_10009	1	[4]MNN-vds	[4]IDGYLVKLHQsvGgttSTStAtTATAAITsAS[	15]LIEQLSNVv-LKEL-AEDEIGVAA-SLLF	89
Query_10010	1	MSG-LDT	LQKCLQELANCRQniv-QKS---SPGEAVGEAL[	2]-SQDIDHIV-SNDI-EDQELGLAT-QLLFD	63
Query_10011	1	MDK----	[1]LGSCLASLRDAADsvr-----SSGGFVDLAP[	1]-VVDfSELC--KDLySDEIdSYAI-AIVFD	57
Query_10012	1	MMG--s-	---ELIGFVvVIN---SCIS---GNHLNKYIEAK[	1]NISNIISYfEdKRRkrTDL--DGAF-TVLFd	57
Query_10013	1	MAA----	ARSELERLcRNLE---EQRl---SASPRRFIAA[	2]IVADIAAYT--RTLTPPELALAA-SILG	59
Query_10014	1	MND----	LEKHlQNLFDsLS---NSNV-----DAAKSI-	-VLDIVYLcGEN-L-NPOTITFCL-SLFFK	52
Query_10015	1	MAG--GG	REEE DLLLGLKlQ---SFGP--GNGQGDpDAAK[	1]VIVDIAEELaVDL--TPAQYDLXI-SQVLG	61
Query_10016	1	MNEitNS	LHENILKLNDFIK---YQKD-----EGHSSiY	VIKDISEIClRGTl-SEKDRSYVC-SLILR	59
Query_10017	1	ME---SR	IESLWESLHSSp---QRYS-----YATEVAS	---DLYSLC-VEQI-TESELAYSS-SVLFN	52
Query_10018	1	MA---SN	LSQLLIDLHALLN---LQSL--HRGSNACDIAS	---DITVLC-LQTA-SDAEKAYCS-SALFG	56
Query_10019	1	MS---DL	LNRYLTHLHSLPS---LDAP--HCGLDASNTVV	---DLGSPC-LENV-SSEEQAYYS-SLLFS	56
Query_10020	1	MTT----	[1]LNERLEELTAFLLh-----AQAVESIDETL[	2]LRLEVSEAMiLSRA-SAQQCtILL--FOsA	58
Query_10021	1	MSS--DD	[4]-----IDKFLKdp-----ELKESKA[	1]----FNVI fNEFASsNENFKHKYlAlLFD	51
Query_10022	1	MEQ----	[1]IDELLRQLHKFLF-----A[	4]VVNELRIYlIDESL-KEADLAFIS-SKIFN	50
Query_10001	58	HETGLIKFLKkIIQND	EF---EG-C--KCEILkELLKTFISKf-PGKVH--	PYAVDVKDAcLQL-ITRRRRARVKNNATPL	127
Query_10002	59	KDNGILVFLNKAIpKDEL	---KG-S--KEQLLEFLHAPLLKA-GSRLL--	ARTVEIKKICISA-FMRDKSAKVKTKTPPV	128
Query_10003	60	ESTGITAFLRKtVTLdQY	---LP-A--KVTElSFLLAFLKlEIK-GRKIQ-	PHAVEKKEVCMAV-FSRDRlSRVKCATFPV	129
Query_10004	61	RDFGLLVFVRKSLNSIEF	---RE-C--REElIKFLCIFILEKMG-QQKIA--	PYSVEIKNTCTSV-YTKDRAAKCKIPALDL	130
Query_10005	65	KDTGLLVFLRKSrLrFEF	---RD-S--REDVLLKLLSEVVERI-GEAVG-	RYACDIKGTcVvV-YTKKAAKCKAAALDL	134
Query_10006	58	KEIGVTFFLQKlSKKNEY	---QG-SnaKGLLELLSDFIHKV-GKKAL--	PYLVEIKEASLNS-ymTDRFTKlKSSALPV	129
Query_10007	59	KDTGLVSVFKVVKQAEY	---DS-T--KTAELTLCLFMERA-EKKVL--	PYAVDIKDVcITL-FIREEKAAVKKVVIVL	128
Query_10008	60	DSNSLLSILSSIGKTAF	---AQ-Ti---VYGIIDIRLAIAKT-TSLSP--	SHRKAIMRFtLRL-LKQDYlLKGRRREVLYL	129
Query_10009	90	GDHNLKFLERSVLINTK	----DmIkIKTSILNLI SKLIKSF---KLQvady	VVPiKSICISL-FRRDQSPTVQAATFEP	161
Query_10010	64	PQIGIKFPLNDTSrVTDsv	---lAkaKIALLEFLFLYIRKI-QSDIE--	PYTEIiKKNVFGI-LNVdDEAFVKAASFp	135
Query_10011	58	EDTGIFLESLLRVRDK	----SvVraEScSFLALPIKIKI-GPRALh--	YAALLKEKSLTF-FRREDSKSVQAAALTP	129
Query_10012	58	KKTGILQFIQKALQDAP	---TE-A--IKESFTLLKQLIDVF-KERMS--	CHVVQIKDLCKTC-ILLEKKAEBKDKAFDV	127
Query_10013	60	SDQGVLSFLRQ-lGLELY	---EE-C--KVAALEFLAGFIQRlpADRLT-	PYAVRIKDTCLTC-F-RDKRAKVKSLSPK	128
Query_10014	53	REKNLLTFLRKTITRDEF	---RG-C--KVDLLQFLKFKITSa-EKKIL--	PYSPDIKDTcILL-FSSDKYSDVRCATFSI	122
Query_10015	62	EK-GITfFRGSVSKDQSF	---VE-A--KkSTLDVLTLLDKB-AQKIN--	KVVDIRKYTLYL-YNVEKSTKvQAAALeV	130
Query_10016	60	KEDGFLHFLEEAVNTFSL	---KD-S--LIEALQFLNLFLINI-GNKVV--	QYGI TVKDVVFSI-FLRNQLAKVKILALSA	129
Query_10017	53	PDNGILKFLHDVTTLEEF	---AS-C--KEBELKLLSNVFNKL-SRRVL--	SHIVPITDVcVFI-FMREKSAKVKNATLNL	122
Query_10018	57	QDYGVVTFLRKSlGLDDL	---DK-C--KEDLLKLIATYMEKI-GGKIS--	PYAVQIKDVcISL-FTRDKAARVKNATFPF	126
Query_10019	57	GETSIVSFLRKVITLDEF	---QD-A--KIEELLKFLQSVYMKM-GKKIN--	PYVVEIKDICLKI-FSQDHSNKAKAETLPL	126
Query_10020	59	EPPSLRFLRSTSAETFD	garkReiSaaRIGVPELLSAPLKVYgAHSALsk	QHVIdIYKVCQQT-ARSDPFNRvKAQALNV	137
Query_10021	52	EPEGLVYRlQSKKEVSEDA	-----KDGDLFTTKLIQASpALLATs--	LPKVKCLAESLlFSSISSAKVRIASLNC	119
Query_10022	51	SNSLLEFLRENSLHNDK	l---iSsvKkVYELIIdYlKLR-SQEAT--	PYLKLLKDTCLRS-FRSDHNSLVKESsLQL	122
Query_10001	128	LIEL	LKQTSGL-PVAGDLQVEGMIERLFKELIK-SPKP-AA-	SVKAHIFHLGVIaELFPPEKMv--RGSdKLL--NL	196
Query_10002	129	LIEL	LQRSASA-QMEADLIPKMIELyFQQLMQ-PSKL-SQ-	TVKQGIYCLLGVLaEiYpVLM--sGYSERLL--SV	197
Query_10003	130	LKKV	LQLTIHS-QLGDELrVSDMVDRLFLELTM-RSQT-TA-	TVKSGALSLLGVISeVYpDLMVqkDHPDRLL--Kv	200
Query_10004	131	LIKL	LQTFRRSS-RLMDEFKIGELFSKFYGLAL-KKKI-PD-	TVLEKVYELGLLGEVHPSEMI--NNAENLF--RA	203
Query_10005	135	LITL	LHsKLSs-SCLQFKIGELFNKFGDLSQ-KSKF-PD-	TVLEKVYELGLVLaEVQPTEmL--HNSEKLF--CA	209
Query_10006	130	LIKV	LELSVGS-NMGEDLKIQKfIEKPFMELTK-ASKL-TA-	TVKGIYMLGVIcEiYpEYMT--TDSTKLL--DI	198
Query_10007	129	LILK	LELTegS-QMSEELNMRMElEFLMELSR-KTKH-QA-	SVKRELYKLGIFAEVYpEGMv--NYAERLT--DL	197
Query_10008	130	ANEC[3]	-SSVLStESELNfCYDELLHCLLALIDS-LKKL-QN-	GVKGLFIIIGSLCQEIPEKLIaDYNVALLfAnI	205
Query_10009	162	IQKI	LHMkLTdNINPEVfGVREMSDlLLQFTc--GKL-TQ-	TVKAELLETVGLFTeYFTGIMF--DKSAQLIs--I	230
Query_10010	136	MMAM[3]	----SKRiIDPEKLNVRDIISRYLIEFFV-RTKGT	SgVKAkAGALDMLGIaHRFPeYVdf--EKNKILr--V	206
Query_10011	130	LIVL	LELNMAR-LKPAMFAVDQELSTLLMLeFR-KAKT-NS-	TLKANIiQLLGDLIeYFRSfPE-eENVLAILk--L	199
Query_10012	128	IQSV	LEN-GPW-ND--NLNVDELFDKlINPFKSnKSN	TmPA-KVVQKHLHlGIIVHkYpECLR--PEETRLl-Rl	196
Query_10013	129	LIAL[5]	lFTDENrgQLLEKADDLKkYFDPFKSsASKL-TQ-	TVKRDMLHaIGLIISLSPAFAA-sSLSTSVL--NV	205
Query_10014	123	ISKI	LELTkGEyEITAKMNIpKLADYfLSLAN-QSKL-SS-	SIKANILSILGRICWYHPeVIS--SKSNKCL--EL	192
Query_10015	131	FCTL	LAKCGSH-IDIQEINVDQFIERLICDIRSakGST---	VLQHQLTGLGMLMSYpQNLR--HQSPILl--KI	197
Query_10016	130	LIKV	LELCsGaiS-SEQFNlQKiVnKIIEYLaQsSki-TP-	TVKQAYYLLGISSRHYPdCMQ--LCSERLL--NW	199
Query_10017	123	LINT	LELCGKT-EYSHDLIPKIADKFFQaALL-PTKH-SA-	TRGGYHLLGVLCEWfPPEHLV--KISSRLI--DI	191
Query_10018	127	LLQL	LSSKLSHS-SVVELAISKLVdKYMCAcKQ-PTKY-SS-	TVKSGMYSLLGTIAEKfPEYIM--PHAQQLL--TI	195
Query_10019	127	MAEI	LDVKLES-SVVEKLDVpKIFDKYfGACIQ-pVKH-SS-	TVKYWfSLlGTIAKfVPEcVI--SHSDHLV--KH	195
Query_10020	138	VVNA[3]	----AHKhTTSDEIEPRAYVEKfYLDIQY--SKA-AQ-	TVKGMLKVIGYLVKfSaelK--KTVSPllG--W	205
Query_10021	120	LSIC[3]	----AVGSsVDGNQNLVNGQLCERLrLEVQg--RQ-	TSeGVTQGLYRALGSICKHYpEeFQ-----PVAv-KI	185
Query_10022	123	LIKI[4]	----NSKEIEEMRPNDMMVMLLDEIKL--RKP-SA-	TVKGTITWlIGLLHEKYEQLNdfkVESQDVGf--	191

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	197	YMTTLK-	QETI-SKTHEPEYVLIAGCLOGLTHFMVNFTQSADDEGAEHAY---SIVKYCVKTM-DQNIK--KRYEMP	265
Query_10002	198	YIATLN-	QQMT-SKTRKPELTVIAGCCLTGLGHLVNFNFTQSVEEGAKHSK---DIYRFARMAI-DPTVD--HTRYELN	266
Query_10003	201	YISMLK-	EQME-SRTRKPEMLVIGCCLGELTHFLVNFNFTQSVEEGSAHSK---DIYNFTRRAV-SPQVE--MARYDVP	269
Query_10004	200	FLGELK-	TQMT-SAVREPKLPVLGACCLKGLSSLLCNFTKSMEDPQTSR---EIFNFVLKAI-RPQID--LKRYPVP	268
Query_10005	204	YLNELR-	VQMT-SATREPKFLVVGCLRGLTALMNFNFTKSMEDPKASK---DIFDYALKAI-NPQID--LKRYPVP	272
Query_10006	199	YKSSLK-	SQMM-SKTNKPEMTIISGCLLELTSYLVNFNFTQSAGDD-S-CY---YIVKYCKMAL-KQSTT--LTRYEVP	265
Query_10007	198	CVRTLQ-	AEMN-KKTGKPEIPVMTGCELEALSSILNFTQSAAEG-KHSY---DIFKTRMAL-NLEET---ARYDLT	264
Query_10008	206	MSEELKN	-QMC-SSISHSYSVIIGCFLGLNLLAHFNFLNFDGACRMSCS--ELFNLLKKAC-NPKND--VHRLGHL	275
Query_10009	231	FMETLGS	-QLK---SKTPETFILKSSLLGLSSLLNFTSSDFISGSSKNVql--LYQYLYICL-DPVSS--QRFDIP	298
Query_10010	207	FLANLKE[1]	-SLT-----DAAFGLYSFLFAYSDSIKKENQDAY---IFKCMNSTtMVEND--RTYSIV	264
Query_10011	200	CLDTLEE	-QLQ---KSPQLVLMAGAIKCLDSLMMWFNDSLPAAGGSVnsr-RVFTCIEQVL-WPREE--LKRYPVL	267
Query_10012	197	YLSELEM	-QTS--HATAVKYQIIAGCLNGLTEFLYVFP--LKPQEKEAR---RIYDCVTAVA-RKHKE--M--RVAP	260
Query_10013	206	YLSELRK	-QMT--GTKAPELPTVIGCFRGLTAYLKNPVSQCDPSARQATpgvQIYNYVKMAI-NPDVTVkmTRFQVP	278
Query_10014	193	FLTTLKN	-EIN-AKFRKSDLSIAGALDGLNSYLNNFVSSETSDHLRA----VFNFKLSSLSSSED--LSRYAVP	260
Query_10015	198	FKGQLVK	-HTG---SKPDLNALAGVLRGLALYLHFFPEGIDDDPQLFT---ELYKHIIKVL-NPDLN--LPRRDAQ	263
Query_10016	200	YLKDLKF	-HTT-SN-QKSESSIIGCLOGLNFMYPQLQDTEDEKAY---QLKFIIRIIL--SKRD--VKTYDMA	266
Query_10017	192	YMSNLK-	SEMTvKTKKPELPISGCLYGLTAYLNNFQSASEGSVYAQ---DIYKYTRMAI-DPLAN--LTRYEVP	261
Query_10018	196	YLNTLK-	SQMT-GKNTKLDMPVYIGCLRGLTSLYLNFTQSVEEESKYAA---DIYKYARLAI-DPKVE--LPRYDLP	264
Query_10019	196	YVSVLKL[5]	MQMT-RKTKKPDMLIIAGCLRGLTSLYLNFTQSVEGSKFAK---DIYTYGRMAI-NPQVS--LSRYEVP	270
Query_10020	206	IEHELEK	-QFL---STSPMMIVNGLLFALARLLKCDAGRKQNDALRK---KVYSYLLTIF-ATTNgnLSRYQVT	274
Query_10021	186	SVNVVKS[2]	QKIK-SDHTKVFQAVIASLFEFLSNYFCHFSL---EFPDLKQ---VYHLAVKVI--AGKN--YQRYEAP	252
Query_10022	192	--QELVA	-QMK---SEKPEIKAIIGMLKGFTHSLVG--ECTLDQDQLDA---LYVCLKTAI-QPILD--MQHKGM	253
Query_10001	266	KASLGLMAKHAQMSDYLDDDFKDLYETLSYWTNHNREMAHIGMLALEDFLVQISSALVTKV-KEGN-rkEGAVFKFFI	343	
Query_10002	267	KAGLNMFKAHSAQFGPYLSDHQDLHEAFNKWSAHNNAVAVKAVGFAAQEAFLLQOIAELIVERA-AKRD-asDMAVFRPFV	344	
Query_10003	270	KAALWLLAKHAAQFNEYLNNDYEDMYRKLKQAGYAALESFLKQISDMLVARH-EKENnkaDLAVFKFFI	348	
Query_10004	269	SAGRLRFALHASQFSTCLLDNYVSLFEVLLKWCATNVELKKAALSALSFLKQVSN-MVAKN-AEMH---KNKLYYFM	342	
Query_10005	273	LAGRLRFTLHSSQFGTCLMDNYRTVFEKMKKSCGHTNSEMKKASYSALSFLKQVAI-QVAEN-AHLH---KDKLEYIM	346	
Query_10006	266	KAGLHLLSKHASQFNEFLIDDYEDMYDRLYFWSHHNRDILHVGCMACESFLKEVNMMSYLSFG-D-----SVFFFI	335	
Query_10007	265	RAALTLLKRHASQFQFIYENHERLYSVLLAWSQHNNDVSSLGMEALEAFFKQMAEMLATKA-EBGH--pDRAAFAPFI	341	
Query_10008	276	KAALSFLISRQVFRHHFLDEYSLLNECFSHWAKHYDREIKTYSLNAEDLLIRELATELITKVDSDPN--hVRSVYKYFC	353	
Query_10009	299	RAAMVVQRHTSLFKQYLTEQSEKFFARLEEWCKHINKNRDLAFNTIDSFLQOIANELTVGNRSIES---DQATFKYFI	375	
Query_10010	265	KTGLQYIIDHSNIFGEPFLREYKEIWOQLWKLVRHNPVRNAAFKAIWAFKQIAKVLNSATEFKEF-----DCVDYFM	339	
Query_10011	268	KASIRLRSRHAGLFRMRLESIGIRNMHWLQESGIRNTHSNKAKIRDAALPAYDQFIHVSALVSDSRIRNA----QALYTAIV	343	
Query_10012	261	RCALGLIQAHHGGQLAKFLYDEYKWFHEFLKWLINSAGED-HKAGVSAVDVYVQVSAIILKNDT-SET---GHHVFLYFI	334	
Query_10013	279	RAALDFALHADQFRSMIATDFELLQYLSDSCSEHNNKDIKASAFRALEAFLLQOSVAMHTKAGKSKDT---MFMGFLIF	354	
Query_10014	261	KSALNLIKHCROFDELIYSYKEIFDRISQWSEHKNYDMKKLALYALDLSYKSLAEMIRVKSNTETD-krRL-IFKYFI	338	
Query_10015	264	RAALELIHTHGAKFSSQLFDECTQVQWFWKCGSRNRDDHKAALPAMDTFIIVIANMLELQA-MEDD-krGVKIFKYLL	341	
Query_10016	267	KASITLSSSHASLLAQFLYEDSMNIYKIEIWTNHSNREVRSLGYNALDAVFKQISDML--EQNYE-D-scKAIEFKFFI	342	
Query_10017	262	RAGLSLISRHAQFREYLVKDESHMYESLSSWSNHNKKECRVAVNAIGSFFREVSTQISDPSISREL---GISNYKFFM	338	
Query_10018	265	KAGLNLIKAHAAQFCDYISREAEALYEVIYRWCRRNRDIRAAAFHAEIEMKQIAEYLVNEN-DAAN---INIIFKFFL	339	
Query_10019	271	KAGLYLFAKHAQFKEYIAKDYEVMYGDLYRWCKHNNKETSAYAAAAMEAFFQOISLYVGRHGDKSD---LAMFMYFI	346	
Query_10020	275	NSAEHFLEKHAQVFHEETGNPLSAWFSYMKSCCLSDNKTVKQHFACSNISLQVLNSYLAEAKDDMRK-----KCLNKVL	349	
Query_10021	253	QOAILLFAHCHPVSAFLADNYEVLFEQFRWSTHNRDLMHVGMALAFDAFVKSVTAATSAAESAtpaQKKLFLPFL	332	
Query_10022	254	KSAMKLLGTHINLFKXHVTRHAIELVALTKLVCVENMETRDAANEMLVHIMQETISDGIITDSSIHDD-----IFRKIM	327	
Query_10001	344	QKFRDIM	DNSSGSKIEAIA-IRKGYFAAPCKLYLTPKDVLFMFNEMLHRSEQLFLSGAEQND- D RIG 411	
Query_10002	345	KRFRDII	DNPQSDTRALSSA-VRGYGYFAAPCKLFMTDDVVKFMFTMIQRSEQLYLYQGEVSD- D KLNQ 412	
Query_10003	349	REFRTHI	DTTEAGTRMSIA-VRGYGYFAAPCKFLSESDVKFMFSEMMQRSEQLYFGTADSM- D KVHH 416	
Query_10004	343	EQFYGII	RNVDNNKELSLIA-IRGYGLFAGPCKV-INAKDVDFMYVELIQRCKMFLTQTDG- D RVYQ 409	
Query_10005	347	QQFYDII	RTPESSMKELSLIA-IRGYGLFAAPCKA-VNSNDVDFMYIELIQRCKQMYLTEADTED- D NVYQ 413	
Query_10006	336	KQFRQIM	NSESSDKVSLA-IRGYGLLAGPCRRFLNVTDVVFMLESEMNLKCEQFLLKVGENQ- D KMYG 403	
Query_10007	342	RQFRETL	DNPHAKQNEIASA-VKGYGFAAPCKFHMTAEVVKYMFNEMQRSEQLFSSKNELV- D NVYN 409	
Query_10008	354	NQFDEIF	SNT-SGIREISMT-SKSYGRFLQIGKALYGEDEVCKVLRKLLQRFEYLSSTSSDAQQ- - NVYH 420	
Query_10009	376	RKFSYIF	ENNNSTPFDISIA-IRGCGRFAAPVKAFIGEWELKSLGLNSLFKFSKMLVMVLENIDE[1]- -VMH 443	
Query_10010	340	LIFDDIA[13]	BEQDQKTKDISIA-VRGYGYFAACVSKRLDKBEERYKPIFQNLNPSNWAISNLNEIET -[1]ALSQ 421	
Query_10011	344	KDALDVL	NAQDADPKRISLA-IRALGKLAACIVKFSGKESLKDFAHRL-----EPPGSLSRGDDG[1]E[1]SLGY 409	
Query_10012	335	NFFKNKM	LSSSTSSYDMKLV-IKAFGLFAAPCRVYMTQDVLVYMTVVMQRAEQIYFKEGDISNA - DLQN 402	
Query_10013	355	KRFAKAM	GEA-ASHRAVAIA-IRGYGYFAAPCKFLCGVDDLRLMSDDLARRCQQLYFHMGGDDEQ E SAIH 422	
Query_10014	339	QKFYKNL	TDGNKDLKEMVIS-IRKGYGAFAGPCRELDMDTKMLMLNIIIDICDRNFLSADSQAQH[4]E KIFQ 411	
Query_10015	342	QBY-NKM	LETSESTKLSLA-VKGYGLFSGCRQFLSPVEVHTMFSQVLTASQHAFYQSGDMLN - KLTN 408	
Query_10016	343	EKF-QLM	IKCSENIIRVSTA-IRGFYFAKVCCKLYLTEKDVKYMLNDVIKKEQHCLNEE----[4]E RIQI 409	
Query_10017	339	KCFRSLM	NDPDIHRLSLIA-VKGYGHFAKPCRLYGSMDLHFMFGEMVQRSEQLFLASVDFTT[6]E ERFL 413	
Query_10018	340	KKFRSII	DSGSSDAKELAIA-IRGYGYFAAPCKFLKDKDVRFMFTDLLQRSEQLFPI--DVNV- D KIYQ 405	
Query_10019	347	KTFRRII	DSVLSDRQMAIA-VRGYGYFSPCKFLMQADVRFMFSEIMQRSEQFLYSQSDLSE- D RAQH 414	
Query_10020	350	KEVLPVI	SDFGASTLAITFA-LQCLGWFASSIFLYLGAQYIKIEDKLRVYGESLLALDAKTTAW[6]Q YVQC 424	
Query_10021	333	QQFRQIL	DSDGSSWKTVSVA-VRGFYFAEAAKLLAPPDMRQVYADVQRAEDLFTAESDSSDE - RLSN 400	
Query_10022	328	KEFETIL	NNEFNQTNVLLISIKAVGIFSKAIKTKSGDDALREYLSKLTTLSONKIIFEFQEQED[6]N[6]KQKQ 409	

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	412	LHSFLGALASIVEELD-QFSETFPLSFLERLVLVQMQLVPRLLHVKLHNQCYIALIRLFLGLALKGSIF-KGFLERIIYQSL	489			
Query_10002	413	LPNLFLEALASIMRQLD-VLSDSFQAALERLIVVQVSSYPVPNKYQYLSHQAILLVLLIALSSKGATL-KNFLSRVVYQAL	490			
Query_10003	417	LPSFLEALASIVQHLD-QLSDMFVTSLERLIVVLFENPFRLLGKMKHLICQKAILLVKLLALVPKQATL-RTFLLSRIVYQAL	494			
Query_10004	410	MPSFLQASVSVLLYLD-TVPEVYTPVLEHLVVMQIDSPFQYSPKMLQVCCRAIVKVFLLAAKGPVL-RNCLISTVVHQGL	487			
Query_10005	414	LPSFLDSIASVIFHLD-KIPGVYTPVLERLVLVQIDSPFLYSARMQPICCRSIVKVFVLAGKGPVL-WNFISTVVHQGL	491			
Query_10006	404	LHSYLDIAIANIIEID-ELSEYASIVEHLLVMQIQNMPMLNEKMHFMATKSVIHVLLALMCKGSVF-NQVLSGVVYQV	481			
Query_10007	410	LPSFIESLARIVEQME-EVSDTFVLSLEKLVVCLFHEHFPSPYKRRKLDVQSLIGTFHALISKGPTEL-RIFLGRVVYQGL	487			
Query_10008	421	ITNLEAASRGLLCLLS-KIADDDVQLVERHVTFLFHNPLMEKRFYHRCKVINVATFAAYGHNYQY----IQSVFQSL	495			
Query_10009	444	LSSFINAFSNILFEFK-DIEYWMHELEQAVATYIIFPYLFTVQRERYYSAVNRLSSLYHRPEFL-KTFLGRIVRQGL	521			
Query_10010	422	IPLYIEAYCFIGREYD-EIFQQLMETEHEVGLMIMYFPRMVDLRLRCVNAFEHLLWMLYLKGRGTMEKFFVSKSFYQTL	500			
Query_10011	410	AVTLVGAFADMVHLD-DVEDLMVTFIAQVTGSIFFLYPQLFAKQKQAVYALRNLLKALYPKGPVY-FLFLRNIVWSSL	487			
Query_10012	403	LPDYIESLSSIMRQLS-AVTNEQLVALQCLTVQLIENFPKLPKPFHWAIRAITVTLSSLAESGGTFLEEFLLSNVVYQGI	481			
Query_10013	423	LASYIDAARLIQQLE-ESGDVFLPILERLIVLLLEAYSPMYRQRASVHSLTRCFVSLYHKGSTL-QTFPAKVANHFV	500			
Query_10014	412	LSSFIESLSCVCSQID-SFSEGLNTMEKLVILADDCYPKLVKRYNNQISLAIAALFLAIQRTNFSDE-FISRIIYHSL	488			
Query_10015	409	LPSYIESLAIISNMDD-VVNSGILENVQQLSSALVEHPFLVSSFKRFLAYRALFRLLFTVYNKPGYF-HEFLHTFVYKAI	486			
Query_10016	410	LPSFLDLSLSSIVLQIQ-EISESSLNLEKLVFLQIQNFPYLPSPYHFLCYNLSLIKFLAVHSGKQPQF-KGFLNKIIYQGI	487			
Query_10017	414	LPTFIEALANISKEIQ-NLPDVLSSLQKLAVLLFHCYFPLSSHVRVQGSVLLKLMVAVSSIPTSF-KQFLSHIVYQGV	491			
Query_10018	406	LPSFLEAFASILNILD-EVNDTFTFRSLEKLTIVLLENFPYVINELHSLCYKSLFELLALHSGKAAL-KNFLSRVVYQGL	483			
Query_10019	415	LPTFLEALSIIIEELD-KISDSVLACMERIVRFLDKFALNLKGGKFLCCKALIRVLFHLSKSGAIL-RNFLSEIVYQGL	492			
Query_10020	425	IGQFVQHRSDVV--LD---EGYIMFLERILCHLMTAYPQCLLKSQSVVHKSIALFVLSKWTVID--PVDQFVHHQL	496			
Query_10021	401	LPSFVESLNLALHLD-ELTGVFLSLLARLSVTVCHMYRVSERMGRVGLRAVAHALHRLSRRRPGLAKAFAAQVVYQSL	479			
Query_10022	410	LLSFIESYSFIQNLDSBPSDEVQHFLLKVCIEGRHRRKLYQKYRDRFYEALASFTKALKSHQKAF-ANWIQKFDVLSL	488			
Query_10001	490	IRSCSH---PIVLESQEAENSEN---[3]VNLKVD-----VSYQDYLPMWKA-LLT	AGKLLKAL-GQYNTAEQHR	550		
Query_10002	491	VQTCSH---PVVLDGSPSIDH--pV[5]QDTEADLLADARHITYKDYVKLWAS-LLD	STHVKEV-QAMGMVADR	561		
Query_10003	495	IRTCSH---PNVIEGVTAEDEDLp[5]GGEQQPQPPVKRFSYRDYLELWGN-LLD	SSRLKEL-SQWGAVADRQ	567		
Query_10004	488	IRICSK---PVVLPKGPESSESE--D[1]RASEGVRTGKWKVPTYKDYVDLFRH-LLS	SDQMMDS-ILADEAFFSVN	554		
Query_10005	492	IRVCSK---PVLPGEGTGRSE--S[5]SASEVMWTGKWRVPSYKDYLDLFRS-LLN	CGKMKEF-GLVD-TLQMIN	561		
Query_10006	482	IRTCSH---PIVLETEQTD--T[5]VEKQ-----S-ISYKSYLNINWL-LMD	SPEIKVV-LFESFALFILT	543		
Query_10007	488	IRTCSH---PVATDLGEVBAE---V[5]MESE-----KKITYKDHLDFWNM-LLG	SSHLKV---HSKLPDPEKS	548		
Query_10008	496	IRVIAS---LPSVKGESIVLAK--	DENTFHLDVLPFKM-LLD	S-EFVLD-SEFGDRAMQV	550	
Query_10009	522	LITFSKp-heALLPHLNTGDTpw--	-----YEVYKEVWYH-LLH[19]-QQVKQeGeKEKEKEKDS	593		
Query_10010	501	ILTCSD-----ATQLQNRSDLQn--	---ANSKDMIDVVAEHTYQEMFFFWEN-IFK[16]ENRLNTI-----QLE	571		
Query_10011	488	QRAMEP-----VDEALEEGSVq--	-----LWRHYVELWSH-LID[15]AESFEEL-----VTSVGVNGV	549		
Query_10012	482	IRSCKH---QIAIDVELQKGO---	-----FHENVFTYKDYIPLWKE-LL-	--SVNNE-QKRSTSLTQTQ	534	
Query_10013	501	LSACSQ---SIRTEMDAEDAva--	-----QSYLALLRQ-LLK	GSLVKI-PDEEMTHEDRN	550	
Query_10014	489	IRIFSY---KTLYSVN-QEDTS---	---DETNEPMSSIIVTTSLDYVMLWSN-FLG	LQEFSDL-K--SDNLDDK	549	
Query_10015	487	VLSCSH---APEVESSPGEQgs[5]	GGGEWKLKNEVVTFSSYIPLWTA-LMV	PEKVAANI-KVSGMPATVMT	557	
Query_10016	488	IRTCY---PVRFEIAREENM---	---SGDNLELKNPREITSKSYSSLWEN-LLY	IVRNKEL-NHIGISLEDRH	551	
Query_10017	492	IRSCSL---PVLIEAVQGDSDAv--	VTAAPLADSNKVMYSLDYVQLWIQ-LID	YKLLKGM-GM-DFDLELR	556	
Query_10018	484	IRTCSH---PVYV---ETDSDLN--I	QSDSVILE-TGPRISYKDYISLWKN-LIQ	ATEIKGL-TLITTDREHI	546	
Query_10019	493	IRTCSH---PVIL-KEVNGEST--I	MEEREVEYEGSTPNTSLKDYLELWEY-LLD	KDKLKE-SLLGKQLEERK	557	
Query_10020	497	LFSIADi-----TDPDqaV[5]	TGELVTNLIYIEYEGFVNLNLRSErVVG[9]SGDVRQI	-----ELK	563	
Query_10021	480	VRTCSH-ehPVYQSDTEAGSSPgsA[5]	DFVQGDASNVSVRKISYKDYLRLLWLC-LFS	LEKEKLLcDYVGMASAAER	555	
Query_10022	489	IQTCKIip-dAVLFGGESPYESL---	-----QDAVQFWRE-WLN	----KDK-----	524	
Query_10001	551	---V LASAIYDELLSAMLKMSRLDLS---	---NSTDS--ESQP-STSEP	ALS---D--PSSDPLR	G	601
Query_10002	562	---K IHII LYDELIFSILKILSKLDLTS---	NKDTTE--LQDQ-VAVDS	DS--TEDA--S-SDPLH	G	614
Query_10003	568	---A MHRMLYDELVTSVLKVIDKLDLTT---	SKQVET--GNEE-LTADV	SG--SQDG--TTSDFPA	G	621
Query_10004	555	---S[4]LNHLLYDEFVKSVLKIVEKLDLLEIQ	TVGEQ-----E-NGDEA	PGVMM---iPTSDFPA	N	611
Query_10005	562	---L[4]LNRLLYDEFIKSILKIEKLDLSVEKQ	NQREE-----E-GSGDV	SNGNV---iPTSDFPA	D	618
Query_10006	544	---Q LTDALYNELMNSVIKVLNKLDTSTTS	TELTEQ---MDEA-STSGG	DVTMETDD--VSADPTS	S	602
Query_10007	549	---A LTQAIYDELVRSTLKILSKLDLTSVRG	SEDTDQ---GDSA-GTADK	GRPVTFPL--ASSDPVS	G	607
Query_10008	551	---N IQRTLDFCFMKSFLIVRLLSFTKAF	QDVEEY-----	-----FRFGNPVF	N	594
Query_10009	594	isis[4]IIQLLYNEIVQMTIIFIRKLDLSYSTK[5]	RQEKELKKQVVKDKDK[7]EEEEQEEDvTMedQOQ[6]G[3]		G	685
Query_10010	572	---F[2]---ILYDEFISSVFKMIRTLNLAVTEA	---SDVSLGTSI-TKNND	---HEVDAn-LIAIAGS	G[4]	630
Query_10011	550	---I[4]LRESIYDAIMHCVKIVMNENLNKYRVK[5]	KYSEELDagPIE-PEAEM	-----DLi---EPLD	M[2]	613
Query_10012	535	---L IVVKYVDELVNTVLMKMLDLSRTRK	TNI BALS-----TS	GGASEADI--PVLDPES	G	587
Query_10013	551	---A MLRTRYDETLKAFHLINKLNLTHSS	SDAEDedvAVAT-SGVKE	EEPHEVDD--VSDPTT	S	610
Query_10014	550	---K IVEIIFNEYIESI IKIMNKLDLAAVKN	-DGSEVN-GGEQ-ENVDK	NQSDLVGS---SDPII	G	606
Query_10015	558	---E MSNFIYSEFISCIIEISEKLDLTYKQ	NELQQQE-----EV	SGSDEGDP--ATSDFVS	G	616
Query_10016	552	---N LLQLIYDEIISVVLTLKKLDLATTHS	MQDSEIK-DSNL-KQFET	SELYSTVS----DPFY	G	608
Query_10017	557	---S IHQLVYSEFIQSLNILRKLDFLTAST	TEDGIM-----A-NEGEG	GGSGGGDPvITSDFVL	G	614
Query_10018	547	---Y VQRSLYNELHSLQRIILKLDLTSQQ	TTIEE-----ET-QSSE-	-----GTSNPPVA	G	593
Query_10019	558	---V MHGLVYDEMIRGILRIIHKLDLSSSKG	SSAQEGDILGKS-SASDV	PASATADKevSLSDPVP	G	620
Query_10020	564	---K[4]LQAIILDSVTKHVLAIIRKLNLSYQFD	SQTCKIG-----RDA	-----T	G[1]	608
Query_10021	556	---S ISGGIFDAFLSSVLAIERLDLSACRL	SDLEPPvtAASE-ASSSS	-ATTDEDSaaPTDTLQ	G	617
Query_10022	525	---I[4]TCQYVYGELIQRVIRIISDLNLTYYKVE[3]	NNNKDGN-QDES-NNNN[7]NQELTGLHq---	DYIV	Q[2]	598

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	602	LHAVCPEDFILLTNLVELSRDLDPVQTQKFSWVFPFHGQLIQSSKLPVSGLYKMLTVCMQVASQISYFKEMNRQLK	681
Query_10002	615	LHATKPKDFQIVFNLVDFCRDLLPANQPALFKQWYRFGHEVILLSCRLLPVSGFYKLLTVCMMNICTQIQFPCIKPSDL	694
Query_10003	622	LQAACKPRDFVVFINVVDFCRDLLPHHHPEYFQRWVFTVGHALILHSTRLLPVSGFYKLLSVCMTICTKVYFKDLDTKLL	701
Query_10004	612	LHPAKPKDFSAFINLVDFCREILPEKQAEFFPEWVYSFSEYELIQSTRPLPLISGFYKLLSITVRNAKKIKYFEGVSPKSL	691
Query_10005	619	LQPVPKDFTFVFNLVDFSTELLERHEEFDQWYRFGYELILRSTQFPLVSGFYKLLSATMKIAKKINYPKGVSTNTW	698
Query_10006	603	VHASKPKDFQVFINLVDFCRELLPNKHINLFWNIFTFSHNIIILLSTQNPLVSGFYKLLALCMKISNKLAYFQ-VN----	677
Query_10007	608	LQASKPKDFQVFINLVDFSRDLHTDQYCALFRKRWFTFGHQLIVMSTKYPLVSGFYKLLAICMHATRKTAYFKDIS----	683
Query_10008	595	VIPTKQADYDLFYNLVELCRSVIMNGSHVLPFEPWQPMITMLLELSSRYPMVSGFYKLLAVVWMAEKLNYPNYTEQDKT	674
Query_10009	686	LKPTPKDFLELFINLVDFSKFLARQHTNLTFRWYIFGKEIIFGSRKYPLISGFYKLLHVLKVAKKVNYFGLDGSNK	765
Query_10010	631	LQPLVSKDFVLFQNLVDFWQLFLPRIKSELFSRWVYLAGDTLITFSTRYPVSGFYKMLGSCLTVCESIKFDFGIKNVDN	710
Query_10011	614	-----MRKYLNLVDFVQELLSGPCSSILKWTYFPGRCIIELSNRYPLISGFYKLLTVNITIANREGFLSVAGEDSL	685
Query_10012	588	LEANNRKFQVFINLVDLTRDVLMSVDCSQFEKWIQPFYKVISESSNKPLVSGFYKLLTVALQCDRLHYFSDYGDMLN	667
Query_10013	611	LIASRPKDIQVLIHLVTFQDLPSTTRTDLFPRWLHIYSDIVTRSSKYPLVSSLYKLELANTIAETEQRVFEAASKTAE	690
Query_10014	607	LKPQKPRDFEIFINLVDFTRLELMNKNDFLKSKEIILLSTSHCLISGFYKLLATIMKMAVKIYGFNKIDSIQT	686
Query_10015	611	LKAKVPKDFQVYMNVSLEKVLPLQSEDELEPHIPLRVLVFVTVQSSLOPLVSAHYNATTVLTCSRQTFQSR-----K	685
Query_10016	609	LQLSKPKDFRVFMNIVDFCKDLFISKHNELEFKWILIFGLEIIQFINKNPYISGFYKLLTVCLTICKKTRYFQHVKKLT	688
Query_10017	615	RQPAVPKDFQIFINLVELAKKVIPSVLSLDFPFWILTFSEVVSLATRYPLVSGFYKLLTVCLSICNKHDFDHDVLSGQ	694
Query_10018	594	LKPNVPKDFQVFINLVDFTKALLPTTRTFEFLPWIFPPGKNLVLSTRLLPVSGFYKLEICMTISSLSYFEDLNITV	673
Query_10019	621	LQPKVPKDFLIFISLIEFCRCVLPNVCYTLFERWYMFCDRIIMSSRFVPSGFYKLELVCKMKCEKLSFFKNIAEKPK	700
Query_10020	609	-SPTVPRDHSIMLNLTVEFVERLISQVLSLQCPWIPLMNQIFASICDLPLVSCFYRIGTVMTKAMHELNYFKSASSVDT	687
Query_10021	618	MVASTPKDIVIFVNLVDFVEELLQHRAGIAKWYPLCQALIEKSAENPLISGFYKLLATFVTRARDAGYFNDQPVPV	697
Query_10022	599	--ANNNTDQQLVRLSNFAEMFLQCKDQWFLKWIPQFTDTLTKKASQTPRISKLYQLLVKCMQICSNHRYFEVSDDDIS	676
Query_10001	682	IETSDES IKMEMDEEF[ 6 ]-EEACFMFLTKFAKEVMTRLRQYKDDLLASCLGFILALPREII ALDVTTFVPAIQ	756
Query_10002	695	SSQDES MDIDLNFSS[ 8 ]DRICICTFLFAKFSKEVLVRLKQYKDDLLASCLLVLSLPHIEIV QAEVAHLVPAIQ	772
Query_10003	702	SFSEES MDTEGRDRP[ 5 ]----CFTLFRFKVKEVLVRLKQYKDDLLASCLLVLSLPHQIV QSEISTLVPAIQ	772
Query_10004	692	KHSPED P----- EKYSFCALFVFKGKEVAVKMKQYKDELLASCLTFLLSLPHNII ELDRVAVVPAIQ	753
Query_10005	699	KQCHTD P----- ERCACAFALFGKFGKEVTRIMKQYKDELLVSCLEFVLSLPHDII ELDIKAYIPAIQ	760
Query_10006	678	----ES NEQE-----[ 1 ]QKAVCYLLIKKFSKEVLVRLKQYQDDLLAACLLTILSLPKEII TEQMIIDVVPAIQ	739
Query_10007	684	----LS RDQDGENS[ 9 ]DRETAFLFKFSKFEVLVRLKQYKDDLLASCLGFILALPPELV IADLPTIVPAIQ	758
Query_10008	675	ENCSNH[ 2 ]----- LLLSDFIDHVLVESCRCFQGELELVSCLEFLLALPTCIV VEKLEISKEAIK	731
Query_10009	766	SSSLDA[ 7 ]MEIDSGKSS[ 7 ]NKRNCFKLFAKFLKEVEAASTQYKDELLASVIQLLSVPKQLV --NIPVLIPTLS	847
Query_10010	711	ENIMIQ ---DSQDL[ 7 ]-QRASYFLFKFPEVVSARLPQYKDDLLASCLLVLSIPRELI --NIQNLISPIC	781
Query_10011	686	DADGGY[ 2 ]LDKRRDSTT[ 10 ]DRRLCKELFRKYQEVLVASRRYKLELLAACLLRCLSAPPLLV[ 1 ]----IASIIPSLR	765
Query_10012	668	QPTEQK ----- --ASYSLLCRFLQELILRLQYKNDLQVACLETVIAAPVALV VPLLPHAVPAFE	725
Query_10013	691	TAAGNM[ 1 ]----- EVAATLELMSKFGKEVMVRSKQYKDELLASCLTFVLSLPLPVV KRDFAVFPFAVE	752
Query_10014	687	NLNTSS[ 7 ]EDCDDPSS[ 9 ]ERVNCFHLFKYLLVSVIRMKQKDELLVSCLEFFLSPKPEFV IISMDHSFHCLN	772
Query_10015	686	QDPEVE ----- --TSIHLLMSYVRELLQRCRQYRDQLLASCLTILHLPLCIV QEIPFQVSPAIQ	743
Query_10016	689	KEEIS- -----[ 4 ]DCLIAFELFAKFSREVVVIYQQFKDELLASCLTEVLSLPTCIV ESEINILIPAVQ	752
Query_10017	695	GTGDD VAMETEERT DRQSCYFLFSFKEVVMQYKDDLLASCLHLILVLPKEMV VSDFKELIPAIK	764
Query_10018	674	TKCTIN AAAB-AESI DKYLCYLLYKFKYKEVLVRLKQYKDDLVHVSCLSLILALPYQLI QDDMALIIPAIK	742
Query_10019	701	DLKNEN LNLEIVDTS DCMKCFVLFHFKTKEIVVRMQYKDDLLFSCLLVVLSLPLELV TDLVSSVAALK	770
Query_10020	688	TEQTRF[ 3 ]----- -----FTSCVQVSSQARFYQDELRLTSSEFVLTAPLGLV A--IHAMVGIVK	741
Query_10021	698	MEIADT[ 7 ]QDSQTDNDE[ 9 ]NRIQCCAMLSGFAASISARCGQLRGDLLAASLRFLLSLPASLL[ 5 ]----DDALAAVR	784
Query_10022	677	VAKKGE ----- -KQNTYNMLLTFPKELIGKQEYQDELLTSCLDLLLHVVPQIL[ 6 ]TDNVQNWQVIMQ	742
Query_10001	757	RSLEIGRSYLPLARTSLNLTLESWSKSLPP DVMKPFYEQLLPYLDAYLTSSAD-e----KNDTLSLVQKDPMKK-	824
Query_10002	773	MTFQLGLSYLPLAETGLEALERWMLHLLP DSIQPHLKVILPCLDGYLRGTIT-qGSDAA-DVVQVTMTASASS-[ 3 ]	846
Query_10003	773	TTFRLGLSYLPLAHAGLDALERWLLNLRP SALEPHPHILPYLDYSYLQTAAD-tGSSSSSDGFQTVSLTAASS-[ 1 ]	845
Query_10004	754	MAFKLGLSYTPLAEVGLNALEWISYIDR HVMQPYKIDILPCLDGYLKTSAL---SDETKNNWEVSAALSRAAQK	824
Query_10005	761	AAFKLGLSHTPLADAGLDALDAEDWSAHIPK HIQPHYKIDILPHLDGYLKTAAF---N-EAQNNWEGMNNINNSRKR	830
Query_10006	740	VALTGLSYPPLAGIADDALEYWCHNLPD HIVKPHYSCILPYLDNYIRTVDK---GMVAFSIFMWYISCGMHIR[ 8 ]	818
Query_10007	759	ITFKLGLSYPPLAVEGMEAIETFRWDSIPN HLLQPHYQDILPCLDGFKLQSE-qAGNETNSPSTLIMEPKGH--[ 2 ]	831
Query_10008	732	SSLKGLNLYPLIACVCMDALEKWSAEFPE[ 3 ]-QLDSFYTDLLPFDYDLRCTSV-----DETLVNNGFNPNYNGEK[ 8 ]	809
Query_10009	848	TAFKLGLSYLKLGHVGLNLTIEYWLVLPE[ 1 ]--VDAHLYAILPALSDFYLISTT-----NTSSSNQESANSK[ 4 ]	815
Query_10010	828	IALKMGLSYPLATIGINAIEKLVLDQCP[ 2 ]--ISKWLEQILPLFNEIYLLVKAD---SVDNSIKEFAQIRNTQRNRR[ 8 ]	861
Query_10011	766	NALKIGLRYPLAEVAMDSLERWQAAPPE[ 1 ]--LQHWLPQIIPCLNDYLVVVEK---ADVSEPGTSLVTEAVGTSAA[ 6 ]	841
Query_10012	726	VIFVRGRSMLPLARGLSALDKWYCTLPS HLVTPLLMQVLPYLDPPFRSQGV-aSGQDV---IIEKELMVLGRR	795
Query_10013	753	TALSGLGYPPLADAGLSALETW---LPP[ 2 ]-AAGRPFPAQLMPLLDKYLAAVA-aSDIDVEEMKLLSRSITRKR	823
Query_10014	773	QALDLGVNLYKLAEADALSLEYWLNKNIPT QLKPYFPSILTQFENYQASSN--KS---DLDSDEKVIILLNLSHK[ 2 ]	844
Query_10015	744	MALQGVSYLPLAEVCILALQWTKLSLQ EILSEHMPHVLPPLLPLRYSQET-gGEAEVHTRVITVTKMAFARQR	816
Query_10016	753	ISLQGLSYPPLAMSCDALETWLYSLPS ELLKEHYCKLLPSLNDYFLASDN--ISLQPLKNIHTFKQSLVVK	824
Query_10017	765	LALRIGMGYPLAKSVMBALEKWTLYIVE[ 4 ]DTLQCLVNEILPLLADYIKATSV-----QLKEEEDGGRG[ 3 ]	834
Query_10018	743	VSMNIGLSYPLANAADAALERWSSDLQP NIGYPYSDILPCFDYDKSSAD----DTTETLVEDSSKALKART	812
Query_10019	771	LTFKGLSYPPLAEAGLDALESWTALPA DKLKPLRILQCLDGYLTTGLDqLHNECNTIVLITSSH---HT	841
Query_10020	742	SMLDLGRSFTPAAVVAIAALERWQKQRP[ 1 ]--LNDVICDIPLISAYLNQEKL-----GER[ 4 ]	799
Query_10021	785	RGLRLGRFWPPLLEAALALERWTAAS-S[ 3 ]--LSDRVAPVLGGLLAVLLSDGG---SGAGLTDAAEAERGGADR[ 8 ]	864
Query_10022	743	KAILDLGRNNQLAWNINMLEQWPNLLPM NVTLDDLYHDILPKLSDFLQVDDN-----KK[ 8 ]	804

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	825	QGKKKIPVKLLKQ--TMFDA--EKGESESELK <b>IQRKCVLL</b> LGS <sup>1</sup> GG <sup>2</sup> INHALVFGQTANP---PD-L	VAVDSDQ--	890
Query_10002	847	AGKRKLPAKLIKGG--ASTTG--QT-QDSPLR <b>VQRIHL</b> LGS <sup>1</sup> GG <sup>2</sup> QTNAYLE--TSAEETA-KV-A	VAWDTQP--	912
Query_10003	846	RGKRKIPVKLLKA--AAERN--QV-QESDLV <b>KARIRI</b> VR <b>FLGS</b> LGG <sup>2</sup> H <sup>3</sup> TNA <sup>4</sup> CLV <sup>5</sup> GLTAEDLA-KR-A	VAWDTQT--	912
Query_10004	825	GFNKVVLKHLKKT--KNLSS--NEA--ISLE <b>EIRIRRV</b> Q <b>MQLGS</b> LGG <sup>2</sup> QINKNLLTVSSDEMM-KS-Y	VAVDREK--	890
Query_10005	831	GGSKVLFPHHTKKS--KEPLV--SED--S <b>AITAVRRRI</b> VR <b>L</b> LGS <sup>1</sup> GG <sup>2</sup> QNNRSLVTVASAEEMM-KR-C	I <sup>1</sup> AWVTEK--	896
Query_10006	819	SSFSTINCLLY-----F-----Q <b>S</b> -YESKLG <b>V</b> <b>IKDRV</b> V <b>Q</b> Y <b>L</b> GS <sup>1</sup> LGS <sup>2</sup> IN <b>Y</b> TL <b>L</b> T--NSQDDIS-KQ-A	I <sup>1</sup> AWDTEQ--	878
Query_10007	832	SRKGLPIKLLKP--AVA----KE-ESSQLE <b>LKQRI</b> V <b>K</b> FLGN <b>L</b> GGATNA <b>M</b> VQ--CTSEDIA-KI-A	V <sup>1</sup> AWDQEK--	895
Query_10008	810	QTVNSDSENVLVK--ADDND--SEEEMPN <b>V</b> <b>SLRR</b> IAS <b>FVGS</b> LGCYMN <b>R</b> FLID----SADVG--S-L	LRWQGS--	873
Query_10009	916	MGRGAAPRMMTKSi--DFVdpkLLAQR <b>TV</b> VE <b>ES</b> L <b>EIRIRRV</b> Q <b>M</b> LGS <sup>1</sup> GG <sup>2</sup> QIN <b>Y</b> HL <b>M</b> GS <sup>3</sup> VNL <b>M</b> AP <b>G</b> Te <b>S</b> G--	V <sup>1</sup> AWDTIS--	985
Query_10010	862	TRKERVKSGQK <b>K</b> syK <b>T</b> VVGi <b>g</b> Q <b>L</b> E <b>K</b> N <b>V</b> L <b>G</b> LD <b>Q</b> <b>V</b> R <b>V</b> L <b>R</b> L <b>L</b> GR <b>L</b> GG <b>V</b> SK <b>F</b> L <b>F</b> ED <b>I</b> V <b>N</b> ND <b>L</b> V <b>Y</b> r <b>P</b> i <b>P</b> T[13]	L <sup>1</sup> AWDPEK--	948
Query_10011	842	SRGQLAQIKRARA-----KDSQP <b>S</b> PFEL <b>I</b> <b>Q</b> <b>V</b> R <b>V</b> Q <b>R</b> FL <b>G</b> CIG <b>L</b> GH <b>F</b> L <b>V</b> PE <b>V</b> DL <b>L</b> Q <b>L</b> T <b>G</b> t <b>G</b> G--	L <sup>1</sup> AWDPID--	904
Query_10012	796	--NRTNYSQNKV--ARKKI--IEHPESEL <b>S</b> <b>Q</b> L <b>H</b> E <b>I</b> S <b>F</b> L <b>G</b> Q <b>L</b> D <b>S</b> T <b>T</b> C <b>L</b> A <b>L</b> L <b>K</b> D <b>P</b> Q <b>Q</b> E <b>B</b> D <b>V</b> -----	I <sup>1</sup> PWATCLqq	860
Query_10013	824	-----E--QA <b>AV</b> V--AVDS <b>DN</b> PF <b>A</b> L <b>K</b> M <b>R</b> I <b>L</b> R <b>L</b> L <b>G</b> A <b>M</b> G <b>Q</b> H <b>N</b> E <b>L</b> A <b>A</b> L <b>E</b> H <b>Q</b> G <b>A</b> G <b>G</b> g <b>S</b> D-V[8]	V <sup>1</sup> AWDMTN--	888
Query_10014	845	--GRK <b>L</b> PVK <b>L</b> FD--K <b>S</b> SLN--FDV <b>Y</b> E---- <b>Q</b> <b>I</b> <b>Q</b> L <b>K</b> I <b>L</b> K <b>I</b> L <b>G</b> Q <b>L</b> S <b>A</b> D <b>M</b> S <b>H</b> V <b>L</b> -ID----N-ED-I	T <b>S</b> W <b>D</b> T <b>V</b> Q--	901
Query_10015	817	--R <b>K</b> V <b>D</b> ---Q <b>K</b> M--K <b>E</b> SKR--VE--ET <b>M</b> A <b>M</b> Q <b>I</b> <b>Q</b> L <b>M</b> I <b>R</b> L <b>L</b> I <b>G</b> S <b>L</b> D <b>L</b> -L <b>R</b> L <b>H</b> T <b>I</b> P <b>Q</b> D <b>P</b> D <b>T</b> I <b>a</b> S <b>A</b> -A	V <b>R</b> W <b>N</b> --V <b>g</b> k	876
Query_10016	825	---K <b>L</b> L <b>P</b> L <b>K</b> I <b>S</b> Q <b>Y</b> --K <b>K</b> I <b>K</b> Q--T <b>D</b> L <b>H</b> E <b>S</b> E <b>L</b> <b>I</b> <b>Q</b> <b>V</b> R <b>I</b> L <b>R</b> FL <b>G</b> S <b>V</b> G <b>N</b> Q <b>Y</b> N <b>F</b> A <b>M</b> I <b>S</b> -AD <b>K</b> I <b>K</b> I <b>H</b> -E <b>V</b> -A	V <sup>1</sup> AWDINKqn	890
Query_10017	835	AT <b>K</b> R <b>Y</b> AR <b>M</b> K <b>M</b> VE <b>K</b> --W <b>S</b> E <b>G</b> L--H <b>Q</b> S <b>R</b> D <b>I</b> T <b>S</b> <b>V</b> R <b>R</b> I <b>L</b> L <b>L</b> L <b>S</b> L <b>G</b> G <b>V</b> V <b>N</b> E <b>S</b> L <b>K</b> V <b>S</b> P <b>S</b> Q <b>H</b> S <b>L</b> -----	V <b>S</b> W <b>S</b> D <b>R</b> --	898
Query_10018	813	G <b>K</b> A <b>K</b> L <b>P</b> V <b>K</b> V <b>V</b> N <b>P</b> --T <b>G</b> G <b>S</b> K--D <b>D</b> ----S <b>P</b> L <b>A</b> Q <b>V</b> R <b>I</b> <b>R</b> I <b>Q</b> L <b>L</b> GR <b>L</b> GG <b>N</b> T <b>N</b> -S <b>L</b> V <b>I</b> R <b>K</b> G <b>N</b> D <b>D</b> G <b>S</b> -S <b>Q</b> -A	V <sup>1</sup> AWDNHK--	875
Query_10019	842	S <b>R</b> G <b>R</b> K <b>L</b> P <b>L</b> K <b>L</b> L <b>K</b> M--K <b>P</b> Q <b>S</b> Q--H <b>Q</b> M <b>T</b> L <b>T</b> P <b>L</b> A <b>K</b> <b>V</b> R <b>I</b> L <b>R</b> M <b>F</b> G <b>L</b> GG <b>E</b> T <b>N</b> -I <b>L</b> L <b>L</b> -D <b>N</b> T <b>G</b> I <b>N</b> S--S <b>N</b> -A	V <b>V</b> W <b>D</b> T <b>E</b> N--	907
Query_10020	800	---D <b>K</b> A <b>T</b> L <b>Q</b> G <b>V</b> K <b>A</b> --T <b>G</b> S <b>D</b> E <b>a</b> -----S <b>D</b> L <b>A</b> Q <b>V</b> R <b>R</b> L <b>L</b> L <b>L</b> GR <b>C</b> D <b>G</b> --N <b>L</b> L <b>L</b> V <b>S</b> E <b>P</b> S <b>L</b> I <b>T</b> ---A[2]	I <b>V</b> GG <b>I</b> F <b>S</b> -p	859
Query_10021	865	R <b>A</b> L <b>R</b> S <b>I</b> E <b>A</b> R <b>R</b> A <b>R</b> A--K <b>N</b> Y <b>A</b> A <b>g</b> g--G <b>A</b> E <b>L</b> G <b>L</b> A <b>E</b> M <b>R</b> K <b>V</b> L <b>L</b> FL <b>G</b> S <b>A</b> Q <b>S</b> S <b>F</b> K <b>Y</b> <b>Q</b> L <b>A</b> E <b>I</b> A <b>S</b> D <b>K</b> A <b>A</b> T <b>a</b> D <b>L</b> -A	--A <b>W</b> S <b>P</b> G <b>G</b> -r	933
Query_10022	805	GG <b>D</b> T <b>Y</b> Q <b>L</b> E <b>M</b> L <b>K</b> S--G <b>N</b> E <b>V</b> N <b>1</b> -----E <b>R</b> K <b>D</b> I <b>A</b> L <b>K</b> V <b>M</b> D <b>L</b> L <b>G</b> K <b>I</b> G <b>G</b> NA <b>H</b> -S <b>I</b> S <b>N</b> E <b>Q</b> N <b>K</b> L <b>I</b> E <b>r</b> E <b>N</b> -F	I <b>K</b> W <b>D</b> P <b>E</b> K--	867
Query_10001	891	<b>H</b> L <b>K</b> F <b>D</b> L <b>P</b> F <b>L</b> D <b>I</b> K--T <b>S</b> I <b>Y</b> F <b>D</b> A <b>F</b> L <b>P</b> K <b>V</b> V <b>T</b> L <b>A</b> T <b>V</b> S <b>S</b> D <b>R</b> Q <b>T</b> K <b>V</b> A <b>C</b> E <b>L</b> L <b>H</b> S <b>L</b> V <b>L</b> L <b>M</b> L <b>G</b> R <b>S</b> A <b>S</b> Q <b>P</b> T <b>G</b> ----	N <b>P</b> M <b>A</b> A <b>I</b> Y <b>R</b> K <b>L</b> F	962
Query_10002	913	<b>R</b> L <b>Q</b> F <b>A</b> V <b>P</b> F <b>V</b> D <b>M</b> K--P <b>N</b> I <b>D</b> L <b>D</b> R <b>Y</b> L <b>P</b> R <b>V</b> V <b>E</b> L <b>A</b> L <b>T</b> S <b>G</b> D <b>R</b> Q <b>T</b> K <b>V</b> A <b>C</b> E <b>S</b> L <b>H</b> T <b>L</b> V <b>L</b> F <b>M</b> L <b>G</b> K <b>T</b> Q <b>Q</b> P <b>E</b> Q <b>R</b> Q <b>A</b> K	S <b>P</b> M <b>Q</b> H <b>L</b> Y <b>R</b> R <b>V</b> F	988
Query_10003	913	<b>H</b> L <b>Q</b> F <b>A</b> V <b>P</b> F <b>M</b> D <b>M</b> K--P <b>T</b> V <b>C</b> L <b>D</b> P <b>L</b> P <b>R</b> I <b>C</b> E <b>L</b> A <b>T</b> T <b>S</b> S <b>D</b> R <b>Q</b> T <b>K</b> V <b>A</b> C <b>E</b> L <b>L</b> H <b>S</b> L <b>V</b> L <b>M</b> V <b>G</b> Q <b>A</b> Q <b>Q</b> P <b>G</b> M <b>O</b> A <b>K</b>	N <b>P</b> M <b>T</b> R <b>L</b> Y <b>K</b> K <b>V</b> F	988
Query_10004	911	<b>R</b> L <b>S</b> F <b>A</b> V <b>P</b> F <b>R</b> E <b>M</b> K--P <b>V</b> I <b>F</b> L <b>D</b> V <b>F</b> L <b>P</b> R <b>V</b> T <b>E</b> L <b>A</b> L <b>T</b> A <b>S</b> D <b>R</b> Q <b>T</b> K <b>V</b> A <b>C</b> E <b>L</b> L <b>H</b> S <b>M</b> V <b>M</b> F <b>L</b> G <b>K</b> A <b>T</b> Q <b>M</b> P <b>E</b> G <b>G</b> Q <b>G</b> A	P <b>P</b> M <b>Y</b> Q <b>L</b> Y <b>K</b> R <b>T</b> F	966
Query_10005	897	<b>P</b> L <b>S</b> F <b>A</b> V <b>P</b> F <b>M</b> D <b>M</b> K--P <b>I</b> I <b>Y</b> L <b>D</b> S <b>F</b> I <b>P</b> R <b>I</b> S <b>E</b> L <b>A</b> S <b>A</b> S <b>D</b> R <b>Q</b> T <b>K</b> V <b>A</b> C <b>E</b> L <b>L</b> H <b>S</b> V <b>V</b> A <b>Y</b> T <b>L</b> G <b>K</b> G <b>S</b> Q <b>I</b> S <b>E</b> Q <b>Q</b> E <b>T</b>	--L <b>Y</b> K <b>L</b> N <b>R</b> O <b>L</b> F	970
Query_10006	879	<b>H</b> L <b>R</b> F <b>E</b> T <b>P</b> F <b>I</b> D <b>M</b> K--P <b>I</b> I <b>Y</b> F <b>D</b> P <b>L</b> P <b>H</b> V <b>V</b> E <b>L</b> A <b>T</b> S <b>S</b> D <b>R</b> Q <b>T</b> K <b>V</b> A <b>C</b> E <b>L</b> L <b>H</b> S <b>L</b> I <b>L</b> C <b>L</b> G <b>R</b> G <b>A</b> Q <b>Q</b> P <b>E</b> V <b>Q</b> K <b>R</b>	Y <b>P</b> M <b>Q</b> K <b>L</b> Y <b>K</b> K <b>L</b> F	954
Query_10007	896	<b>H</b> L <b>K</b> F <b>D</b> M <b>P</b> F <b>M</b> D <b>V</b> K--A <b>A</b> L <b>Y</b> L <b>D</b> P <b>L</b> P <b>R</b> V <b>V</b> E <b>L</b> A <b>L</b> S <b>S</b> D <b>R</b> Q <b>T</b> K <b>V</b> A <b>C</b> E <b>L</b> F <b>H</b> S <b>M</b> V <b>L</b> M <b>L</b> G <b>R</b> S <b>V</b> Q <b>Q</b> P <b>E</b> G <b>Q</b> Q <b>R</b>	H <b>P</b> L <b>D</b> G <b>L</b> Y <b>R</b> N <b>V</b> F	971
Query_10008	874	<b>H</b> L <b>E</b> F <b>I</b> C <b>P</b> F <b>P</b> D <b>M</b> K--F <b>E</b> I <b>A</b> L <b>D</b> T <b>L</b> L <b>P</b> R <b>L</b> C <b>D</b> L <b>A</b> L <b>S</b> A <b>G</b> D <b>R</b> P <b>L</b> V <b>A</b> C <b>E</b> A <b>L</b> H <b>A</b> C <b>V</b> L <b>F</b> L <b>L</b> G <b>N</b> E <b>S</b> F <b>R</b> T <b>V</b> K <b>M</b> K <b>D</b> R	E <b>S</b> I <b>S</b> K <b>A</b> Y <b>S</b> H <b>L</b> F	949
Query_10009	986	<b>R</b> I <b>Q</b> L <b>Q</b> I <b>P</b> L <b>G</b> D <b>L</b> g <b>V</b> S <b>I</b> Y <b>L</b> D <b>S</b> V <b>L</b> P <b>V</b> I <b>V</b> T <b>L</b> A <b>T</b> R <b>S</b> T <b>H</b> R <b>Q</b> T <b>K</b> V <b>V</b> A <b>C</b> E <b>L</b> L <b>H</b> S <b>I</b> V <b>L</b> M <b>V</b> G <b>T</b> N <b>S</b> -----G <b>G</b> S	--F <b>H</b> R <b>L</b> Y <b>R</b> R <b>V</b> F	1054
Query_10010	949	<b>R</b> V <b>S</b> F <b>N</b> I <b>P</b> F <b>P</b> D <b>T</b> K--V <b>E</b> I <b>L</b> F <b>D</b> E <b>F</b> L <b>P</b> R <b>I</b> V <b>E</b> L <b>A</b> E <b>S</b> A <b>P</b> D <b>R</b> K <b>T</b> I <b>V</b> A <b>C</b> E <b>S</b> L <b>H</b> S <b>I</b> V <b>I</b> M <b>I</b> G <b>K</b> S <b>A</b> F <b>Q</b> L <b>S</b> R <b>Q</b> S <b>S</b>	--P <b>F</b> Y <b>R</b> I <b>Y</b> R <b>K</b> I <b>F</b>	1023
Query_10011	905	<b>R</b> V <b>K</b> F <b>T</b> L <b>P</b> F <b>R</b> Q <b>D</b> K--P <b>I</b> I <b>W</b> L <b>D</b> S <b>L</b> S <b>R</b> M <b>V</b> E <b>L</b> L <b>S</b> S <b>D</b> R <b>R</b> M <b>K</b> V <b>T</b> A <b>S</b> E <b>L</b> L <b>Q</b> A <b>V</b> C <b>L</b> L <b>M</b> L <b>G</b> N <b>A</b> A <b>K</b> G <b>P</b> M <b>R</b> R <b>R</b> E[4]	V <b>H</b> F <b>D</b> K <b>I</b> Y <b>K</b> R <b>L</b> F	984
Query_10012	861	<b>P</b> L <b>V</b> F <b>L</b> P <b>F</b> K <b>D</b> M <b>K</b> --L <b>N</b> I <b>N</b> L <b>E</b> K <b>L</b> I <b>P</b> R <b>V</b> V <b>E</b> L <b>A</b> L <b>T</b> S <b>S</b> D <b>R</b> T <b>R</b> F <b>T</b> A <b>C</b> E <b>L</b> L <b>H</b> G <b>I</b> V <b>M</b> L <b>L</b> G <b>S</b> Q <b>E</b> A--Q <b>T</b> S <b>H</b> P	D <b>K</b> Y <b>S</b> E <b>L</b> F <b>R</b> K <b>L</b> S	934
Query_10013	889	<b>H</b> L <b>K</b> F <b>A</b> V <b>P</b> F <b>Q</b> E <b>M</b> K--P <b>E</b> V <b>Q</b> L <b>D</b> P <b>L</b> P <b>R</b> I <b>V</b> E <b>L</b> A <b>L</b> H <b>S</b> D <b>R</b> Q <b>T</b> T <b>A</b> A <b>C</b> E <b>L</b> L <b>H</b> A <b>M</b> V <b>L</b> M <b>L</b> G <b>Q</b> G <b>A</b> G <b>S</b> G <b>T</b> H <b>A</b> A <b>E</b>	D <b>N</b> L <b>S</b> P <b>L</b> Y <b>R</b> H <b>V</b> L	964
Query_10014	902	<b>H</b> L <b>K</b> F <b>Y</b> V <b>P</b> F <b>T</b> D <b>I</b> K--P <b>T</b> Y <b>F</b> D <b>R</b> F <b>L</b> P <b>R</b> V <b>I</b> H <b>L</b> A <b>Q</b> N <b>S</b> P <b>S</b> R <b>Q</b> T <b>K</b> I <b>N</b> A <b>C</b> E <b>L</b> L <b>H</b> A <b>L</b> V <b>I</b> Y <b>T</b> I <b>G</b> K <b>S</b> V <b>S</b> E <b>P</b> E <b>G</b> Y <b>S</b> Q <b>S</b>	Y <b>Q</b> M <b>T</b> K <b>I</b> Y <b>H</b> R <b>L</b> Y	977
Query_10015	877	<b>L</b> V <b>K</b> F <b>A</b> T <b>P</b> F <b>E</b> H <b>I</b> K--I <b>D</b> I <b>Y</b> L <b>D</b> D <b>L</b> L <b>P</b> H <b>V</b> V <b>D</b> L <b>A</b> V <b>N</b> S <b>S</b> D <b>R</b> Q <b>T</b> K <b>V</b> A <b>S</b> E <b>L</b> L <b>H</b> S <b>I</b> I <b>L</b> M <b>I</b> G <b>T</b> G <b>A</b> Q <b>S</b> Q <b>E</b> T <b>Q</b> L <b>K</b>	E <b>P</b> M <b>A</b> P <b>L</b> Y <b>R</b> H <b>I</b> F	952
Query_10016	891	<b>H</b> L <b>K</b> F <b>Y</b> L <b>P</b> F <b>I</b> D <b>V</b> K--L <b>E</b> I <b>F</b> L <b>D</b> K <b>F</b> L <b>P</b> R <b>I</b> V <b>T</b> L <b>S</b> T <b>A</b> S <b>Q</b> R <b>Q</b> T <b>K</b> V <b>A</b> C <b>E</b> L <b>L</b> H <b>A</b> V <b>V</b> T <b>F</b> M <b>V</b> G <b>C</b> G <b>A</b> Q <b>Q</b> P <b>E</b> L <b>Q</b> A <b>I</b>	Y <b>S</b> M <b>K</b> K <b>L</b> Y <b>K</b> N <b>V</b> L	966
Query_10017	899	<b>R</b> L <b>G</b> F <b>A</b> V <b>P</b> F <b>S</b> D <b>L</b> K--P <b>E</b> I <b>Y</b> L <b>D</b> P <b>L</b> P <b>T</b> V <b>C</b> E <b>L</b> A <b>L</b> S <b>S</b> D <b>R</b> Q <b>T</b> K <b>V</b> A <b>C</b> E <b>F</b> L <b>H</b> A <b>A</b> M <b>L</b> F <b>V</b> I <b>G</b> K <b>T</b> Q <b>Q</b> --S <b>V</b> Q <b>K</b> K	N <b>P</b> M <b>T</b> K <b>L</b> Y <b>Q</b> K <b>L</b> I	972
Query_10018	876	<b>R</b> L <b>F</b> F <b>A</b> I <b>P</b> F <b>Q</b> D <b>M</b> K--P <b>I</b> I <b>H</b> L <b>D</b> P <b>L</b> P <b>R</b> V <b>V</b> E <b>L</b> A <b>L</b> S <b>S</b> D <b>R</b> Q <b>T</b> K <b>I</b> A <b>C</b> E <b>L</b> L <b>H</b> A <b>L</b> V <b>L</b> M <b>I</b> G <b>R</b> G <b>A</b> Q <b>Q</b> P <b>G</b> G <b>T</b> G <b>S</b> V	T <b>N</b> Y <b>G</b> N <b>L</b> Y <b>K</b> K <b>I</b> F	951
Query_10019	908	<b>H</b> L <b>T</b> F <b>A</b> V <b>P</b> F <b>Q</b> D <b>M</b> K--P <b>S</b> I <b>Y</b> L <b>D</b> P <b>L</b> P <b>R</b> V <b>V</b> E <b>L</b> A <b>A</b> K <b>S</b> S <b>D</b> R <b>Q</b> T <b>K</b> V <b>A</b> C <b>E</b> C <b>L</b> H <b>S</b> L <b>V</b> H <b>M</b> L <b>G</b> K <b>S</b> V <b>T</b> Q <b>P</b> Q <b>G</b> F <b>K</b> K--	S <b>P</b> M <b>E</b> H <b>L</b> W <b>K</b> H <b>V</b> F	982
Query_10020	860	<b>H</b> F <b>R</b> L <b>E</b> L <b>R</b> L <b>S</b> E <b>T</b> P--L <b>S</b> L <b>A</b> L <b>D</b> S <b>V</b> L <b>S</b> H <b>L</b> G <b>N</b> L <b>A</b> Q <b>S</b> S <b>V</b> R <b>R</b> V <b>K</b> S <b>V</b> A <b>C</b> E <b>S</b> Y <b>H</b> A <b>M</b> I <b>C</b> Y <b>L</b> C <b>G</b> K <b>T</b> A <b>T</b> H <b>P</b> H <b>T</b> A <b>G</b> T <b>K</b>	T <b>D</b> F <b>Y</b> D <b>L</b> W <b>R</b> E <b>V</b> F	935
Query_10021	934	<b>L</b> L <b>Y</b> F <b>D</b> L <b>P</b> F <b>A</b> D <b>S</b> K--P <b>R</b> L <b>Y</b> L <b>D</b> P <b>L</b> S <b>R</b> C <b>V</b> A <b>L</b> A <b>T</b> S <b>S</b> S <b>Q</b> S <b>K</b> V <b>A</b> A <b>A</b> E <b>C</b> L <b>H</b> S <b>I</b> V <b>L</b> L <b>M</b> T <b>G</b> R <b>C</b> A <b>F</b> A <b>S</b> T <b>S</b> G <b>N</b> D <b>R</b>	E <b>Q</b> L <b>L</b> T <b>L</b> W <b>R</b> R <b>L</b> L	1009
Query_10022	868	<b>R</b> L <b>K</b> F <b>T</b> V <b>P</b> L <b>Y</b> T <b>K</b> K--V <b>D</b> I <b>Y</b> F <b>D</b> A <b>C</b> L <b>P</b> K <b>V</b> V <b>D</b> L <b>A</b> K <b>N</b> S <b>S</b> D <b>R</b> E <b>T</b> T <b>V</b> A <b>C</b> E <b>F</b> L <b>H</b> A <b>L</b> I <b>Y</b> M <b>I</b> G <b>K</b> N <b>A</b> S <b>A</b> P <b>K</b> R <b>G</b> N <b>G</b> [5]	A <b>P</b> F <b>A</b> K <b>I</b> Y <b>K</b> K <b>L</b> F	948
Query_10001	963	<b>P</b> Q <b>M</b> L <b>K</b> L <b>A</b> D <b>T</b> D <b>V</b> E <b>Q</b> <b>V</b> S <b>R</b> Q <b>L</b> F <b>E</b> P <b>L</b> M <b>L</b> Q <b>I</b> I <b>H</b> W <b>F</b> T <b>K</b> N <b>N</b> S <b>H</b> R <b>P</b> D <b>T</b> E <b>A</b> L <b>L</b> N <b>S</b> L <b>M</b> D <b>G</b> I <b>V</b> S <b>P</b> S--D <b>S</b> A <b>V</b> R <b>D</b> F <b>S</b> A <b>K</b> C <b>I</b> R <b>E</b> F <b>L</b> A <b>W</b> S <b>I</b> K <b>Q</b> --	1039	
Query_10002	989	<b>P</b> A <b>L</b> F <b>R</b> L <b>S</b> C <b>D</b> V <b>E</b> Q <b>V</b> A <b>R</b> Q <b>L</b> F <b>E</b> P <b>L</b> V <b>M</b> Q <b>L</b> I <b>H</b> W <b>F</b> T <b>G</b> N <b>K</b> F <b>E</b> S <b>E</b> T <b>T</b> A <b>L</b> L <b>N</b> A <b>M</b> L <b>D</b> G <b>V</b> I <b>H</b> P <b>T</b> --D <b>T</b> S <b>L</b> R <b>D</b> F <b>S</b> A <b>C</b> L <b>K</b> E <b>F</b> L <b>K</b> W <b>S</b> I <b>K</b> Q--	1065	
Query_10003	989	<b>P</b> V <b>L</b> M <b>Q</b> L <b>A</b> C <b>D</b> V <b>E</b> Q <b>V</b> A <b>R</b> Q <b>L</b> F <b>D</b> P <b>L</b> M <b>M</b> Q <b>L</b> I <b>H</b> W <b>L</b> T <b>N</b> N <b>R</b> K <b>F</b> E <b>S</b> E <b>T</b> T <b>I</b> T <b>L</b> L <b>D</b> A <b>I</b> M <b>D</b> G <b>L</b> V <b>H</b> P <b>T</b> --D <b>T</b> A <b>L</b> R <b>D</b> F <b>S</b> S <b>R</b> C <b>V</b> R <b>E</b> F <b>L</b> Q <b>W</b> S <b>I</b> K <b>Q</b> --	1065	
Query_10004	967	<b>P</b> V <b>L</b> L <b>R</b> L <b>A</b> C <b>D</b> V <b>Q</b> T <b>R</b> Q <b>L</b> Y <b>E</b> P <b>L</b> V <b>M</b> Q <b>L</b> I <b>H</b> W <b>F</b> T <b>N</b> N <b>K</b> F <b>E</b> S <b>Q</b> D <b>T</b> V <b>A</b> L <b>L</b> E <b>A</b> L <b>D</b> G <b>I</b> V <b>D</b> P <b>V</b> --D <b>S</b> T <b>L</b> R <b>D</b> F <b>C</b> G <b>R</b> C <b>I</b> R <b>E</b> F <b>L</b> K <b>W</b> S <b>I</b> K <b>Q</b> --	1043	
Query_10005	971	<b>P</b> V <b>L</b> L <b>R</b> L <b>G</b> C <b>D</b> V <b>D</b> T <b>T</b> R <b>Q</b> L <b>F</b> E <b>P</b> L <b>V</b> Q <b>L</b> I <b>H</b> W <b>F</b> T <b>N</b> N <b>R</b> K <b>F</b> E <b>S</b> R <b>D</b> T <b>V</b> A <b>M</b> L <b>E</b> A <b>I</b> L <b>D</b> G <b>I</b> V <b>D</b> P <b>V</b> --D <b>S</b> T <b>L</b> R <b>D</b> F <b>C</b> G <b>C</b> V <b>R</b> E <b>F</b> L <b>K</b> W <b>S</b> I <b>K</b> Q--	1047	
Query_10006	955	<b>P</b> A <b>M</b> L <b>K</b> L <b>A</b> C <b>D</b> V <b>E</b> Q <b>V</b> A <b>K</b> Q <b>L</b> F <b>E</b> P <b>L</b> T <b>M</b> Q <b>L</b> I <b>H</b> W <b>F</b> T <b>N</b> N <b>K</b> M <b>G</b> E <b>S</b> E <b>T</b> M <b>A</b> L <b>L</b> D <b>T</b> I <b>F</b> D <b>S</b> I <b>Q</b> A <b>D</b> --D <b>T</b> S <b>L</b> R <b>D</b> F <b>S</b> G <b>C</b> L <b>R</b> E <b>F</b> L <b>H</b> W <b>S</b> L <b>K</b> Q--	1031	
Query_10007	972	<b>P</b> A <b>M</b> L <b>Q</b> L <b>A</b> C <b>D</b> V <b>E</b> L <b>V</b> A <b>R</b> Q <b>L</b> F <b>E</b> P <b>L</b> T <b>M</b> Q <b>L</b> I <b>H</b> W <b>F</b> T <b>R</b> S <b>K</b> V <b>F</b> E <b>N</b> P <b>E</b> T <b>A</b> A <b>L</b> L <b>N</b> A <b>V</b> M <b>E</b> G <b>L</b> D <b>I</b> D <b>P</b> S--N <b>A</b> S <b>L</b> R <b>D</b> F <b>C</b> A <b>S</b> C <b>I</b> R <b>E</b> F <b>L</b> K <b>W</b> A <b>L</b> K <b>Q</b> --	1048	
Query_10008	950	<b>P</b> V <b>L</b> L <b>K</b> L <b>A</b> C <b>D</b> V <b>E</b> V <b>P</b> K <b>Q</b> M <b>F</b> S <b>T</b> L <b>F</b> V <b>Q</b> I <b>L</b> H <b>L</b> F <b>S</b> G <b>T</b> D <b>F</b> A <b>N</b> K <b>M</b> V <b>T</b> S <b>C</b> L <b>L</b> D <b>T</b> L <b>M</b> D <b>G</b> L <b>C</b> F <b>W</b> --N <b>S</b> L <b>R</b> D <b>Q</b> S <b>A</b> Y <b>I</b> S <b>E</b> Y <b>C</b> R <b>W</b> S <b>M</b> R <b>K</b> --	1026	
Query_10009	1055	<b>P</b> A <b>L</b> L <b>R</b> L <b>A</b> V <b>D</b> V <b>E</b> V <b>A</b> R <b>Q</b> L <b>F</b> S <b>P</b> L <b>I</b> M <b>Q</b> L <b>V</b> H <b>W</b> F <b>T</b> K <b>N</b> V <b>R</b> A <b>E</b> N <b>D</b> E <b>T</b> M <b>A</b> L <b>L</b> N <b>S</b> I <b>D</b> A <b>V</b> N <b>G</b> E <b>D</b> --G <b>A</b> L <b>R</b> E <b>F</b> G <b>A</b> R <b>G</b> E <b>F</b> A <b>S</b> W <b>S</b> L <b>K</b> H <b>A</b> --	1131	
Query_10010	1024	<b>P</b> A <b>L</b> L <b>R</b> L <b>A</b> I <b>D</b> T <b>D</b> P <b>V</b> P <b>R</b> R <b>I</b> F <b>R</b> R <b>L</b> V <b>A</b> Q <b>L</b> I <b>H</b> W <b>F</b> T <b>N</b> N <b>A</b> Q <b>Y</b> E <b>N</b> E <b>T</b> M <b>A</b> L <b>L</b> Q <b>C</b> L <b>D</b> A <b>I</b> C <b>D</b> T <b>W</b> --G <b>S</b> L <b>R</b> D <b>Y</b> G <b>A</b> E <b>C</b> L <b>E</b> F <b>L</b> I <b>W</b> S <b>I</b> K <b>Q</b> --	1099	
Query_10011	985	<b>P</b> A <b>I</b> L <b>R</b> L <b>A</b> D <b>T</b> D <b>V</b> L <b>T</b> R <b>Q</b> L <b>F</b> D <b>V</b> K <b>Q</b> L <b>I</b> H <b>W</b> F <b>T</b> S <b>N</b> Q <b>K</b> E <b>N</b> P <b>D</b> T <b>M</b> A <b>L</b> L <b>D</b> A <b>I</b> L <b>D</b> G <b>L</b> V <b>D</b> P <b>D</b> --N <b>G</b> S <b>L</b> R <b>E</b> F <b>C</b> A <b>D</b> M <b>F</b> S <b>E</b> F <b>L</b> Q <b>W</b> S <b>V</b> R <b>H</b> --	1061	
Query_10012	935	<b>L</b> P <b>L</b> L <b>K</b> L <b>G</b> C <b>D</b> P <b>A</b> I <b>H</b> R <b>L</b> Y <b>N</b> L <b>L</b> V <b>Q</b> L <b>V</b> H <b>W</b> Y <b>S</b> S <b>F</b> K <b>L</b> K <b>S</b> E <b>V</b> E <b>L</b> I <b>Q</b> T <b>L</b> M <b>D</b> G <b>I</b> T <b>H</b> P <b>T</b> --D <b>T</b> A <b>L</b> R <b>D</b> F <b>S</b> A <b>R</b> C <b>H</b> E <b>F</b> V <b>W</b> S <b>V</b> K <b>Q</b> --	1011	
Query_10013	965	<b>P</b> A <b>V</b> L <b>R</b> L <b>A</b> C <b>D</b> V <b>D</b> L <b>V</b> P <b>K</b> Q <b>L</b> F <b>S</b> P <b>L</b> V <b>M</b> Q <b>I</b> I <b>H</b> W <b>F</b> T <b>S</b> N <b>Q</b> K <b>S</b> S <b>R</b> T--E <b>R</b> E <b>V</b> V <b>L</b> L <b>D</b> A <b>I</b> L <b>T</b> G <b>L</b> A <b>D</b> A <b>K</b> --N <b>A</b> S <b>L</b> R <b>D</b> F <b>C</b> A <b>S</b> C <b>L</b> R <b>E</b> F <b>L</b> H <b>S</b> I <b>K</b> Q--	1040	
Query_10014	978	<b>P</b> C <b>L</b> F <b>L</b> G <b>S</b> C <b>D</b> T <b>D</b> H <b>F</b> V <b>R</b> N <b>L</b> F <b>Q</b> P <b>L</b> V <b>M</b> Q <b>I</b> I <b>H</b> W <b>F</b> T <b>G</b> N <b>K</b> Y <b>E</b> S <b>S</b> D <b>T</b> I <b>Q</b> L <b>L</b> G <b>C</b> I <b>M</b> E <b>S</b> L <b>V</b> D <b>E</b> K--N <b>A</b> A <b>L</b> R <b>D</b> F <b>S</b> V <b>N</b> C <b>L</b> E <b>F</b> L <b>K</b> W <b>S</b> I <b>K</b> H--	1054	
Query_10015	953	<b>S</b> A <b>M</b> L <b>R</b> L <b>A</b> C <b>D</b> P <b>D</b> H <b>T</b> R <b>Q</b> L <b>F</b> L <b>T</b> S <b>Y</b> Q <b>T</b> I <b>H</b> W <b>F</b> T <b>N</b> N <b>R</b> F <b>E</b> N <b>V</b> D <b>T</b> M <b>V</b> L <b>L</b> E <b>A</b> L <b>M</b> E <b>G</b> I <b>V</b> T <b>A</b> N--D <b>P</b> A <b>L</b> R <b>D</b> V <b>S</b> A <b>Q</b> C <b>L</b> A <b>E</b> F <b>V</b> K <b>W</b> S <b>R</b> K <b>Q</b> --	1029	
Query_10016	967	<b>P</b> A <b>L</b> L <b>Q</b> L <b>A</b> C <b>D</b> V <b>D</b> E <b>V</b> A <b>R</b> K <b>L</b> F <b>E</b> P <b>F</b> T <b>F</b> Q <b>L</b> I <b>H</b> W <b>F</b> T <b>T</b> K <b>R</b> M <b>I</b> N <b>S</b> D <b>E</b> T <b>K</b> Q <b>F</b> V <b>Q</b> V <b>I</b> D <b>G</b> L <b>V</b> Q <b>S</b> K--D <b>A</b> V <b>F</b> R <b>D</b> F <b>S</b> A <b>R</b> C <b>L</b> E <b>F</b> V <b>Q</b> W <b>S</b> I <b>K</b> Q--	1043	
Query_10017	973	<b>P</b> V <b>A</b> L <b>Q</b> L <b>A</b> C <b>D</b> V <b>D</b> R <b>V</b> A <b>E</b> Q <b>L</b> F <b>R</b> P <b>L</b> V <b>F</b> Q <b>M</b> I <b>H</b> W <b>F</b> T <b>R</b>		

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	1040	-TSPKEMEAG---PANPKRLKRMYSLLEHPSMSMRLGGALMFNSIYVFPREED-ALVDQFSLQILSVLVNSLAL	AH	1111
Query_10002	1066	-TSKKQKEKS---PINTKSLKRLYSMALHPSFRRKLGAAALTFNHIYTVFPREED-SLIDQFVCEILVTFVNSLAL	AH	1137
Query_10003	1066	-TTPKQLEKS---PINIKSLKRLYSLALHPNPFNRKLGAAALGFNSIYRDFREES-SQVDMFTFELLVTFVNSLAL	AH	1137
Query_10004	1044	-ITPQQKEKS---PVNTRKSLFKRLYSLALHPNAFKRRLGAAALFNHIYRDFREES-SLVEQFVFEALVIYVESLAL	AH	1115
Query_10005	1048	-TTPKQKEKS---PANTKSLFKRLYSLALHPNFKRRLGAAALFNHIYRDFREES-SLVEQFVFEALVIYVESLAL	AH	1119
Query_10006	1032	-SSVKTVEKS---PVNASSILRRIFSFALHPSVFKRRLGAAALFNHIYQIFREEK-SLVDRYTFQILVHFVESLAL	AH	1103
Query_10007	1049	-TSPAQLRKN---PHIPKSLFKRLYSFALHPSFRRKLGAAALFNHIYRDFREED-VLLDIFTFEILETFIESLAM	AH	1120
Query_10008	1027	-VDSGGQNMRS---SKVELLQVTTIITFAVHSSAVKRAGAAVAFNNLYKIITREKR-WLVDIYILEMFTFMKSLYL	AE	1099
Query_10009	1132	pTTKSTANEArS-AFNFKSLKRVYSLAAHPSAQKRLGAAALSLIRVREDD-NLISVFLFEIAHQVLLSMRL[4]	AD	1210
Query_10010	1100	-SSATHAESsm---NMKSLKRLYNLCAHPDYTKRRLGASLAINNIYRDFREED-YLVDQFTEFELLXWMLVNLRL	SD	1171
Query_10011	1062	----APEDDnf---VNVRSMLRRVYNRMAHPNYPHRLGAAMAHRMYPVLENG-AIVDRYIFELFFCIKSLKL	GE	1131
Query_10012	1012	----GSHV---NIRYVMKQIYSFCLHPCPSKRLGAAALVNLVSVYTVLREEE-SVNVQFWLELLYHMVNLSL	SE	1076
Query_10013	1041	-STDKLLLETGgl---VKVLLKRMISMAAHPNAFQRMGAALAMNNIYTRLRESD-TLVDVYVCELTDFVFMRSIQL	AH	1111
Query_10014	1055	-TPLSKQENSaptPVNAKSLKRIKFNPLTHPNANKRLGACLAWSVYTIFFREEE-SIVNRHIFELIFLYIESLAM	SE	1129
Query_10015	1030	-CRQQNSVSS---NIKSLIKRIFSFCKHPSAFKRLGALAWNSIYREIREDE-QAVDIWTLDDLAAHMTSLDL	AQ	1099
Query_10016	1044	-TSVREQEEN---PIGMNVVNLCLNYSLHPNYPKRLGAAALFNHIYTIFFREEE-SLVDKFTLEIAVHFIESLAI	AH	1115
Query_10017	1050	-TSKKQKERS---PMNTRKSLKRLYSMSGSPSTAKRRLGAAALFNHIYTVLREEE-SLVDVFIIEMLVVMVNSLKL	AH	1121
Query_10018	1029	-TTKKQKEKT---SNNKSLKRLYSLASHPSAIKRLGAAALFNHIYVVLREEE-TLVDTFIFELIFVNYVNSLAL	SH	1100
Query_10019	1060	-SSEKQLEKS---PINIKSLKRLYSLALHPSASKRRLGAAALFNHIYTVFEEA-SLVDIFVLEILVNLVLECLCL	SH	1131
Query_10020	1014	----PQIKAhY-----IFNRVFLSCLRHPSVQRSGAAVSIYFLQSLREDEGALTRFAMPCLKNLLYALRL[4]	AQ	1082
Query_10021	1090	pSSGADSTASfy-SRNVKSIVKRVCLLALHPASSKRLGAAALFNVAVYQFPREED-SLVNFAFELLHCFMRSIQL	AD	1164
Query_10022	1026	-MTDKNKEMpn---NIKSLIRRIESENHPDPKRLSSVLCFSKIFSVIREHD-PLIDRFCEIAHNVLGSLKM[4]	IE	1101
Query_10001	1112	R--DDKS-L--GTQEQCKSALLHVERILIKKQVT-FNKDSKQRIKPRGWRDA---NMTTAVRWLLRQCGRPQTECRHVC	M	1182
Query_10002	1138	H--DDKS-L--GTQLHGVEVLNHLERIRVKADM-LNKESKIRRIIPREFPAGVcpTfSMLVIWLLRQCGRPQTECRHQC	M	1211
Query_10003	1138	H--DEKS-L--GTQAQCVAVLTHLDRIITTKADL-LNKHSKVRVRPRGFDFRESppTLPDLVWLLQCGRPQTECRHRC	M	1211
Query_10004	1116	A--DEKS-L--GTIQCCDAIDHLCRIIEKKHVS-LNKA-KKRRLPRGFPSPAslCLLDLVKLLAHCGRPQTECRHKS	I	1188
Query_10005	1120	A--DDNS-L--GTVQCCDAIDHLNRIIQKASA-LNKP-IKGRVPRGFPDPqVSLKEVVMWLLTQCGRPQTECRHCK	M	1192
Query_10006	1104	K--DDKS-L--GTQEQCAAVLKHQRITIFHEDI-LKKPSKIRQEPREWSNK---TLDIAVRWLSRQCGRPQTECRHTC	M	1174
Query_10007	1121	H--DDKS-L--GTQEQGVQLNHLERIVNKAAL-LCKDSKIRKPKTWNKA---TLDICVRWLLRQCGRPQTECRHSS	M	1191
Query_10008	1100	N--DDEL-Y--GTVDECKQAIKHIRRIILHNLDM-LQIHNVQRRTPNEWSQST---VHCCDVLNLTMTWSNYQFRSECF	F	1170
Query_10009	1211	EgdDTTTL-levEVGKQLQAVIDAVRVRIERRAEL-LNRTNNNRREHKD-----LADFVGVVFGVGRMERVRVFTSI	I	1280
Query_10010	1172	N--DQSA-T--GTROQATSAIRHLKIIIVKSNLFLNTPRRKFP-NLNRSD---LSSLTEWLFIESTRKKEQDYSSMCR	R	1242
Query_10011	1132	L--DDEQ-I--GTMKQMNKTVELLQRLVNTITL-LEKKRKRDP-----fLALDDFLQWVFEQTRLLHGRSRTCKM	C	1197
Query_10012	1077	E--TDSI-L--GGQEQVTVAVNHIERVLEKKAAL-FNKEEDNIRKPPPEFGIL---LKDVAWSLLKCGTVQVCECRHK	M	1147
Query_10013	1112	R--DDPA-M--GTATHTERVLTNLRRIILKKAADK-FRRTSRVRRIPPGWETCT---LAEMTRWSLTHACDLHAARHQAM	C	1182
Query_10014	1130	N--DDKM-F--GTQEQCKLGLDHVERIIFKSDI-LNQIDPKRVKPPGWSEPV---LEVAVRWLMRQCGRPIETECRHK	M	1200
Query_10015	1100	V--DDPA-L--GTHEQKTSAINHIERIRVKSAL-FQKSGQRRLPPGFGDGT---LVDVLEKWLFEQCGRPKALVRHTC	M	1170
Query_10016	1116	Y--DDKI-L--GTQKCKQALNHLERIMEVKAL-LSKTSKRRPLRNINDSI---LPEVVEWILKNCGCLQTECRHK	M	1186
Query_10017	1122	K--DAQS-I--GTVQCSLSLDHMSRIKAKASTLSKANKNRVPRGLKSAD---SILTAWLLAQCSIETEARKMCR	C	1193
Query_10018	1101	T--DEKS-L--GTQEQLSKALAHIERIKAKAST-LNKEKNRRTPTGFMTKT-vTLHHVVNWLVRCGRPQTECRHQC	M	1173
Query_10019	1132	G--DEEA-R--GTQEQCVQVIHLEIRIHSRAQL-LVKENHRRRIPRGFASQSSiTLIHMIPLWLVKQCGRPSSSECRHQC	M	1205
Query_10020	1083	Y--QIGGD--VAREIMLNAVMTERGICRPHLFLSDSSTYSIDNDY-----TLHQMTQWLFQQTGARBEQFRRLCR	R	1152
Query_10021	1165	A--DDAS-L--GTVQCNLALDHLERIMVRAAT-LKVESQGRRTPLGSKWTS-aTLEAVRWLVQCGQPEACRAAM	A	1237
Query_10022	1102	F--SQEV-----IDNCIMMLEKVKVQVKRISi-LMNVNPKRSIHQSVFTFI-----LYLFEKFVAIETVCRRES	M	1164
Query_10001	1183	RLVCQ LAPCVHEIN GARE--LFSVLLSQ-gNSEYFIQR---FFGMYSSA--RMGLNQK PTMSRVQENF-	1243	
Query_10002	1212	ILYIK ITPLLPGHN TPQT-WMQNTLKEK--GPDYFVSR---FEGGVKAA--SRSGISKY PTLSTMGPFF-	1273	
Query_10003	1212	LLVCS FLPLPRKP SPQA-WMKERLKEK--GPDFFITR---FEGGGKGTG-hKSGIKKH PKISTMDSTF-	1273	
Query_10004	1189	ELFYK FVPLPGNR SPNL-WLKDVLKEE--GVSPFLINT---FEGGGCGQP--SGILAQ PTLLYLRGPF-	1248	
Query_10005	1193	VLFFE LVPLPGNK TPSC-WLADVLKER--GVSPFLIGK---FESAGSDHA-kCSGIYQ PTLKDLQEPF-	1254	
Query_10006	1175	KLVDN LCT----- HVQgnyFVSDFLDIdvDKSTIiKK--vFEGGATATH-gKSGISAC STLKISDTF-	1234	
Query_10007	1192	MLVYK LAPCVPEFQ SVQG-FFQKFLKMEsgGPYFMERfeaYRLGTT-----RFTLMKC HTLGDIGEQF-	1254	
Query_10008	1171	ISYQE LTD-----A EIEKlHFLRHIGTS---IFGFIH---FERSHRSVL-gSFSICNL[2]---RVIKGTS-	1226	
Query_10009	1281	ALFGR VVQLLPQK[4]----WIHQVQVKH--GLAALIKMaepQTASAMP-kLDGNHR[2]-----EF-	1339	
Query_10010	1243	ELFGE FVKLLPDFK[4]----WVKARIDRD---PKFLES1--YKNSLSFSPS----NVLHK[2]----AAQTK-	1298	
Query_10011	1198	MLFCS ICLSIPEGCS --SahwFRANICSG---DQVSSVtkqFDVGMYPH--HPSLQOE[2]-----	1250	
Query_10012	1148	ELVYV LAPLLPVCK SGSE-FVSWYAATH--SWTSIIRIc---EGSKESG----GITV-[1]PTLKHLRNEV-	1205	
Query_10013	1183	AFPLH LAPLDTPDE TGKP-YATAHAWAafGAPALVALv---EQGSAPQP--LLGVSVK[1]--LEHSRAS-	1244	
Query_10014	1201	ELSHK LAPNISGVK DIKD-YFNIRLVNE--SEMYFLAR---FEGSTLKEKilKTSLSY QSFLLDTGPKF	1264	
Query_10015	1171	ELLAS LAPQVKGCA SVRD-FIEKYVGIS--KLSIIVDil---EGGGEAKM----GISYV[4]ASSKHTGSSL	1233	
Query_10016	1187	ELVFO LAPCLSGFN GCKD-YFSHLLETk--GKKYFLNC---FESEMS-----NTV PSFNQNDGLK-	1240	
Query_10017	1194	KLLVQ LAPPSSNVV NIRM-WIRKTIeSE--GGNYFVTR---FESHG--SGILKA PTLAPAP----	1247	
Query_10018	1174	QLVYG LAPQLSGVS TIQD-WVKTIIGSS--SVTFFITR---FEGGGS--K-y--GIGRY QTLADISDTF-	1231	
Query_10019	1206	RLISK LAPLLPGVN SAAD-WMKKT-HED--HPSSFVHR---FEGGGGGGQ-k--GIEQY PTLAHIGGHF-	1264	
Query_10020	1153	RIFLS FSALISTST[2]----WM-----hnYETVQERKS-----ILAV[2]--PLSALHTV-	1196	
Query_10021	1238	RLLVN[18]LLTSSPEKR[5]SPRQ-FFQALLKQF--KANYFYDRwerFOPHQQQQ-----H[2]-----GPL-	1313	
Query_10022	1165	KLWEA IVRCIPPTN[5]DNPVkwIFSYYLQR--rGEKAIKRRfakISFIEQNYDdqNDQLNHR[2]--SEAVRKT-	1237	

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	1244	SLNAACTWFDLVLATLDCYTWVFGERLMSPNQILIA--SNNKGTIFQV--LHFMELAQED	ISAAVKLPFKVAD	1315
Query_10002	1274	SVKLACSWFDHLLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	ITGAARGIEYGGG	1345
Query_10003	1274	TVKTAVSWFDLHLLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	IDAARCFESHGA	1345
Query_10004	1249	SLQATLQWLDLHLLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	IIAAEKCFGTG--	1316
Query_10005	1255	SVRAALQWMDLHLLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	ITAABECPFTLG--	1322
Query_10006	1235	TLTSTMNWFDLHLLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	IDEVSKLF--KESS	1305
Query_10007	1255	TVKNAVWLDLHLLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	IEAADLFTKGLD	1324
Query_10008	1227	--IYALVWLELTYEAMIDGYVWALNTRNRSNEAWNyDl---NEDPRVKSL---LFPCLDLVN	SIVIGRHPFNLSK	1295
Query_10009	1340	-----WMRRLQASLAVYIWFYFGEFTDPAVLSKdGs-----SRIISN-aFYVFLSEYALVD [1]-----SH	1329	
Query_10010	1299	-----DWCRRFLSVLANYKWLKKEKFPVPLDLITsSg-----IMFRK-aINyFLENLVQND [6]PDAMEIDDGEYFI	1367	
Query_10011	1251	-----WLTkFVRSIHWFEWALKENIISIKELLDfEr-----DCVIHIAssISAFLOSLKLSQ [1]-----	1303	
Query_10012	1206	ALVYVVRWLRMVLTSFDYVWLFGEELVKPSVLWGtEgLDNTSETRIFEA--AVYFIQHVCCHR	ASSVTCLLSTDTK	1280
Query_10013	1245	-----HEWLERLQATLEAYQVLVLSHRLLTFRALLVdC-----EMLLWGA--LAHFLST-----	-----FPRRSE	1296
Query_10014	1265	QFGLLVVWLVKLLNALDLCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	--SFLQNFLSHNY	1329
Query_10015	1234	KDINI--FLEALASLDGCVWLIQGQVPAAKLL-----NHSELQLFSA--TTYFINEMALLT	VSEY--MKNHDAF	1298
Query_10016	1241	---PIINWLEGLHTMLDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	FDQRVQAFTHES	1308
Query_10017	1248	NINIAINWLVKLLNALDLCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	IEGVVNSPEVAMA	1317
Query_10018	1232	NLRSMKNWYDYFLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	LDETVKLLYS---	1301
Query_10019	1265	SLKETILWFDFVLAALDCYTWVFGKLLSPSVYTT--TARRPSVLFQSS--LQYFLSSLSLLD	ITEATSCTSPGSS	1338
Query_10020	1197	---ITVDWLEQLSACIENYVWVEMVLGDKTRGfEiCgSgHKKKSVKRRH---TAEVSDTNNEN [4]LSWALLTFLNCDG	1271	
Query_10021	1314	-----RALLACLECHLWLSGGGLLDPQLLgCgG-----GSRLCQR--LSDFFNETDPSLD [5]LSASVASSASSYN	1380	
Query_10022	1238	--LSEKQKTIEMLTAQIECFNLLSRKIASFQDKSkLv-----dINVMGNLYSFL	ADHIVIKFVKNER	1300
Query_10001	1316	GSS---IFTPREREDYNTKCTVIVRVLTSFICTMFGNYKTESL--KVIP-TSFWSSNLWD	TLLAVLQPTQLGFNM	1385
Query_10002	1346	KTSAFhgVFTPRESEENYQEKCTVIVRVLWDFLTVLWLNENYSDAV--KVIP-KSIWQDLSFD	LVVTCILQPTVVGFM	1419
Query_10003	1346	AT----LFTPRETEDYNYGKCTVIVRIMDFFTVLGAFPODAY--KVVV-DAVWKSPLYD	MVVTAALEPTALGFNM	1414
Query_10004	1317	--AAG--nRTSPQEGERYNYSKCTVVVRIMEFTTLLN--TSPEGW--KLLK-KDLCNTHLMR	VLVQTLCEPASIGFNI	1386
Query_10005	1323	--VKG--vLFSPPREEREYNYCKCTIIVRIMEFACLLLEGLQDQCW--KVLE-KDLNLSDLFE	LIAMVVIDPITSIGFNI	1393
Query_10006	1306	VE----YFTPREIEDFNRNKCTVIVRITMFKFCSQVLSNAQSGYAE---IE--DLALRTIFM [9]LLCLVLLKPRQLGLVI	LLIKCVVQFQSVGFNM	1374
Query_10007	1325	RK----IFTPEWEVEENYSKCTVIVRHLDFLCSITKSGSHLE--QVVP-SNIWCPDLWI	LICDCVLPNSALGFNM	1393
Query_10008	1296	AQRNR---LTPREIDQFMTLKCOLEHVYALFMSMLEKYNYSQV---ME---TGGQL [2]VVGVSLLTFAVVGVRT	AFLLKACDPKSLGFRL	1358
Query_10009	1393	PAMAS---LTPREIDQFMTLKCOLEHVYALFMSMLEKYNYSQV---ME---TGGQL [2]VVGVSLLTFAVVGVRT	1459	
Query_10010	1368	KNVGSenILTPAQKQLWLRKIQIIEHLIGFINVLFKFTDPNNV---E--SLWNSNING [5]LLGNCLLRPNIIIGLDVI	1442	
Query_10011	1304	--SIHQfLDLTPGEGQDELQKLRYEATTHLMFKFCSQVLSNAQSGYAE---IE--DLALRTIFM [9]LLCLVLLKPRQLGLVI	1382	
Query_10012	1281	DVLEI---TPWEAESFDRTKCFVTRIRLLDMISTVLSKSGE--AG--CLP-HSLLNLQQLQ	LIVACVLPNSALGFDL	1347
Query_10013	1297	LKRHD---TPAEAAAYRYLKCTVMVRVLDVFRVTLR---VDAS--ALP-TKLWSGECLD	FICKAVLMPIMLDPDF	1362
Query_10014	1330	TEAHI---LFPSPNEIDEYNGLKCTVMIRLMDFLETIFANYSKESL--DVIP-MSLWNSNFFG	LVINMTLNPSTLGFSL	1401
Query_10015	1299	QNAVA---TPYEETAAKLKCTVIVRLLDFVSVVLQEGPKVAS---LLSghSVWSATMFS	LLTCCTLDPASVGFDI	1368
Query_10016	1309	VKKIQI---FSPKEKDEYLRNLCFTIIVRLINFTVTVENYHLQ-L--NTIP-STFWSINFNW	CYSSVNLNPKFLGFDL	1378
Query_10017	1318	TSIH-----ERGEYQRNKWCTIEELLFLFIELLTKHTSETVsfEYVP-DSLWSQELYT	IVMDSLIFYSSLVNTAP	1385
Query_10018	1302	KSTL--gLFTPREIEEYNMCKCTIIVRLINFLVLLQMSPTSII--EGIP-RQIWSQNLFE	MIFLCILTPHQGFNV	1373
Query_10019	1339	RGMG--dIFTPEKEIDDYFRACKCTVIVRVLNFTLILQDFPQETL--KTLP-PEFWSDELFE	VVLSVLEPVGVGFMH	1410
Query_10020	1272	SRNYR---SKRSRTEAYLSVLSLNCNFKVAKIMSHGNILLR--ETLD--LDNQLSRD [5]--LVEVLFHRADLQLHN	1341	
Query_10021	1381	SISEI-----RASAVTIRCOLIVRTAALLQLLAGRQPDCCLK--DLLQ-GKENSNGLWT	ALARATLRFQLGFDA	1446
Query_10022	1301	KSLDK--LILCKEKDKELTKSRMVLITQFLRYLDQDDEESLLAHA--DLFQ--KNCNFNLIH [4]IMVIPLEHLQQLVLED	1376	
Query_10001	1386	GQ VEV--MNNLPEQTLGTL-QILMRSLP--RNLAQELIESVSRHLN-MSS--CDLSSKLPIPLEPNQD---YIQL	1450	
Query_10002	1420	AD VEI--MNNLPKTRKLL-RLMSTHLP--ANLRNLSKAVTRSVT-KSS--HDDLTKLLQSMFGRDEGLTtdHGKL	1488	
Query_10003	1415	GD VEI--MEKLPQETKQLL--LLMCRSLP--SQQALLKSQVAAKLS--RRS--SHDLFSLHPEVSAGVAEL--dhVRL	1481	
Query_10004	1387	GD VQV--MAHLP---DVCvNLMKALKM--SPYKIDILETHLREKIT-AQS---IEELCAVNLGPDPA-QV-dRSRL	1448	
Query_10005	1394	AD VQV--MKNLP---NVCvKLLNAVVK--SFFKETLVDSVRKRIIT-AQS---IEALCNVDLYSPGA-RT-dHAGL	1455	
Query_10006	1375	GD VEV--MNNLPSEV-----PTKLMFVGITNFSDL-----FSLLTNLS--LNNPVG---HTVL	1421	
Query_10007	1394	GD VEV--MMKLPPEQEQVL-RIFSRYP--PKALDMFRRLTKSKLD--SGS--YNLLKLLPHILSGPSED---HMTD	1458	
Query_10008	1359	YS KQF--FECISQDVRLLF--LALKNN--TFLVQVSTSMKEYLT-QFY---PWPVDESVVTCDSHLAI-----EQ	1419	
Query_10009	1460	ID [11] KQV---NELP--LNLIVGRVQCLIVNggDKYRDTLIGYLGEMIS--SSS--CNLGEIGVLTITTEGTRDLV-----	1533	
Query_10010	1443	LP DNIVk-QKLIQVETLkLkFKKLSInn-----VTQLAQDLA-EIAfSDEVDLLNIELDKSDPTY-----	1502	
Query_10011	1383	SE LQRa--QDLAKLASKVLTFRFTQLSSN---LKSEMVELLRRVIEgDDPaYDLNLSFSSKISGPKAL--sLEDA	1451	
Query_10012	1348	YN TEV--TDGLPKNLIQLL-GVLQETLP--SSQLEELNSELGKAVS-SMQ--TDVLSNLQDLSLRNDK-VTL---EQ	1411	
Query_10013	1363	SD PEI--EEKLPKKTGVVL--ASLCLS--GEPRTNLVAAMRTTLA-SDE--CDLLRMVRDVLPGRDIPAP-----LI	1425	
Query_10014	1402	KN LEI--FSNLSPKIRQFY-KLVIQNLN--TDQKQFRQCCQKIIL-DSR--EL-NLDFLVL--KDIDWI---NL	1462	
Query_10015	1369	KD TEV--SKHLPsrLEEL-GILARRG--SEVIGESDYLEKQLQ-KPE-NDIIKMLPLSSETAHINL---RS	1433	
Query_10016	1379	TD TYI--FNQFPKERTTLL-KVFEYMP--KPLYNNFNEHLYEYLE-SKC--NIYNLFPNLNDKDKLIL---EI	1442	
Query_10017	1386	DD PST--AKRISAAVACL-TAMCTSLP--IDVKQNLIAIASNIN-TGQtNNVINLFPVSYNESTSSG--ILLF	1453	
Query_10018	1374	GD IEV--LEKLPQMKDVC-GLFQMKLP--KDKRELMIITLKELLS--RNRsYNLFDLLPDRDRDNELNI-dFERL	1442	
Query_10019	1411	GD VEV--VEKLPQOTGVLC-KLLVQRSL--GECLERLKGTLEKIIA-LGShCNLFQDLPPLSNPQEHQV-dCLRL	1479	
Query_10020	1342	AS SFRa-IEEFCSYVTMHS-----P-----QWAhQMOKMID-----	1371	
Query_10021	1447	QD PSL--MAKLLDCLSTAWhAVARNLAPvgSSYRSSIVSAIAAELELAGQsvTVQRLLPQSGLASDPPSV---AGA	1517	
Query_10022	1377	QE LER---QLRNQSELLLrKLVAIKNL-----DLAKIKKDIvPSMrc-INLNHMTVDNVQNYHPKML-----	1435	

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	1451	RLLVSGFEQLNSAG--LLSNTLV--SKS--SEKLAEDLVNTVYFMGMVTAVD-[	2]	P--TLVNLSPVQSQAELKMLTLAFQMG	1522		
Query_10002	1489	IHLVKGYEQLHQAN--LLSAAAT--GHGS--GSSFAVKLFITVFDLSLTTSDA[	2]	L--LQANLTPLSVELAESLLDLAFLLG	1562		
Query_10003	1482	TLMVSGYQQLHNAQ--LLLPAAL--GTNS--GKSLASELVSVAVFKSMVTRKD-[	2]	P--TAASLPPSFHQTLGGLQLACQLH	1553		
Query_10004	1449	AAVVSACKQHLHRAG--LLHNILPQSQSTDLHHSVGTLELLSLVYKGIAPGDER[	1]	C--LPS--LDLSCKQLASGLLELAFAG	1522		
Query_10005	1456	RSILSACQQLHKAD--LLNSAIDCEAGS1SSSLSKLLSVVYKGIAPGDER[	1]	S--LPS--IDVSSKQLADGLQLSFLG	1529		
Query_10006	1422	OHLVSGYILNKTG--LLVNT-----NVQQLVFDTVYKVCYIPQFF[20]	LkKIITLTPAASILGKSLDLDFSKLG	1506			
Query_10007	1459	SHLVAGYQLLHKTD--FLLPALKDGGHS--AENVSNNLQTLFQAVSEEDQF	--RLECSPTLELSRKLLELACSMT	1530			
Query_10008	1420	QQLNLVYGMLEVEG--LQSVVELPRAE-----PLWTTVADVYFSKRIG	-----LP	SKLGMVQKIFRVAALTG	1480		
Query_10009	1534	--SLINGNYILKQGS--ILDVTLKFAKLS-----LPAMTSLFDYLNHEVDN[10]	EILVFIISTGKIQPSKLELYLPPIK	1612			
Query_10010	1503	DQITHGYQLVLSGG--LFPDLLEHEMTS YMNTSSKLYDYILAIWKKFNNL[4]	-----DPLWVELIRKLLS	FALSLS	1573		
Query_10011	1452	LSLVKGYQLSQVK--ILPLVLNKVRA--GKGERLILTVFGI--EBOC	-----P	PSLQPTAEMVNLA	MSLG	1513	
Query_10012	1412	RQLLKGLEILHKCGmLS--HIVKELPNqYGLDGNQIVRHLFSALVEGQGD	MkRrTVVLQPAAFSYAESLLKFAFLLN	1485			
Query_10013	1426	KNVVSgvQQLAAL--LLDECFFTRDA-----FGRELFAALFEYSHTSTPE	-----P	TNLAAGLADLALS	1488		
Query_10014	1463	SQTLAGYEILSDFK-----IF-----KITDMSKFFVFDLSQEEET-	-----S	QTTEAKRKLFLSCL	SIN	1516	
Query_10015	1434	RHVIEGIGILEKSGwLAKCNVVS-----YSV--AKELVDWLCSIFKEQNL	SvSGISPHSLLKFLYTMQLAITLD	1503			
Query_10016	1443	QEMLLGYEQLIENK-----LWPQNCnkSKKLACDMLSSVWESCILETEI[6]	-----L	NPVSLNFGKMLT	VALKLD	1513	
Query_10017	1454	NRLVQGHYIYAESNkCLSDVLKSMGISeEESLAMLRYNALTIIRGSSS[8]	Y-----M	PLHLKGTGELLKGL	LELS	1532	
Query_10018	1443	TFLKGYKQLMTMG--IL--KTPADHKDIE---KIMKSVVDCFRGEMN[1]	D--RTA--LS	PAQQVAAEELLDA	FELN	1511	
Query_10019	1480	AYLVQGYQLHCAG--LLLPAKAQARNaseEYAVKLFYAVYDGVKPSLNV[1]	L--Q	TVILD	PASQHLASKLLD	LAYDGL	1554
Query_10020	1372	RMLKLHQSADVY-----SSES1SDHSTTISLTLFGSKVLSAEL[7]	-----N	YAEFALVARKIVE	YGH--G	1438	
Query_10021	1518	VRLARGVRLRDCG--LLADSTADEKRQ-----ILDRCCQLAES	--aQT	TQHPVLPQLIDAAL	TYALNDS	1579	
Query_10022	1436	MSMCLGYLKLASID-----PNIKEYLIS--FRSLPTIVEA[5]	-----T	PISNESARVLV	KFLSKLE	1492	
Query_10001	1523	YPEKN-LVELIFN----S NNIHVGVSTSVQ--CSPHGTLFYAAFKSVVNLNFVKSVEKFPV-L	LMRHL---	EANGE	1586		
Query_10002	1563	IEAKM-LVSCCLD---K TALFSG---RS--STSKGAVFYTFKTKLNGEMAVQQAQKVIP-L	LVREG---	GKDV	1623		
Query_10003	1554	PQGH-D-LDCCVMN---K TPLMTGGKKS--STFGMVVFTFKKCIQQLIKCAESLP-V	LLQQL---	TAEV	1617		
Query_10004	1523	GLCER-LVSLLLNPAVLS TASLGSSQSGSViHFSHGEYFYSLFSSETINTELLKLNLDLAVL-E	LMQSS---	VDNTK	1591		
Query_10005	1530	GQCEE-LVSLLLNMTMELS VPLSGATQKNPIFPHSGEYFYSLFSSETINAELLKSLKSAV-T-Q	LMRSA---	SENPK	1598		
Query_10006	1507	LKDNV-LITALLN----D DMIQGT--PSV--NLQHGAEFFLMFKSVICQYIAENGRITIP-I	LTNQA---	KTHGR	1569		
Query_10007	1531	DKPAQ--ILGAILS-----KT--KTK--NQMSQNFYAMFQSTINLNYLMKEAKYFIP-E	LVNYA---	SKESS	1587		
Query_10008	1481	SAVDV--LIKLLMNTESVR[1]-----AECTS1CS2TKGELLIANFSDEIALILCTNAELFLK-H	LKEAS---	RCQPD	1544		
Query_10009	1613	VLTkStTTTTTSSSTTD[1]--QLEEEERQKQKQIEHQSNFYKTYHSELIEYISKNFESFSP--[1]	FIDRIv---	QTNHI	1683		
Query_10010	1574	KYNGA-----INQWLLS[2]--IGFPSDNSTFYQNGTICYQCFSSVINEHMSRNFREPAK--[1]	FPDNIndf	TRDQM	1640		
Query_10011	1514	VNATF-LVNLIMDEKSVG[1]-----STNkKLTSGEVFYMNFRTIVHQVKKFYKDCDLR--[1]	ILEAA---	AEQQA	1575		
Query_10012	1486	KEVTT-VICCIKE---Q ASIQNPATHSS--VLHGDYFVFTFRSAILLEHMLACPEHVS-C	LIKDI---	SANNV	1548		
Query_10013	1489	VPKQ--LLAMLRDDTPTP[6]STMLASASATQTTFRGALFYDTRVYRVLAPYIARHFVKVFAE-Q	LLPHV---	GDGR	1563		
Query_10014	1517	---RN-FLTKIL-----QNYLEEQSAN--VTLSYFRAEIFTQCLNFYSDI-----LD-F	LFSKL---	SSKFS	1568		
Query_10015	1504	PKVKH--HLLSWAT---N SDTVSGTSGQE--VEKGSVLSILGPKLIPQFLESPPDALS-R	CVEMI---	REGFR	1566		
Query_10016	1514	ISSED--LLKKLLD---E QQLYVSNLTKT--TRRGLIMFSTYQQEISLAFAGINSKHFTS-S	LMNNI---	NKSKT	1576		
Query_10017	1533	DSVEP--FLLLVLS-----PAQSGAGASF-AAQDGLLTYITYSGLINNKLLHIXMAID-V	LLKNA---	KTHPQ	1594		
Query_10018	1512	---E-TIDIFFRYLLDE TKLYGDSN-SA-TATLGMTLYYTFNSNINAFIAKVSNRCLP-M	ILLDA---	ASKSF	1574		
Query_10019	1555	---LQFNLYMES--IMEDSDNDP--GSSCGCTFYRIFQTSINGFFIERANIHMA-N	ILMKA---	ASSPF	1612		
Query_10020	1439	SPQDR1VVVAALQAAAAS[4]TDIFLKSADR-----MYAS1YSDV1HMLPTLAVWKRA-S[1]	LVLSL---	LENDF	1506		
Query_10021	1580	ALLQP-LVTLLTETH-----SAGLDDLNSYKPCLTAYFSAQCSSLITL[4]	LLKAA	tapATAAA	1638		
Query_10022	1493	IPNET-----LCEWLTTK[1]-----LLYQKSSDLLSYIAENWNHFKL-N[3]	YLIK---	AKDEP	1543		
Query_10001	1587	T LSSVILGIAESVSDRPLRKS HG KMVS---RAIVSQWSSLSV--WWSEstASLDTKMVAITLLTKLMQIDST	1654				
Query_10002	1624	V VSSVILNGLLDYLTSSREVRRTHG MSVH---SAITQWDILMG--GLKKG-CSTDLQSFALLMLRKLQIDAK	1690				
Query_10003	1618	W VGSILHGLLDHLAKDKTLRKKEG AALC---DSLQSWPKLAP--WWEAG-ALQDSKSTALTLLRKLIIIDSQ	1684				
Query_10004	1592	M VSAVILNGLMDQSFREANRKHQG LKLA---TTLIQHWKCKDS--WWAKD-SPLETKMVAIALLAKILQIDSS	1658				
Query_10005	1599	M VSTVILNGLMDQSFDRDHTVRKTQG VLVV---SAVLTNWRTRLES--WVSEN-SAESKMAVLTLLSKVLQIDSS	1665				
Query_10006	1570	K VNSILGATVDYIARDRTLRRKKEG PNVA---STILSEWKKFST--WWTNN-APPELQNLASLLTKVLLIDSR	1636				
Query_10007	1588	V VSSVLSLLEQISDRDQLRGGEG KTVV---DVVLSQWSQLQLqgWWAPD-SPQEVKMSALNLLRKMMLINGK	1656				
Query_10008	1545	V VGRALPHTVSVVVHSSDLRKKYG[4]LQILqifGQELLYWADVVS---NET-----LKMLWLKMLKLIIDSK	1613				
Query_10009	1684	D[1]ILNDICIKYQTIQFNNTDKQSK[4]KSMI-----IPFLIKISSh--LLKDVGKTTDKENIIEFTKLNLIIDTP[3]	1756				
Query_10010	1641	N[3]IKDLILNLDYLSIKKDTQQSVE[2]IHLNs-ePYFLKSPVKICTy---D-----RHYLLKLNGLFMLEPN	1706				
Query_10011	1576	A[2]---LLIALLDSYLPSEMGA SIT[1]SSFM--QEFLOHIEVLAP1--CRAG-SSLTLRFLFDILHKLMLDTPG[4]	1646				
Query_10012	1549	Q[1]VISTVTEILQYAKIRKSKHQDS TTLCaferSIVNHWNKLSF---VEG--NSEQVQLLGLVNLARIMPQ	1617				
Query_10013	1564	Y -RLVHDAARALVTGDTEMRRYAD -EFF---TAFIDGFLHVRT--WINSAGAGSETHVQGMRLVSAVVRCGVR	1629				
Query_10014	1569	K ILQFLNFAVTEIAMDKHLRKTNG QKLI---LKVFESEWPKFEP--YW--S-FESQFKILMISLLTKLMIESP	1633				
Query_10015	1567	L[1]AMGFASALFTCLIRNADIRKKYS KVVV---QSFLKHWPTLAAq---AGK--TSLHQDSVLSLSSLF-LIDR	1631				
Query_10016	1577	T VLKILVGLIDSIGKVKFRYRYG KDII---NSIMCQWSRLQQ--CLEES-CDIDLQNDIITLTKMFLIDSS	1643				
Query_10017	1595	Q VSVILNLLLEVYQKDFKINNNN[5]ETLC---TELISHWQDLSL--WWEDG--SKDLKSAAVTLQKTIALKPK	1665				
Query_10018	1575	M ISGILSGLLDFILDRGLRQSKG IHIS---KKILSSWNSLKC--WWLTP-DAQDQLQISVNIILNCLTLDPQ	1641				
Query_10019	1613	L VLSLLNGLLDQVIRDKILRTRNG SSFV---AVLNNWTSFEI--FWRTH-SAMDLRHSILMILKMLILDP	1679				
Query_10020	1507	V VNVLIDILRQVACSKVYLRKSAE[5]ATLvaavKQYLNKLETVKRe-----TQPALSLQILKHFVLELSON[4]	1579				
Query_10021	1639	T[6]LAAIIBSLAASVKRTALS KQCCA[4]KQLA---SQILAKWTLVQQ1--WDYG--NDDCRFSVLGILLTNLAVDAE[3]	1717				
Query_10022	1544	S LYAVFIPILKLS-----QKTPN[2]KILL---RSVPENWNEQERtdFTEFE--VATSKVKIYTFIANACQVQLR[2]	1610				

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	1655	VIIDVKNKSFYY--VFKMFLGLISDKNNTLQFKTRVLDL	LIFFCNLPA-EE-NKQLESCLNQFVADCFPMSSLEFHA	1727
Query_10002	1691	FVSDTTHPVFSS--VLEAYSFMLTDASTSLTFKAQVMDV	LFFFVSLPD-GE-QTKLRSALDRLVADNFPNLSSEFRV	1763
Query_10003	1685	VSRDPAPAFQFP--VFTMYLDMLTDSRSTLPPFKSQVLDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1757
Query_10004	1659	VSFNTSHGSPFE--VFTTYISLLADTKLDLHLKQAVTL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1731
Query_10005	1666	VSFSSNHEAFAM--VFHTYTTILTDSTLALNLKSKQAMII	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1738
Query_10006	1637	VTSHTSHSPFST--IFEMYQTMVLDNKTNLAFKFKVLDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1709
Query_10007	1657	VVSDPTHSPFSC--VFSMYQSMVTDKSSLLTFKCKVLDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1729
Query_10008	1614	I---TEWPEFAP--MASAYFMLNDQSLTLDPFKNRMLSL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1682
Query_10009	1757	----ETKEGYDF--VYQLLINYL--NRSSMLNPFKQAVTL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1825
Query_10010	1707	LLRRFSLLEEFQN--LFRENYEAFLDKQIDLEFKNEALDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1778
Query_10011	1647	LNKYSFDFITetFLSFLSCSSEDATNSRLGSSAAFTL[7]	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1730
Query_10012	1618	PDRGLARCEA----LQKWVITQLTSRDLQHLHYKSRALDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1688
Query_10013	1630	PETAQRHANFEQ--VLEVYKRFLEDRVDMTSKRVALQL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1702
Query_10014	1634	QPK-----NEK--VFQMYFEFLTDEKSMNPFKAKVLDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1694
Query_10015	1632	BGTIKSHCTSSRSdMVVFNGLLADAKVELQAKIKVLLK	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1706
Query_10016	1644	FIKGNVNYDSNS--IFFSFITLLQKSEMSLEFKICVLDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1714
Query_10017	1666	LLLKSADVSKPL--V--DMYTAMIGDEKLELNFKAVMIDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1738
Query_10018	1642	FATQSNDSSEFET--VFNFFLSLLSGTSTSLAFKVVLDL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1714
Query_10019	1680	KVTDENRFTFSS--LFEMYLKDKKSTTLAFKQGVLEL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1752
Query_10020	1580	YVEALRLGPILD--IQNACSLLLKQRCGSYLVKADLLQI[5]	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1658
Query_10021	1718	-----SndqVFVLFVGVV--NSRVNFDLIVNGLAL	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1782
Query_10022	1611	---SNHYAIKFIh-LCEKVSXYLKER--YPVHIKREVLNL[6]	LFFFVTLVQ-TQ-LAELKARLDRMVDNFPNLSAEFPA	1687
Query_10001	1728	-GSQQQKCYITALDRLLLAMELSGSLMLLNLIIISVICBEVK	HTHENAIQSAFQRFIKRLLPKD--KHKLALDVPYKIFSQ	1803
Query_10002	1764	-GSPKYIDYAGALNKILTAMVLSGNLMLLEIIVISIMCRDEK	HAYEDEIQESLAAFIKRLPEE--KQTAAGVAVYQIFCQ	1839
Query_10003	1758	-GSSKYNDYVGAATKLLSALELSSMLLELLISVMCREAK	HVMEDDIQTSLSRFIKGMGAD--RQKPAIDIPYGIFFR	1833
Query_10004	1732	-GTPRFNNYVDCMKKFLDALELSQSPMLLEMTVEVLCREQQ	HVMEELFQSSFRRIARRGSCV--TQVGLLESVYEMFRK	1807
Query_10005	1739	-GTLKYNVYDAIKKLLDALELSQSPMLLKMALTLICRDRS	HVMEEGFQVCFQRLAKTGSCV--KQVALLAEVYGMFEN	1814
Query_10006	1710	-GSPRYQDYISAINKILLALELSGCSMLLELLISIMCREKN	HVYEDIEQDVIARYTKRLVIY--KQKALADIPYKIFTS	1785
Query_10007	1730	-GSPKYQEVYIAMSLLSALELSGSLMLLELLISIMCREKN	HVYEDRMQESLVTFIRRLPTG--HQKAAALDVPVPIFTK	1805
Query_10008	1683	-GTAKYDQHVAAQCILFAFGCTGSPILLELLISVMCREQS	HPIASDLKHQKQFSSRANSAV--LENILELLDKMFFN	1758
Query_10009	1826	-GSPYINEYITTIERFLNTEITCNPMIIDLTLQVL-KEQD	HSHITHINQSIERFIVRTNDSg--AKEVFSHCAMPMS	1900
Query_10010	1779	fGTKKLENYLSVLDKMFAMVRSVSLFKILIPPTFCRES	DMYGEFQRFQVSEFLSKLSRS--KYVDTMDIAYGYFTN	1855
Query_10011	1731	-GSTQRSNYQTLNQLLAAVSVSHSLELLELVIPVQPNK	---ISSRAVSEAVEVFAESLEDNc---EAAFNLCLTMDM	1804
Query_10012	1689	-GSVGYSGFVNAFRKILRGLBESGCVILLVSVVMAAESQ	HVLEDDIQKSLQIFIKRLEDD--QVQKAEAVPKVFDV	1764
Query_10013	1703	-DSTQHREYIAALDDILDALVLSGSELELLEVVTVHQREDR	tHVHEDHIQRKLRDFIVQEHDTa---PQAIELMFPQIVLA	1778
Query_10014	1695	-SDDIYKDYLVNVHKLVCLELSQNLDMLSLLIGVCREKE	HLCEENQFRESLERFIKRLSSE--KQMEIVKQNWSNFQS	1770
Query_10015	1707	-GGARDREYHQAICEILLTALELSLSPILLQFVNVNFCREDK	HRYEVMHESLEKFMKLLPLL--KQOETLKHIDMYND	1782
Query_10016	1715	-GSSQYNDYHLAFTKLLIALELTGNEMLLKVIFPLVQPNK	---LLEENIQKCLKKCRKRTDSN--KERLIAEVPKMPLE	1790
Query_10017	1739	-GGPLLENYTNIIEKLCNSLVASGSLVLELLIINIMCREVR	HVCEEKIQTSLHQFISKLSSEd---VDALKEVYEMFKN	1813
Query_10018	1715	-GSPDYNDYVACIDKLLANLIKSGNMLLELLISVMCREKK	HAYEYEQIKSLLFISSLIPK--DALGALDVPVKMFDV	1790
Query_10019	1753	-GGPQYNDYVQTDKLLAAFNRSRLLLEVLVIAIFCERERQ	HVHEDQIQSFIAFIQRLHPEpdQAAMDVCLSIFFN	1830
Query_10020	1659	-GSRDFDVHMLFTVVLHVVEQKSIAYLKVYIPSLKEAES	HLFRVEIKQMLARFSLSSGvLNGSVASSEMIRQRM	1736
Query_10021	1783	INSHAYCQYVRAFRLGLLAQETSSEHPLLLEKSLIDVCREK	HALAEAFQCCNRLRLRIPIQ--RCRDFVTPWVRLFRG	1859
Query_10022	1688	NSNVYNNFHIIIEGFLSLLLSKLNKLVLRMIYNSMRWKT	---TYEPLSKRTKIIEIQSINKm---PLHQFKQTFQFMD	1763
Query_10001	1804	-E---G RFTNNOARRACEKVSLSLMLRVHSSALTEFYC	EHIKEIMKIE-AKPTKAPHAPE--AQLTSKIGCF	1870
Query_10002	1840	-E---N SFSKEIRRAAIEVRCIPLMLRAGVKALSSFFK	ECIKNIGGIIID-AKTIKAPETVFQ--NQLVSKTCCF	1906
Query_10003	1834	-E---G SFPNEIRRSALERVCLPMLRLVHSPALTEFFFT	DHIREIMDIIIG-APQTKAGEASLR--TQITNKLCCF	1900
Query_10004	1808	-D---D PRLSFTQRSFVDRSLTLLWHCSLDALREFFS	TIVVDAIDVLK-SRFTKLNSTFD--TQITKMGYY	1874
Query_10005	1815	-E---G GLSNIARQSVMDRSLTLTIHCSLDALKEFFIS	KIIKDAMGILQ-SRFTKTSSEFFD--NQLTRKIGCY	1881
Query_10006	1786	-S---G SYPSAIRIAALSRCVCLPMLRLVHKSAVIEFFL	THISFIMSEIE-AKLVKGSKAELE--NQLTKIGCF	1852
Query_10007	1806	-E---G LFPNEIRRSLSRVCLPMLRLVHKSQSLSEFFV	DHIKFTDTLS-AKLSKSSSEIE--RLTLKLCFF	1872
Query_10008	1759	pS---G CVCASLRLLALNEFFIPFLHSKNDLLLTWTC	RKIQTLSVCLT-ESLTLIEQTDCE--QQLITKIGCY	1826
Query_10009	1901	nd--- YQDELKITLIDKFCVPLIGHMKQVLIQIFA	QHLSSLMSIIQ-PLAPKYLATAIErkSSVVEKICCF	1968
Query_10010	1856	-StyvS ----FORIIEMILSPPLLQGPRIKVFMDFFK	KHINEIIDLLE-RDKYTRNVEKII--YILNERAGSL	1920
Query_10011	1805	----rG[1]--SYELRRAVVEVIFPAFVNSSSMEFVTVWYS	QHISDLFYVLKqSSSMLSABREFE--DLFKKICHY	1870
Query_10012	1765	-D---Q YDPSVRLGCAQRFVRLPLQGCRRSAAREFFR	HCIQDLQGMNLN-VTLDCRQQLQTE--HQLVGRIGAW	1830
Query_10013	1779	-E---T N-PLFVRRGVAQRVLAFLQKASLKITKDFYI	KRLPAM-SLM-SQDTPLTSTP----AIVIKMCTF	1840
Query_10014	1771	-S---D ---DERKFLLEKVFIFLKNCKDPVIFDFMS	ANIVSLIEVLN-WNL---NEPAFE--MMCLNKKCVF	1831
Query_10015	1783	-P---T CIKQNLRYNIMDLVLTPLLQKAEAVPTQNFLL	TVISAVMAGMS-APVHGRSDEQHK--TLVT-KVGSF	1848
Query_10016	1791	-D---E LYPSDMRIHIEKVFVILGMISKITLKYFFI	DHIKEIMQIVK-SKR---KSSN-E--IQLVSKIGAF	1853
Query_10017	1814	-D---K DFPVAAVKGIGERLVCPLTECCSALLLEFFV	LKISDVMCIME-AKLARVTPKLE--CQLVHKSCEF	1880
Query_10018	1791	-D---K SYRPDIRRAAIEVRCVCLPMLRLVHKSQSLSEFFV	SHIKEIMDIE-TKLVT--ADLE--NQLISKLGYS	1855
Query_10019	1831	-E---R DYQTHSRRAVIERICITPFLLSSTARRAFYQ	DQIQKLSVLE-AKELKPTDPNFE--SQVTSKLCFF	1897
Query_10020	1737	-E1ldL[4]NVDVAIRKILLEEVFPFVEQCTGEILLHFYM[4]	KRKSTIISLLA-TLISASAEAFGE--ESRIGTVVAF	1814
Query_10021	1860	-DsggG GFTGFHREQVSRVLAPELLTQCHGSVLLDFVA[2]	DCLEYCLTCLLNAQQSLAGFNESEteEILLARGCC	1934
Query_10022	1764	-Bfmr[4]SVDDNLRWAIMKKIILKILESCSQEFLADIMI	EFTPQLEIVLK-ESIFLQKDPLEQ--YQLIKQKWI	1837

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	1871	ELL-EVMYSRLPNTD VNSKD-SSINKSFCS---PKIPDtGKE-MNLTITKVS NATKGE	NMNGE	TLLIP	LRR	1935						
Query_10002	1907	QLV-ELMYSRLVKEDLF SMD-SAINKAYTM---D-VKT-GKE-MTQALTKACHAAKSE	DIRGE	TNALE	LRR	1969						
Query_10003	1901	QLL-ETMYSRLPKKEVNTPT-SRIVRAYKG---EEVKA-GNE-MTTAITKFGHAAKSE	DLRGE	TGLLD	LRR	1964						
Query_10004	1875	KIL-DMVYSRLPKDDVHAKE-SKINQVFGH---SCITE-GNE-LTKKTLIKLCYDAFTE	NMAGE	NQLE	RRR	1938						
Query_10005	1882	KLL-EVYIIRLPKDDIYSED-SKINRAFHS---SSNME-GNKLSKSLKLCYDAFTE	NMTGE	KQLE	KRR	1946						
Query_10006	1853	QLM-ELLYSRLTKDEVNSKN-SVINEKYCQ---TKVET-GKE-MTQKATKAHEARSE	DTRGE	TELVE	LRR	1916						
Query_10007	1873	ELV-EVYSSFPKKEVNSKN-SKINLKYCG---GKVEY-GKE-MTTAVTKAAHEAKKE	DVRGE	KRLPE	LRR	1936						
Query_10008	1827	KIC-AI IYNELDLECL---K-NEVNVSYLG---SDSQM-KNE-LIKKLVSCAVGAKKE	-FCRD	SKLDS [3]	HFF	1889						
Query_10009	1969	KLT-QQFYESLPATVIRDQV---NPFYTG---DKHADAGTE-LTAAIMKAAHLAKSE	KLATD	EQHVT [2]	LAT	2029						
Query_10010	1921	TLL-QILYLRLASEEDVYSEK-GEIVKAYCQ---SRPLK-AKE-LTRKVMKAAHKIKSR [3]	DDNGN	KLVSE	ANL	1987						
Query_10011	1871	ELL-ELLYSK--ADNV SITK--TIAPVFTS-----SD-----LMRQAGDTCHG	KLDPE	YKLYP [4]	TWR	1925						
Query_10012	1831	NIL-EPLFVHLESSELESKE-CVITAAAPP---GAVKT-GKE-LISDLTSRAYKYRSE [4]	NLTGN	VQLVE	LFR	1898						
Query_10013	1841	TLI-DILFQRLNNTDLYGKS-AVINNKTYCE---AAGEPVNDK--GSNLTNTCRALRD	TVTGE	ESLVE	LRR	1904						
Query_10014	1832	QLI-ELAYKRLLHKDEIFFTN-AKICMAYETIKFGCVKD-GKE-LTKEVLRKSRKHMNE [4]	NTGQE	SAKIE [4]	LRR	1906						
Query_10015	1849	MIL-QIMYSKLPREKLFPAG-SEINAAFRP---GDQ-S-GKE-LTKEIFRYAAKISKG	ELRHD	QTHAE	LRR	1911						
Query_10016	1854	QLI-EILYTYLSKEEVS-KE-SQINCSY-----GKE-LTREATGFAIARKE	NLEGE	SISKE	FCR	1909						
Query_10017	1881	QFI-EILYTRLPSLLISSME-STVNTAYCN---GSPQT-GKE-LTQAITKCHGAKSE	DLHGD	RTHYE	TRR	1944						
Query_10018	1856	KLM-KLMYSKLSKQDVNTAE-SSIVKAYTG---RPTED-GKE-ITKSIKAAHACKSD	DLRGE	TRSV E	LHR	1919						
Query_10019	1898	KLV-EILYSQFSKSELNMLE-STINKAYVG---GDVKT-GKE-LTLAITKAARAASIE	DMRGE	TVALE	LRR	1961						
Query_10020	1815	SLV-EILYRLVDPDIIR---SKINSAFLG---HSGGK-GRE-LTMLVCKCAKVVVTK	---IY	LEADD	LVR	1872						
Query_10021	1935	RVL-SCFYMACGLTELALPRLCDAPKRFYSrqpAAQKpDRS-LSNWSVQCYDLVTO	---SF	PRCRE	SAR	2000						
Query_10022	1838	FLVFEVFRRLKPETIR---QQVHKLYG---ESSpGNE-LTKSLIQLCTLAKKA [6]	NFVKD [9]	NSPLQ	VQR	1913						
Query_10001	1936	ALHCAAYNTLIAALICTQEKKFFYAFLEF TENPAKNQLHWNN-IVDQDVVYEFPLEMENPIERRK-KFVSIRTSVRD---				2010						
Query_10002	1970	QFHCAAFNALIAIISCTQTDIKFYNAFLFSENTAKGQFLLDN-IIDITKNYSFEIEMNAPMERKK-RFVSIRNQIQDtvA				2047						
Query_10003	1965	EFHCAAYNALIAIISCTQTDIMKFFYAFLEFQENTAKGQFLLDN-LVDTNKLYQFDEAFSAPVSRKK-QFVSIRQELRG--Q				2040						
Query_10004	1939	LYHCAAYNCIAISVICVFNELKFFYQGFLESEKPEKNLLIFBN-LIDLKRRYNFPVEVEVPMERKK-KYIEIRKEARE--A				2014						
Query_10005	1947	QYHCAAYNCIMAVVTCIFTEAKFFYQGFLESEKAEKNIFIFEN-LIEAHRTYNFPVEIEVPLEKKR-RFISIRKEARE--A				2022						
Query_10006	1917	QYHCTAYNLMISIICTQTEAKFFYNCLFSEKPSGQFLLDN-ILDCEKMFEPKAE LDSPLERKK-RFVSIRNEVKE--T				1992						
Query_10007	1937	QYHCAAYNALMAIITCTQTEMKFFYGLLFEKDLKAGDFLLDN-IVDQNRKYLFEIETSLPQRKK-RFVSIRSEAGA--Q				2012						
Query_10008	1890	EYRCAAYNCVIALVKKTSQSHLKHGEGFIFND-----CWES-IIDKEVKEYFDREIQEBEQSAG-KMYYSRQDRFS--S				1958						
Query_10009	2033	KYHGAAYSALASVVAATTQTKETFFHVFVFFKENKEKNEYLWEN-IIDLEQQYNF-----EPE-TNFQISSQLSSa-v				2100						
Query_10010	1988	HFRCEAYNALAAVILCTQNNPVFYKAFLEAENTSKGELLWEN-IIDLEIHYKIKIELNQPFPRTK--VDDFRSRSGK--T				2062						
Query_10011	1926	ELHAAAYNCIAAVILCTQIDEKWFNAFLFSEKSRP---LWQH-LVDLTKVYDDMSVETSQSSIIStVKGMRADLKArrT				2001						
Query_10012	1899	QSQAAYNFALVAIISNVKCELRFFTLGLFSEKAK-EVWPWGK-LVNCNKYALALDWEIEPARRK-QLVSIRQCARAq-r				1974						
Query_10013	1905	QASCAAYNALTTVICKTQTEKFFYAFCFKEATGKVRDCVFN-LIDEKRYTYTTELPAAMSVRS-KCVLALREGEA--T				1980						
Query_10014	1907	ILHCSAYNCLISLFI RTQTEPKLYLAFLEFKDDPSKNEFVFEP-LIDKIKSYNFCVEMENFTQRKT-KFISIRDEFKV--I				1982						
Query_10015	1912	LCACAAYNAI IAVLACVQNDLRFYTNFLF SANPTKGEVIWEC-LVDCSRNFEPDVEINFKPASKK-RFVAVRQKLREt-A				1988						
Query_10016	1910	KYCCAAYNAVIAIICCTQDSKFFYTFLEFNFAKGEVIWEN-IIDLNKEYTFPIELDT-IERRR-KFISVIRHENRE--L				1984						
Query_10017	1945	QYHCHAYNTLMSVISCTQSKPQFFVGLFKEDVVKQQLLDN-IIDCEKYEFPQVDFEAPLKRKQ-QITSIRDDLKQ--S				2020						
Query_10018	1920	QCCCAAYNALISVISSTQTEMKFFYVFLF TENVARRKQFLWEN-LIDMERYAFIEI ELSYQDKQKT-TFRSLREQSRS--S				1995						
Query_10019	1962	QFHCAAYNALVALICTQTDLKFYVAFLESENPIKQQLLDN-LVDCNRITYTFEVELSAPLERKK-QF---RMQARI--P				2034						
Query_10020	1873	LACCAAYNLLAAISQTOQEKLYDQILFQSA-----LWSN-LVDSALEYELR-----AETE-NFVTIPLSSLSL--t				1935						
Query_10021	2001	QMQCEAYNLLVHCIIKAASDKERSYLP-PFTEANAARKKFI FDN-IVP TDVDVYQLPLQVEQSASRRKQ-YLVSVRPAASS--A				2075						
Query_10022	1914	QYACAAYTCLTAIIMGTQKENLFSQFLFATLRDKGEECFWNvLIDQEKAYNF--EVETNFNYIQIKQIEFFKLEQAN--G				1989						
Query_10001	2011	--VDPS---DEPTGPA	SSV	H---	YLASQYL	AGSSSLTDV	SQDFD	--GSSMV	----	SSADSSFDNQRSRES	2066	
Query_10002	2048	GSSSQ---QSDSGIE	MSG	Ytgr	YLSTQYL	ANSSLE	DIQQDFD	--SGGVQ [2]	NASIDKDQSL	QRSVPHIK	2114	
Query_10003	2041	DSPEA---GSSR---	---	---	YMQSYL	AGSSLE	DISQDFD	--TSGTQ [5]	TFSSSTQERY	SASQSSRLQ	2099	
Query_10004	2015	ANGDS-----	--D	Gpsy	MSSLYL	ADSTLE	EMSQDFD	--STGVQ	SYSYSQ---	DPRPATGRFR	2067	
Query_10005	2023	GDGDS-----	--G	Epy	LSSQSY	MADSSLE	EMSQDFD	--STGVQ	SFSYNSQ---	DPRGHAVRSR	2075	
Query_10006	1993	RQGE---EESL--F	SS-	Y---	NISTQNI	ESSL	DDLNQYDF	--NSGTA	-----	QVFFSDIFSDLTSS	2045	
Query_10007	2013	GDGNS-----	QDSA--E	SSV	H---	YLASQYL	ADSSLE	EDVSAYDY	--SGSLQ	-----	GERSEPVPPRTRQR	2066
Query_10008	1959	AEYEN-----	---	---	QVELSVLL	DSSLSV	MATQDFD	--TESLE	-----	ARTHCAETR	1998	
Query_10009	2101	ANSQ-----	RDLK	---	YLLSSQYL	QDSSLSQSL	---	DVG--	TSSS	-----	TTSTSTTTTSNNNL	2147
Query_10010	2063	FKEAIv-----	SNNLK	---	---	YLASQYL	ADSSVT	QGGALFKelp	--DES [1]	GESKNEHSDTIPMEIDERVE	2120	
Query_10011	2002	RGKGS1---	SQTAST	---	---	ISSQYMH	SASLSQ	EPAVIRTF	-VGGIR [4]	SAAGVAGLLLEDDEESPSTET	2063	
Query_10012	1975	RETLG---	fASQSVRY	IPS	A-----	SQYLF	DSTLNE	DTLDFD	--STSVV	-----	STEARPGHDE	2024
Query_10013	1981	PSTTA-----	A	ATA	Aarr	YLPQYL	ADSSLD	QDAQLEA1-RQGAD [5]	AVEDASSSSSSPPSDGPTF [5]		2050	
Query_10014	1983	KADTN---	SRSANYI	STM	Hts	SLSL	QNL	YSSSLE	EELSVDFs-VHSSY	NQAYNSQSSVTRA--PSFRR	2046	
Query_10015	1989	REEGA---v	VAGTVQY	MAS	H-----	YLS	DSSLE	EDMSQDFD	--NNSVV	-----	LAMSQTEKSERVT	2040
Query_10016	1985	NSNKTln	fASSLKYA	PSQ	Y-----	LSDSSL	DDVSQDFs	--SSEL	AFSQDMKNKRKEIDSNPS		2043	
Query_10017	2021	TAAGN---	SSGSPSS	SPS	Trpr	YISSQYL	DSSSLTDV	SQYDF	--TLF [5]	GSVGGGASTESSIDEDSGN	2088	
Query_10018	1996	VSSS-----	ADEADSG	IAM	Apsy	HSSQYL	ADSSLE	ESISQDFDn-SSTVQ		GYSENNSIKQIPTRTGDS E	2061	
Query_10019	2035	SPKEN---	SLEESAG	RTG	Lsh	MYSQYL	ADSSLE	DISQCDF	--SSPIQ	VHSGTEA-RQPLYKNESSS	2098	
Query_10020	1936	VSLKQ-----	RY	EVD [7]	Krnns	SAALQFY	TNSSL	MDAVTLTA-----	-----	STVTEPPDINR	1988	
Query_10021	2076	SDSSA	seaGPTDDL [5]	SSS [8]	Nslq	FAASQFL	ADSSSL	DDVSQDFRlvTADLS [1]	TSDSASLATASLSLSTSSSF [3]		2162	
Query_10022	1990	GDISQ---	IKANNL	INE [5]	-----	NSAMV	FETQPK	INI-KSEE [5]	LLNNKKQNTNQDQMLDKSA [1]		2053	

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	2067	SIEVPDVLV	KEHYI	EMEMDALNQHPCCMAALSALLOHMKL	--NQITPEVEKGSK	--PS--SMPDWMKNIHDKIKN	2134
Query_10002	2115	KRYSNEETP	[4]QGBYI	EMESDALNQHECMASITAVLKHMQA	--NKITPEIPEGSK	--AS--EMPGWMSLQKKLTS	2186
Query_10003	2100	ARQQDSS	-[3]HGFDFL	EIEMDELNQHECMSSMSAVLQHMAR	--NNITPQVQKGA	--PK--DMPWMSCLNKKLGG	2169
Query_10004	2068	RREQRD	--P[2]HDDVL	ELEMDELNRHECMAPTALVKHMR	--SLGPPQGEEDSV	--PR--DLPSWMKFLHGKLG	2135
Query_10005	2076	RKEQAE	-LS[2]QSDIM	ELEMDELNQHECMAMIALINHMQR	--NNITPQVEEGKT	--PQ--ELPPWMKFLHGKLG	2144
Query_10006	2046	SNVQPDQFN	[1]TVDYI	ELELDELNQHECMATMVALIKHLHT	--VITPVL--KGTI	--PK--EMPTWMSYFSEKMTN	2112
Query_10007	2067	AVSEEAAPV	[1]EADYE	EIEMDELNQHECMATMALLKHMQR	--NKITPD--VVS	--PL--DVPPWMKFLRDKMNS	2133
Query_10008	1999	----TESAN	DKDVM[4]	KTKLDDLQDHPCCQAMVNVINYM	DMH---MGMIPEYAGT	-----PLPAWMSMRDLFNF	2062
Query_10009	2148	----KIVA	SAMNM	PIELDKINSRHECMAPLTKIDIQE	----RFRSIDIQS	--PN--DMPKWMNLYLKVKN	2207
Query_10010	2121	----NKS	SKL	----ELEDVFNHPCMKMILRVIDRLHS	----QLTKPVDPEVSM	-----PGWMDIYKFTN	2175
Query_10011	2064	NIETAQEE	IIETT	ALDRDDFDQHECMPVILRLABHVH	-----KFGANYQ	--PG--VMPDWMKSLHATLDN	2126
Query_10012	2054	KTPKESGAE	-STEV	ILESDDLNMHECMATLCGVIBHMVD	--QGISPDPDAETK	--DA--ALPPWMSMRKCLSD	2091
Query_10013	2051	ASEDSPGAP	SNIPV[6]	QMEMDELNSHESMVSIMGLIQRIQ	QavhaESKQRNEDGLiLpSS	TMPSWMSNLLKSPES	2129
Query_10014	2047	SYTDELNS	QEQS	EIEMDELNQHECMADLISLQSLVV	--NKITPIYDNGQL	--PS--EMPPWMTFVHKILD	2114
Query_10015	2041	GANEMGGVL	-M-PV	HMDQSDYNGHEVMANLTAVINHMVD	--AEIMVLHGQD	TV--PKatDVPQWMTLVQQKLD	2108
Query_10016	2044	NEAKDDDLV	SNITV	ELEIDDLNQHECMNSLTMVQHMEX	--IQL----LSS	--PK--EIPKWLNDLCKKLED	2106
Query_10017	2089	SQEASSYV	SEE--	QLEBDDLNIHECMATPLMKLLDHMI	--NKITPAPSGDSP	-----EMPPWMTPLHKRMTD	2152
Query_10018	2062	KAKTII	-N[2]ADE--	ELELDDLNKHECMELALLTLDHMAV	--NKINPAIEKDTT	--PS--EMPSWVALRKKASD	2127
Query_10019	2099	TAEQTDWY	[2]SED--	TLEMDDLNRHESMAKMSLKLHLVD	--KKISPEQVKDVP	--PT--DMPAWMAPLHKKIAS	2166
Query_10020	1989	----GAS	LYQDI	EIEMDLKINQHPCCMILRLVQMKT	-----EFGSNwnAK	--VMPGWMQRFVIVD	2045
Query_10021	2163	QRRSRKNRF	[2]QTLTA[6]	VLESDSVNDHPCMATMVALLRHMKA	--FKLIQPLEGVK	--PT--QMPKWMQNLQVSN	2048
Query_10022	2054	---DEDSEI	-QDVS	QLELDNVNKHCECMKIKRCDKMS	-----LFKREWQDN	--TN--EMPKFMQNLCLIKR	2114
Query_10001	2135	SD	T--QLNIRLFLAKVILNGESVFRPYAS	FWFQPLVA--LAT	N---	-MGLNLY--FSLDIIVTLL-SWH	2192
Query_10002	2187	QS	T--HVNIRLFLAKVILNNEVEF	TPYARFWLSGLLQ--LLI	SGNT	GGEGIHY--MVVDVVVLL-GWV	2248
Query_10003	2170	SS	T--HKNVRLFLAKVILNNEVFRPYAR	FWLSPLME--MIV	RGEA	SGEGINY--FIVDLLVTML-SWG	2231
Query_10004	2136	PI	V--PLNIRLFLAKVILNTEEVFRPYAK	HWLSPLLQ--LAA	SENN	GGEGIHY--MVVEIVATIL-SWT	2197
Query_10005	2145	SS	I--PLNIRLFLAKVILNNEVEF	RPYAKFWLGPLLQ--LVV	SGNN	GGDGIHY--MVVEIVVTVL-SWT	2206
Query_10006	2113	PA	T--SFNIRLFLAKVILNTEEVFRPYAK	FWLSPLCQ--LIF	SGPL	SSDGINY--FNVDLVMM-SWH	2174
Query_10007	2134	TT	T--HNVKRLFLAKVILNTEEVFRPYAK	FWLSPLVQ--LMV	R-DM	PDEGLNY--YVVDIMVTVL-SWS	2194
Query_10008	2063	GS	I--NVRRLFLAKVILNTEEVFRPYAK	FWLSPLIK--LIV	MDEF	GSNLIHY--FRVDLIVLW-SWT	2122
Query_10009	2208	DT[1]	L--HPNIRLFLAKVILNTEEVFRPYAK	FWLSPLIE--YVI	SNQT	GGQGIHY--FIRDIAMLL-SWP	2270
Query_10010	2176	IG[7]	T--HINIRLFLAKVILNTEEVFRPYAK	FWLSPLIK--LII	EGRS	YEGINY--FIQDLCAIVT-VWG	2244
Query_10011	2127	AA	T--HINVRRLFLAKVILNTEEVFRPYAK	FWLSPLID	PRKS	GGKPHY--MLRDIQVTL-QWN	2190
Query_10012	2092	YS	Q--PRNIRLFLAKVILNTEEVFRPYAK	FWLSPLIK--AIT	DGVV	--GRAPNY--ILSDTVAMLA-DWC	2152
Query_10013	2130	TS	T--SHNVRRLFLAKVILNTEEVFRPYAK	FWLSPLIL--LIL	KFQD[2]	ETPGPNY--FVTDLCLLL-TWA	2193
Query_10014	2115	VY	T--HENIRLFLAKVILNTEEVFRPYAK	FWLSPLIG--FLV	NSSL	--CRDLAQmdcFTLDLMVLL-SWG	2178
Query_10015	2109	AD	T--PRNVRRLFLAKVILNTEEVFRPYAK	FWLSPLIT--CLV	DGTI	--GDKINY--FFCDVMVMV-GWG	2169
Query_10016	2107	YS	V--HRNIRLFLAKVILNTEEVFRPYAK	FWLSPLIN--LIL	NATF	--GREINY--FILDLMVLL-SWS	2167
Query_10017	2153	SR	T--GRNIRLFLAKVILNTEEVFRPYAK	FWLSPLTQ--FIL	SGDS[10]	GENGLDY--FTVDAIATML-SWS	2224
Query_10018	2128	HT	T--NRNIRLFLAKVILNTEEVFRPYAK	FWLSPLIQ--LIV	-EGK	ESDGIHY--FIVDLMIIL-SWA	2188
Query_10019	2167	SS	T--HLNIRLFLAKVILNTEEVFRPYAK	FWLSPLME--LIL[4]	TEPA	GTEGINY--FIVDLVVTLL-SWS	2232
Query_10020	2046	LS	T--GLNIRLFLAKVILNTEEVFRPYAK	FWLSPLML--LDV	SAVQ	KVSEFNY--ILRDCNVLnSWK	2110
Query_10021	2239	SS	TgshPNVRILFLAKVILNTEEVFRPYAK	FWLSPLCQ--FAL	SDRN	GGQGFHT--LLVDIVVFPV-AWS	2302
Query_10022	2115	PD	I--HLNFKIFILKILNHPQLFSAYAK	FWFEPITAN--YLC	DKNN	GGKPHY--FMRDLCTLLI-SWN	2176
Query_10001	2193	EIAIPEETTVD	IAMTSHLVKVIQNCYHQNRQ	IMRNNLEVLKTLVEAWKSTIR	--IPTQDIYENFIDREDKGRN	2265	
Query_10002	2249	TVAIPED----	KRLANRLLEFLMHFCHHSNRQ	VLRNNLEMIKTLVECWKEIIS	--VPTSVIYRILATPDPVQDS	2317	
Query_10003	2232	PVAIPEDSANG	RALTSRLLHFLMANAFHKTRAV	MRRNLELIKTAVECWKTRVE	--VPTKVIHQDLGSDASSKAN	2304	
Query_10004	2198	GLATPPTGVPKD	EVLANRLNLFMLKHVPHKRAV	FRHNLEIKTLVECWKDCLS	--IPYRILFEKFSKDPNSKDN	2270	
Query_10005	2207	SVATPKGNMKD	EILANRLLEFLMKNSFHEKTA	VFRHNLEIKTLVECWKDCLS	--VPYSLIFSRFSTDPNSKDN	2279	
Query_10006	2175	TSIPQDTAEE	RAMSSRLVQFLMANIYHKNQ	IYRNNLEMLKTLVECWHDRIE	--IPYKIYINQLKSEENKS---	2244	
Query_10007	2195	SVAVPQDTAMD	RAMASRLVEFVVRNVNHDTRQ	IFRNNLEMLKTLVECWGRFD	--MPTKIYDNFSDRDPKSKTN	2267	
Query_10008	2123	A-HLKEDEAIP	KDSLNDLLKFCVKHCSWKS	GVFRKRNRIETMLVLWKSFVT	--LPSLILEKICFPLGNENEN	2194	
Query_10009	2271	N--IYQDQKLD[6]	KSNISKFINFLIQNSCADRO	ILRNNLNILKLFVERWKGFLS	--VEKKSIVEALATPVTNPKSK[5]	2352	
Query_10010	2245	SFITLESYQDK	VL-IYELMEFLMKNSYNESRI	VIRNNLIIKAIFEIWSSELIV	--VPTHVIEHELFSDSVGNVDKV[1]	2317	
Query_10011	2191	VASAPDGR---	--LATNFVHHLISVADHGS	QVLRANIDIIRLPLENWKSGSIS	--LERGQLLQYLISGGRNPDMM[8]	2266	
Query_10012	2153	RKTSLPAKSD	HHLVSQLLYLLISKVHDDR	VFRHNLEIRTVLVEWKPFD	--IPVQLLYDMIMKHKRDPV[1]	2226	
Query_10013	2194	DTAVPQDSE-	KRLARQLVQFLMENARHETTA	VFRNNVNIIRLLERWKDAVQdk	VPFKVINDMSSRQMASKDN	2267	
Query_10014	2179	SIQPKDQSEA-	--KLNIRLFLAKVILNTEEVFRPYAK	FWLSPLIK--LIL	TEPCWDRMIQ--VPVELINNFLTSDHRKLS	2249	
Query_10015	2170	SKAVPQDTVMG	RNIVGRVLRVLCMNTPTQ	TRADIFKYNLDVRSVVECWKSLT	--VPYTVIYSLMEVKDGAERST	2242	
Query_10016	2168	SIAPEDSSIG	RHAASGLLEFLMSKCPHIRK	DVFRNNLEMIKTVVQCWKERLD	--IPTYIIFESMRNLDTKVKDN	2240	
Query_10017	2225	STAIIED----	RFLGSQLLEYIMLRNSHKNRS	ILRNNLEIVRTMVELWKKDID	--IPTRVLYDQFSNSDKSTVSN	2293	
Query_10018	2189	DVAIPKDVDEI	LQ-ANRLEFLMRNAPYPNRS	VFRNNLEIITKTAVELWKNHIE	--IPYRIYKNFSPNDKDKDN	2260	
Query_10019	2233	TTAIPEDSHND	RHLASRLGLFMSRTRHRHRS	LFRNNLEIVKSVVEIWKSLH	--VPAEVIFFNSDPDPVKKDN	2305	
Query_10020	2111	DVVV---	SSLD	TT-TSRFVNLVKLCEDRRNL	VDRSNVLLVTELIALWKDATY	--INVRELIITFMSGDEDTKIK[5]	2184
Query_10021	2303	GTAIPEDDLGD	RDLVERLLDKLITGLCYDKERR	IFKNNWRLLDALIRSWRSRVHc	--VPYKRFIDLLQOTBEYATAA	2376	
Query_10022	2177	YIPEDNPKNRL	LC--TNVNTLIKYSADTKMIF	TNTEIITAGLLHRWKLVA	--INKDVISMMIGMDDKPKESN[3]	2250	

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	2266	VVGLQIVGVFLANGLSPWQPD--A---EVSKD	RFYQYLLAMNMN-QPDKQVYSASAEIVGMALSSQLQNTESDKEWcR	2335
Query_10002	2318	STGLQLLGLVFLANGLHPIDKA--S---DIDEE	RFYAALSYNLK-NRYKMHVAAAAEVI GMAMKQKADVDKVTD--G	2385
Query_10003	2305	SVGQLLGLVFLANGLPPNPS--Ts-gNIDRE	QFFSAVAKNLQTRTYKDVYAAAAEVI GMIMKYMADVEKVTG--G	2375
Query_10004	2271	SVGQLLGLVFLANGLPPYDPO--C---GIQSS	EYFQALVNMMSfVRYKEVYAAAAEVLGLILRYVMERKNILE--E	2339
Query_10005	2280	SVGQLLGLVFLANGLSPYVPK--C---EIGYE	RYFQRLVYNIStRyKEVYAAAAEVLGLVHYFAENKQVd--G	2348
Query_10006	2245	-VGIQLLGVVITCKFPYPGPD--C---PVDRH	KFFTtVAFLMC-DKSKPvYtATGEVVMGLKYLEVKEQETE--G	2311
Query_10007	2268	TAGVQMVGVFLANKLPAYSyT--S---EVDRE	KFYKALASNMG-SYWRtSSAAAAEVI GMCLQDLDQHDKDFK--E	2335
Query_10008	2195	VCGSILCLCLNTVPEtFSFL--Td---PNYES	GIFSHLVRLLS-LKNVLFKLAAFVCGLSLKLNLNELQDCEw--	2263
Query_10009	2353	STALIQLSILLSSNGYAAYDKE---ndsEISEF	KFYQLIDCLT--EYKELYDAASEICGMILQFNHKKNIQSV--	2420
Query_10010	2318	-TGLQLLGVVLAHDKPLFQDEvGldlgGLTSD	KFYMDLTNNLlnhVNSNVRALAAEVCGLVAKHLRKYHCSEA---	2389
Query_10011	2267	LVGLQLLGAIVGSGIPIYDPS--CdfstSEEH	-ICEALLNNLN-VKtKEVYAAAAELVGI TMIYRKESRQVNE---	2335
Query_10012	2227	EPGIQVAAILLANRLEPW----S---STGKL	QFLRALLRNID-TEKRSIFQSSAEVAGMALELLAANGSGML--eG	2292
Query_10013	2268	LSGLHLLSLILANLMPSESD--VeeaaGLSMR	SFFQILVANLM-FKRSRvYtAAfVtGMLLSHFKSKGRGRL--LD	2339
Query_10014	2250	--GQLPGVVLNSNGIENYEYp--M---DISRL	DFFKSLINCIK-EPSKLIHAPAAEVI GMILKR--FDSHDKY--	2313
Query_10015	2243	EAGLQLLGLIILANGLPPYTVd--S---ETEKK	RCESLVIRLLG-AHYASvYgAAAEVGMGLILKYTSQKDGSVE--aL	2311
Query_10016	2241	IVGIQLPGIIVSNGLDPYQNT--S---RIDKI	SYLkTLTVNLK-NKYKEIYESASEVCGMSMKYYPKDDDDVY---	2307
Query_10017	2294	LSGQLVGLI AVANELNPyDTN--Ta-gSLDQG	LyyKI FVNNTf-KYKLIYAAAAEVLGMIMKRMAESYDPFL--D	2363
Query_10018	2261	STGTCQLGLVAVANGLSPYRYD--V---EVDME	L---ALAE NLQf-KYKEVYEA AEVtGMIMKHLASVKKETE--G	2325
Query_10019	2306	AVGQLLGVVAVANGLFPVtSD--S---TIDED	RFYGTLVNNTf-KYKDVHAAAAEVI GLLMMQLAKVKKIFD--G	2373
Query_10020	2185	-VALQIVSAMLVAGLKSELQAEvAnq-TIEDG	-----ILLCMT-TKvTSLYtLAAAEVGLLYLTMDQATSKCF---	2249
Query_10021	2377	-AGLQIVNAYLSNALPYFNEQgVnvgGVDDSP	QILAAALVRLLN-HSRKEVYSPCAGVtGLCLRQMVTAQSVVSP--	2447
Query_10022	2251	MTGIQLLAVSFDILVIEKPenIqqtNvNR [14]	KLMNSLLKILE-SKKKQIQAASETLGKIL-----NKQPS---	2330
Query_10001	2336	AYGDNIQKMLLGIHGNK P----	DQFLTCVHC---IQKHFSAF-VE-RFIHKVLFMLPGLYGVF--RTYALEI IAAGIE	2402
Query_10002	2386	FLHSIVTS ELTKISEIK P----	DVfVTCIHK---LHFHYPEF-AD-HFTNKLLFMLPSLHGQP--KTLCELVVASRV	2452
Query_10003	2376	SFHTQVERYL LPLQVNO Q----	DVFLFCVHK---IHLHFPPF-AD-RFVSKLLFMLPKLHGIF--RTHCLEVIRGRVN	2442
Query_10004	2340	SLCELVAKQLKQHNTM E----	DKFIVCLMK---VTKSFPPL-AD-RFMNAVFFLLPKFHGVL--KTLCELVLCRVE	2406
Query_10005	2349	LVHDLIAKELKQIQSSK E----	DKFIVCLSK---IVKNYPPF-ADrRFLNQVfYLLPKLHGVL--KTHCELVVMHRAE	2416
Query_10006	2312	EFHTYIINNISSLQLSK P----	DNFIVCIHK---IHQQYPT-AD-RLLNKLLFMLPNLNGES--RRLCELVHGRIG	2378
Query_10007	2336	EYVNHLSVLSSTQSSN P----	DYMLCVVK---MHLHKKPI-AE-RFMNKLLFMLPRLHGVE--KRHCLEVISSRV	2402
Query_10008	2264	-LLKNMKAVMSRMTGTM L----	GRSKCLPLHE---ILPNYPEF-AL-QFCNDLLDWFPVSLHGDL--KTACLEGFVAVSK	2331
Query_10009	2421	-FEKLLKDKISSFLSQS ---	DYQALNCLYF---IGLHHPF--IlQFQLKLLNLPQLSNEI--RLVAFNVIYWIAD	2487
Query_10010	2390	-LSSLNPLQKISGMY [5] NeSDMKNF LTVIHH ---	SMHDVDV--AvRFLKYVfGILPRLLN--nKILALEI ISFCAG	2463
Query_10011	2336	---EILEAKLHKILKQL [1] ReAEYERFLN ILNK ---	TIRHPGF-LE-GYAEMVVDQLSRLFGAY--KVNALDTLRFPK	2404
Query_10012	2293	DDAEFLSGIVDKLNLTK D-QSKD	KFLCFLA---VQEYYPRI-VD-AFMNHVLFHLPKTYGIF--RTMCLKI VYSRLE	2362
Query_10013	2043	ELMNEVVARI DQLKDD DaTSKERL I	CLHElcnVNTGHPPV-Vd-HFVQRLFLAPSLPGQF--KtMALETLLWRAE	2413
Query_10014	2314	--FENTVSHLFEILREA D-T---	SLFITSIHK---IQLNYAA--ISerFMTKLVFFLPQLYDGF--KKMCAESILSSIR	2379
Query_10015	2312	EDA-----VVKRVTELT V-KMQG	QGLTCIYN---IHKHYPSF-VK-RLINKLLNLLPKLYGIH--RKNVLECLIAHSA	2376
Query_10016	2308	--SEFCKIVKTEISSLN A-K---	DQYLTCLYK---INTHPFIF-IDsSFLNKILYMLPNVHGIF--RKQILEI INSHVE	2374
Query_10017	2364	RLEEMVGEQFKLSQPL E-KSEDK	FIICLHK---ITLCYNKI-VP-RFVEKLLFLLPSLQGVF--RSSCLEI LDTOAD	2433
Query_10018	2326	PIFDYVTQHLSMVSAG R-MDVAR	FILCLYR---VQLHYPII-AD-RFVNKVLFLMPSLYGEF--KTKCLEI ILARAD	2395
Query_10019	2374	PLHDMVGQKLLSLVNR K-PEL	DKFITCVHK---LKLHYPII-Vd-RFVNQILFTLPSVHGDF--KSYCLEI LCSRVE	2443
Query_10020	2250	-----LCNLKKLIAADS [1] TsEEI	GYFLALRN---VSIHQATI-Vds-----IMLQRLSSVLSKAVSVDVAWALLAA	2310
Query_10021	2484	AFYKQVAEQIGKMLTAR ---	PDAGLHALQ---VSRCCPRL-ACiEFQRLLFQLPHLTGDS--LADCLVCLAAAT	2515
Query_10022	2331	----LVAQS IHPMLMNE [1] -tERHD	VFVHSIEK---TREYPELlIDrKIFMKLLSFVKVtLGTGSM--RASILKSPFRFIN	2399
Query_10001	2403	NI EGAFVEIKNKNIIG-I	LRE RDELSQLMSLKIIRRLLEKL-EASQL	L-YLLP-A--IC EISD 2459
Query_10002	2453	HI PNIFLEMKSKGIVE-L	LTH RDEGTQLVSLKLVNAMLPNL-KPDEI	L-YLLP-P--VR TVFN 2509
Query_10003	2443	VI DNLFL ELKTRNFLG-M	LTH RDDGLQAI VLQIVYALLQKL-KVEEL	L-YLLP-S--VA AFAS 2499
Query_10004	2407	GM TELYFQLKSKDFVQ-V	MRH RDEERQKVCLDIIYKMPKL-KPVEL	R-ELLN-P--VV EFSV 2463
Query_10005	2417	TI IDIFQELKKNFVQ-I	MSY SDEERQVCLD IYVKI IAKL-KAEV	K-VFLV-P--VI AFIS 2473
Query_10006	2379	QI DNVFME LTKG LLS-F	LTQ RNEKSTQVALNIITSIREKL-KSKEM	S-SLMD-S--IT AFSH 2435
Query_10007	2403	NM ENAFTELKKNGLLT-M	LTH RDDSSTQAVCLKIVNELLRLL-QPDEV	L-YFLP-A--VT AFSS 2459
Query_10008	2332	YK GDFPEELKSRGFLK-V	FSL RYERDQMLGMIVKNLHANL-TVENI	S-DLVS-S--FS SVIG 2388
Query_10009	2488	DI PELFLKLSKANLDS-L	IRI REEQTQIIILKILFKLIKNNYTTDQT	LlQIIN---LL [2]RFNQ 2548
Query_10010	2464	HD PDLFTNLQRDRLLG-F	LKH RDENEQLLALRILAGLILQEI---EV [4]	--YYLD---gLI [1]SFQD 2521
Query_10011	2405	AD [1] -NLFNQI-SRSLPK-L	LAH RDEIAQLKILQLLSQLLKNV-NQDTI	RtrVLP-V--LC [1]NFNG 2462
Query_10012	2363	IL [1] DDTYAE MSKNLFK-T	LSD RDPDQNLALSIVHKMVPHL-KPEQL	E-ELVH-E--IC KFET 2420
Query_10013	2414	DI PDLFR ELQKG LLS-L	IKH HHTETLQAVLSILKVLHHL-NASSI	K-QFLPsV--TA NCCS 2471
Query_10014	2380	VL QDPLFKSKN--FID-M	LTH RDTPIQLILCKMIHETLDKQ-TDEDL	D-RLMP-H--LE KFIH 2434
Query_10015	2377	TM EDIYLHLKEQNLDD-I	LAK KEASTQVALQLVNAVMPRL-TPSQL	L-YFMP-G--VV AFAS 2433
Query_10016	2375	NI ENAFELKKNQGLLN-M	LEH RDESIOFSSLNIIQKLLKEL-DIDKM	L-NVFP-I--IS KVHK 2431
Query_10017	2434	VQ PQFYLKLKDKGFHT-M	ISH RDEATQTASLSIVQKLLVQL-TNDQL	S-EIVPiL--IA SFSK 2491
Query_10018	2396	HI KELYTELRNKG LKD-M	LLH RDESKQVSLDLVRAMLKNL-DISQI	L-YLAE-L--VA TFKT 2452
Query_10019	2444	HI PNLFTELKRGRLKE-M	LSQ RDDGIQIGVLR IIKHMKKL-TIPEL	K-FIFP-L--LT SFVS 2500
Query_10020	2311	DT --VYNACGNDGVVKdI	FTH [11] RHSGVQISTLR AITRLLDHL-TMSEL	Nr-LVA-S--TS [9]-YQG 2385
Query_10021	2516	AA [9] PNAAVELKERGFDPsL [5] LSS [7]	AESRVPQNSHALCERLLLLLqQERL [11]	RvWLLAaVgeMC [9]SVAA 2619
Query_10022	2400	VC [9] NEIAMSL LADSEET--	LAD [1] SDENQSFLLITQIAKLLNdtGDKL	MqRVLP-R--LK [2]-FQH 2468

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	2460	LSSVLCRREMYQILMWCYDNYKDEA	QSESR--DEIITLTKDVLRLGLSDDDLNRLTVQNFWSHETR-LPGETVERL	2533
Query_10002	2510	LPSPACREVMYDILMWIYDNFRDEE	SRTEDGADDILCMAKDTLLQGLSDEEASHRLRVSNFWSHETR-LPTGTLERL	2585
Query_10003	2500	HPAPECRRTMYDILMMWIYDNYREEA	SQQES-AGRVLTGAKDCLLQGLSDEEYSLRLLVQNFWSHETR-LPTGTLERL	2574
Query_10004	2464	HPSTTCREQMYNILLMWIHDNYRDEP	SETDNDSQEIFKLAQDVLQGLIDENPGQLIIRNFWSHETR-LPSNLTDRL	2539
Query_10005	2474	HPSPVCRERMYDILMMWIYDNYRDAE	SQSDPDSKAVFVTTKDTLLQGLTDNPAQLYVRNFWSHETR-LPTGTHDRL	2549
Query_10006	2436	HPSTLCRKMFEELLMWIYDNYREDE	CTEGH---EIMLTKKECLLKGGLGDEMMCRLLVQNFWSSDTR-LPTGTLDRM	2508
Query_10007	2460	SPSPMPCRGLMYDILMWIYDNYRDEN	NATVH---QIMSRTKDALLNGMSDKDLGLRLRVNFWSHESR-LPAGTLERM	2532
Query_10008	2389	NFSPSCRRAPAECCCLSYTKLSSLN	-DSSENSKAALATTKMHLLTMLSDSNKSIROLVYNFWDASR-LEITSARL	2463
Query_10009	2549	SPNETCRLLYYDIFPMWIHDYLEKDV [ 3 ]	-----NQIRLTLIGLSDSEAVKRKLEFPWDSKSE-LSLSTVQRL	2616
Query_10010	2522	HPNNDCRRTYYTILVKMYNKINGTV	-----EIKHKLKINLLRGLIDTDESIQALIEFWHAQEE-LSHNTFTKL	2589
Query_10011	2463	HENVDCRRPEYELIYLFPKNKGMK [ 5 ]	-----QSLLMGLCDSSDIVRASLQNFHWLQLS---QDVRHRF	2526
Query_10012	2421	QRSVKCRSVMYELIKWIYDHCQ---	----DIGGVLIBEAKGILLKGLRDDPHLQDSIFEYWNSDAQQRDASTDRL	2490
Query_10013	2472	HDDYTVRAIAYDILMAVYDQKLSHE [ 7 ]	DDETAALTEIAQLARTHLLGLQDSSPKLRQSIYAYWQEKRLARNPVQRF	2554
Query_10014	2435	HPHLSCRYQFVLLISLSLEYIYQFKV	-NQIGSTSQILNFSKENLLKALIEDETSIRVMAQNFWEKAD-MPSNTIDRM	2509
Query_10015	2434	HPAPSCRETMVDVLFVVDYDNFNND	---TEDGHQLESQARGVLLRAVKEEDDALRQIVLNFWLQGA--RSMTLTQCL	2505
Query_10016	2432	SAETICRKMVYELIMHAYEKFNK-	-NGI-KEKEISRYAKEMLLHGLADSDLTLRLTVHNFWSQNH-LPAETFSRI	2504
Query_10017	2492	HGSIQCRIVLYNILLMRVYKLTDD [ 1 ]	SESDSVKMIIRSISRQDQLLGLGDESENEVKLMYBYWNQESN-LSSNLTDRL	2568
Query_10018	2453	NPSAACRQKMYGIFMWIYDTRFQE-	IDKNDTEFKQLINFCQDQLLQGLNDSSEENRQLINFNWNNNR-LPSGTLERL	2527
Query_10019	2509	HPNLLCRDLMYDIFWYDYSYRADE [ 1 ]	FQAEBSGADYLADAKDHLQGLADHRDLQKLRNFWSDEAR-LPGTLERL	2577
Query_10020	2386	HALSVCRGLLFSIAKNLYHKDISDP [ 3 ]	-----KVRI-SLLHGLCDPEELRNEAYNFWSNFAS-IKKSQDRL	2452
Query_10021	2620	LTLPAQARRQAYQLLATWYPTHNL [ 2 ]	-----ALLAGLTDSAESNRWLLFWSRNSG-LPSEASARW	2681
Query_10022	2469	NKNDSYRGLFYDLMVLYDQFENFK [ 1 ]	-----LVKSALIRGLSDKTKVVRNKLISFWSDSRR-LGLDPNQRL	2533
Query_10001	2534	VAMLEAMYTPSTELHY	LSYATNLLLEMTKSSDFNRVLFDTPLSECKFQDF--AVSSTWYQRHAAM-TPMFLDTQS	2606
Query_10002	2586	VAMLECMYSPGTEHQY	LSYASSLLEMTSKSPDYKREIFEHPLSECRFQDV--TIDHSWKQRHLAIsTPMFVETQS	2659
Query_10003	2575	VAMLECMHSPRTEPQY	LSYCTNLLLEMTSRSPDFSRREIFQHPLSECRFQDV--QIDHSWKQRHVAM-TPMFVETQS	2647
Query_10004	2540	LA-INSLYSPKIEVHF	LSLATNLFLEMTSMSPDYPNPMFHPHPLSECFQEQY--TIDSDWRFRSTVL-TPMFVETQ-	2610
Query_10005	2550	MAILRSMYSNKIEIQY	LSIATNLLLEMTSRSPDYTRMFYPISECKFQDY--NIDSSWQQRSTIL-TPMFVETQ-	2621
Query_10006	2509	VAMLEAMYSPSTEQY	LSYATNLLLEMTKSPDYKREIFEHPLSECRFQEF--SVKSSWKQRHAAM-TPLFAAT-I	2580
Query_10007	2533	VAMLEAMYSPSEVSSY	LSYATNFILEMTKSPDYNRREIFEHPLDCCPFEDF--VVRSSWKQRHAAM-TPLFAETQF	2605
Query_10008	2464	LNCLSLQYTCATESTF	LSHASVYMLLELTCSPDYTRLIYDHPLEDCTFENL--PISTHWQSRYASLeVPAVMQPTDM	2537
Query_10009	2617	MQLFEKMYTKETESQW	LGYYCSLLLQLCRTRSDFTKLLFDKPLSECTFREY--NLDASWQNRSTNM-NPLFSSSQL	2689
Query_10010	2590	KELIGNLYSTESENLF	LNACYLLELLEGSKSIENYKPIFDQQLPQSRFDENyDIDTSWRVNNMT--PLFINTQT	2663
Query_10011	2527	LELVNRIYDPCLEENW [ 1 ]	-QISNVLLKLEEDSVYKRPIFSAPLSECVFQDY--TINTSWVGASLPM-TPLFSSSQM	2599
Query_10012	2491	FLILTDLYCPATEHIF	LGYSTQVLLLEATAGTPDYNQKIFKEPLQKCFEDF--RMLLSWAQNAATV-APMPADTLA	2567
Query_10013	2555	YDGLTSLYVPEAEFAV	LQFVTNVMQLTKHSPDYDRELFAARPLSSSTWRNY--RVDASWQRNAHM-TPLFAATQO	2623
Query_10014	2510	LLIVEKLYSPKTEREF	LSYSTNLLLEKTKSPDYNRLIYENPLSECFREY--NLSADWRRRHEIM-TPLFAETVT	2582
Query_10015	2506	LHILRKMYSPESEDSF	LQMAIYVLEATRHSAADYQRNIFDQPLTECKPREM--KVSTAWRARHASM-IPMPAETQA	2578
Query_10016	2505	KAVLQKMYSPLEHMF [ 39 ]	LQYSTFLLLEMTKSPDYNLKIFQNPSTCAFNQY--AISSSWKKRHATL-TWPFVDSL	2616
Query_10017	2569	TQLLVTLYSKKTETQF	LYHATNLLLELTSRSPDYNRVLFDTPLSECRFEDYK-EIDLSSWQRHLQM-TPLFAATQL	2642
Query_10018	2528	SSVLYALYSFNLAESF	LSYSTNLLLELTSRSPDYKLSIFDHPLECKFHXY--NVNYSWQYRHAMM-TPLFAATQL	2600
Query_10019	2578	VEVLRALYSTTTEIQY	LSYTTNLLHLTSLSPDYERLMFQPLSECFREY--QIDYSWQKRHLAM-TPLFAATQ-	2649
Query_10020	2453	LSIFGSLHSSLELDENW	ILYATNLLGLSKECNEFKQPLFKISLGGGEYTTT--QIDASWEAKAQSMA-APLFSVEAD	2525
Query_10021	2682	FALLSRCHGNNTADDD [ 16 ]	-AFCSAVQLDACSSTPDFRRMPDRPLADCPRECF--RIDADWRRRHGGLmTPLFLETQO	2770
Query_10022	2534	LQIMNDLYTQEBENIW	LNSVFLLMQICQSSSDNRRKIFDAPLQDCKFQPLeiNQFGHLNRSQPL-TPLLSQMDQ	2608
Query_10001	2607	SQ-----SNDSPADGLST--G---Q	LRATQ-DAAQFTQ---TQA [ 3 ] SKGPYNWLTQSS LDT---FSDSQN	2660
Query_10002	2660	SQ-ASSGSQSQSPDSSLSLR--G---E	IRATQ-DAHQFIA---TQQ [ 6 ] RKNAYNWLTSQ -DT---FAD-YS	2720
Query_10003	2648	SHPASSQSQQAQSLSQGGQDTLGG---Q	LRATQ-DVLQFTA---TQ- [ 4 ] PKNAYNWLTSQ LDT---FAD-YS	2709
Query_10004	2611	ASQGTQLTRTQEGSLSARWPVAG---Q	IRATQ-QQHDFTL---TQT [ 2 ] GRSSFDWLTGSS TDP---LVDHST	2672
Query_10005	2622	ATQNSSRSQSQEGSLTAQGTIGG---L	VRGTQ-THYEFTP---TQN [ 2 ] ARN-FNWLTSQ LDT---FAEYSL	2682
Query_10006	2581	ST--VSESME-----GVEDSLDG---G	VRATQ-DVQ-FTA---TQD [ 2 ] SNAPPNWLTSQ LDT---YGD-YG	2633
Query_10007	2606	SL--ADGTQS-----GSPDSSFV---G	VRATQ-EIQQFTP---TQG [ 1 ] VKAPYNWLTSQ LDT---FAE-Y-	2657
Query_10008	2538	HSEVSKFISRKRKMPYGRRTYSTE---	---STA--WPGFTP---TAI [ 2 ] LKKNVGNISADT [ 4 ] LDE-----	2594
Query_10009	2690	DEE-----SSQLPTFDGdnfG [ 1 ]	-RATQ--LPMFSL---TQS [ 1 ] ---FSQFYPSSG [ 1 ] ---SSYS	2734
Query_10010	2664	SKQ-----KGNLkrQ [ 3 ]	VRATN-EDLAFSL---TLG [ 1 ] ---GNA [ 1 ] ---LGSQVA	2702
Query_10011	2600	SMDIQGMHGEDR-----DGEFDGathH [ 4 ]	IRATM-----TP-s1QA [ 1 ] ---TQT [ 1 ] ---G	2644
Query_10012	2564	SQLIQTQSL-----SGTGEHTGagpL	LRATQ-ASLAFQP---TIE Q-----SGSS IGS1-----	2608
Query_10013	2628	SQ-----ASAGEGEGEDEFSG---	VRATQ-ASLAFTP---TMA [ 1 ] ---LTGTA [ 4 ] TFEpggISGQTL	2681
Query_10014	2583	NT-----F---TESVDLR-hN	LRATQ-QTLQFQP---TL- --KVYDWLTSQ IDT---FTD---	2624
Query_10015	2579	SNDTSLNSLLSYTQGSDTNMETDgaG	VQATQ-ANV-FSA---TQA PGATYNWVDTST FDTtlaDMEYEA	2643
Query_10016	2617	SA-----ASFSDTSVKNIKID-eQ	IRSTV-DVREDFP---TQD [ 2 ] SKKAFNWLTSQ IDT---LGD---	2667
Query_10017	2643	SQ-----TQSDTSGSITG--M	LRATVqQSLAFTP---TQG QESSYNWMTPSA [ 2 ] VDYSv---PSF	2694
Query_10018	2601	GSQGS-----MA-SFSLSGSSSTDehL	LRATQsTNLQFTP---TQD -TDARNWLDPSS [ 4 ] IQSyaaFLQQSL	2664
Query_10019	2650	GSHGSGMTPTE-SMSVDGSPG----	LRATQ--NLQFTP---TQD [ 2 ] RSTAFDNLNPTQ [ 3 ] LQP-----SQSF	2708
Query_10020	2526	KFW----- [ 2 ]	LQSTQ-----QPScrNNT [ 1 ] ---LQSED [ 1 ] ---	2550
Query_10021	2771	QQQQKASDVEPGTLRQLQDGDGataa [ 11 ]	LRATAvGPGMFEA---TQL [ 1 ] ---PSA [ 1 ] --Tag1LAD---	2836
Query_10022	2609	IFQ--ASQMKAQDQLGVIKVQENepD [ 9 ]	IDTV---SNSNR---NQR [ 6 ] SQIGSQSVVEIS [ 4 ] QDGNFvRATQVP	2690

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	2661	LTLATQ	S-ESSLMFTVGS	SRIKSSKPT	--GSEFGQRKVRSDVr-----VSAMDEANKEIMRLKRRFI	2720
Query_10002	2721	TSSASD	--QSSLLFTLGA	SKMGRGRQG[4]	ATDAGFGKELQSDQprgrDEVD--AKQSGEDKINRLKRRFL	2791
Query_10003	2710	VPLGTE	T-QSSLLFTTDA	SKKVAKARR-[2]	PVGPFGSDRLSA----gRDQTDGRAAEDSQSSEIMRLKRRFL	2776
Query_10004	2673	PS----	--SDSLLFARHR	SERLQRAPLK	SVGPDFFGKRLGLP---gDEVNDKVKGAAG-RTDLLRLRRRFM	2733
Query_10005	2683	PSLASE	PgSSSLLVFSRW	SERLHRPAKK	PVGPDFGKRLALP---gDEVDSKTGIVDQADLILRLRRFL	2750
Query_10006	2634	SMSSSE	T-PSSLLFVSGT	SGSQST--NK[3]	KAGFGLRRQTIIVKQ---LGHAPKS-EGMLNKSdTIVWKLRRRFQ	2700
Query_10007	2658	STQSME	T-QSSLLFNVGT	ANENQTPKKK[3]	KTGPEFGKRLKME---gTDFTDs-REDDSEKEEIQNLKRRFL	2726
Query_10008	2595	----SD	S-SDNLTIVLIDK	ILQDIKKPKS	-----SRVKKVD-----DLGDQLQFLKRRFL	2638
Query_10009	2735	ELSSSQ[4]	SaTTSSTTFVSGS[3]	----SSQSGK[1]	-----GNKRTVNNK--pEDKAGTFKVPANIDPRVEMEMRRFK	2801
Query_10010	2703	GYSLTQ	---SNLLFSQSF[2]	---SLPRSDQG	-----GTQKSVRSS-----EYQRFRRSF	2746
Query_10011	2645	TLSSSTQ[1]	-vNGSVLYGLSQ[3]	----GAGYIK	-----KTGLEEASWMD-----PALHLRRRIL	2692
Query_10012	2609	---PSF	S1SSSLLFTTGS	SFENQRKSYK	-MGPGFGSSKLQPI-----PEVSDEVDSRSDRLAPNLHRRFL	2669
Query_10013	2682	THTQYT	AaQOVLFF----	----ERRRNR[4]	SVDPPLATPQRAAER-----QAREAREIQRLRRRFV	2736
Query_10014	2625	-TALID	T-KSSLLF----	-----KTESKNKd-----SDEKEVLK-----RRRFI	2658	
Query_10015	2644	THAPTQ	GsTSGLLFNVGA	SGQKDRRRFK	-R-PGIGNSRMRPAD---EPDGNDAADTMQSKL---IRRRFV	2705
Query_10016	2668	-WGSSS	T-SSSLLFTIGS	DTKKIPNIDE	-----NNRNSNQD-----TSKNKIEKLSSNTFPHLRRRFL	2723
Query_10017	2695	TVGATP	TsDSSLLFTRPT	MRRPPLKNPLR	-----ISAERLGVK-----ESEGGTEKEKRRKQVLQKRRAL	2755
Query_10018	2665	SSTSTQ	T-SSTLLFLSSK	SPRQRPLMQK	PLSKDFGQQLSSE--taSQPSQSVQSPSS--LNHLRRRYI	2730
Query_10019	2709	AMASTQ	T-QSLLFLFSTSQ	SHKRRKM-LK	PVPGDFGSQKLSQA--ssSRISSHQSQDPNEKKEILKLRRRFL	2775
Query_10020	2551	---SLRD	TmQSQLFSPQA	SVRLKSSYRQ[1]	IVDPTFDLRNRFST---RRNFADSSAVEDKSINDKLSKRFY	2617
Query_10021	2837	----SQ[2]	SvSTLFLRSSG[2]	-----GGGA[4]	SAAGGLAGYHTALGS-----DANPQMGAAQLHRLRRRF	2898
Query_10022	2691	VFTQTQ[5]	LiASNFYGNRRS[4]	TDSISASFYE[1]	---NKFNKKRPNGQ--vNLRIGAQSQDKVADDDGGFIAPVA	2767
Query_10001	2721	H--DQESTAR-FFM	KRNVRMKMREDIQREKMKREH	QVTMYRQYRTGDFPDIQIKYSFLIAPLQALAQRDSTIAK	2793	
Query_10002	2792	K--DHSSSQA-YFA	RREMRRKQLKELILEQREKRFST	VTMYRRYRTGDLDPDIQIKYSFLIAPLQALAQRDASLAR	2864	
Query_10003	2777	K--DQDVTKA-YHA	RQIRMQMKREERVKERQAKRGNL	VTLRYQYRVGDLDPDIQIKYSYIAPLQALAQRDMSLAR	2849	
Query_10004	2734	R--DQEKLSL-MYA	RKGVAEQKREKEIKSELKMKQDA	QVLYRSYRHGDLDPDIQIKHSSLITPLQAVAQRPDIK	2806	
Query_10005	2751	K--DREKVS-L-IFA	KRGIREQRAEQKIKTEKRMERDA	KVTLYRSYRFGDLDPDIQIKYSSLIAPLQALAQKDHSLAK	2823	
Query_10006	2701	K--DKDDR-L-FFM	KKNVRLRKRREERLEKQKSRREN	QVTMYRKYRTGDLDPDIQIKHSYIAPLQALAHRRDSTIAK	2772	
Query_10007	2727	K--DQRARST-YFA	KKNIRLQAMRQEKLEKQLRREA	QVTMYRKYRTGDLDPDIQIKYSYVIAPLQALAQRDNTLAK	2799	
Query_10008	2639	K--DTEEKAF-FV	NLNRKRQKVIERLKQO-----	SLDLRCKYRRGDFPDIQIPYSAVIIPLQAVAEELDRVAG	2703	
Query_10009	2802	KkdNRTDGER--NA[3]	RLAVKKSQRDELMSRAKQAREN	QVTMFRKYRSGELPDIQIKLEILRPLQSLCQMDHVVGT	2878	
Query_10010	2747	Kp-----YS[4]	KQMATAQRQDEIYRFSQHAFPK	NIISMLRQYRVGELPDIGINHGDLINPLQTLAQRLDLSIR	2816	
Query_10011	2693	NprATQSQVL-----[4]	----RASRRREARLKLQSLERAN	KVHMMRQYRTGELPDIIEIKHADIVGPAALAEERDPTFAR	2764	
Query_10012	2670	RrdDQKR--I-FHA	SREIQRASARQBLEQERGRLEAN	IASFREYRSGDFPDIIEISYSAIKPKMQLAKRDAAVAR	2742	
Query_10013	2737	KskDKRAQAaIFHA	KRMAARKQEARLIELRRKARAAQ	VTLFRSYRTGELPDIQIKQRELLQPLQGLAQRPDQLAR	2812	
Query_10014	2659	K--DKAESQSkFFA	RKQIAKKNRDEDLRQLKAKREN	HVEKYRNYRIGDFPDIQIKHSELIAPMQALAQRDSTIAM	2732	
Query_10015	2706	KngDREKQSI-YFA	KIEERRSKLRELEKERRFRREA	QVTMYRQYRVGELPDVQIIPHRALIGPLQALSQRDEKVAR	2780	
Query_10016	2724	K--DREKSRlhFT-	RMEMRKQQLRERKEKQKQREG	QVTIYRQYRIGDLPDIIEINYSYIIVPLQALAQHDDIEICR	2796	
Query_10017	2756	K--DKEMVSS-FFA	MNEARKMRREELKQKRAARQSV	VVMYRKYRMGELPDIQIKHSELIRPLQALAQNDSTGLAR	2828	
Query_10018	2731	K--SEMANNA-FYA	KRERKKQLREIELKQKQIAREN	KVVMYRKYRDLGELPDIQIKHSELIAFPQALAQV--CYLD	2801	
Query_10019	2776	K--DQSQISA-FFA	KQALKRKNFREQALQRQAAREG	AIVMYRKYRSGELPDIQIKYSELIAPLQALAQCDNTFAK	2848	
Query_10020	2618	Q--DKYAMLK----	-----KHDLAMTRERHQB	HVRIKRAYRAGEFPDIQITQDDIVDPIMALCEVSVETSS	2678	
Query_10021	2899	LqqQQQQAAA---S[28]	QAGVERHRRQAREAAEAARKRE	HVRTLRRRYRVGDLPDVQMEQCALVTTLAGLAQKDRMPGR	2999	
Query_10022	2768	KrkGKADLAK-FFA[31]	KQEBIRKMQREYQLRQRIQRQKH	QVGLYRRYRVGELPDIQISYEDVLKPLQAITKSLDTTAT	2873	
Query_10001	2794	LLFVAVFQAIFTQIESEKTER--	ESQSLVAKIESSLNKILETSTQY	F---SPFICGVLEIAYLFPKQL	KL	2858
Query_10002	2865	LLFSSLFRGIFSEIEEVKTER--	EAESTKVEINRHLNMLSTSTQY	F---PPFI SCVQEI SYIYHSEQL	TL	2929
Query_10003	2850	LLFSSLFRGIFFEIENVKTER--	EAEDLTEQI SRLNLALSSEKY	F---PPFTSSLQETCYREPRKL	SP	2914
Query_10004	2807	QLFSSSLFSGILKEMDKFKTSL-	EKNNITQKLLQDFNRLNTTFSF	F---PPFVSCIQDISCOQAALL	SL	2871
Query_10005	2824	QLFSALFSGILLEEVHKSKT-A-	ETKDISKQLVENFNLLKATTLN	F---PPFVACIQDMSCQHKDLL	GL	2887
Query_10006	2773	LLFSSLPKAIYKMDVEVKTER--	EMETIELINTSVETMFNTSTQY	F---PPFLSCIMDIYVVLRSRL	KI	2837
Query_10007	2800	MLFSALFRAIFSELQDVKTER--	ETTEMMAQISRLNAMLDSSSTQY	F---SPFI GCVENIYVYNSGRL	KL	2864
Query_10008	2704	ALFESLMCGVAANVIEIYNFEm--	TEKYFDSISDGLRDIRFSMVTF	H---RPLVSCLMNFLYNFIEYV	HF	2768
Query_10009	2879	SVFSLLFTTIIYQHCPRDQVNFk-	-----KSSCKQSIKIDSNATN	-----SALILSLKIVKNELELT[1]--		2937
Query_10010	2817	TLFTSLFVSIYNQLDNNIQSE-	EIKEYKTI GRHIQDIFSNSTMN	F---PPFVGSCLRICYEVNDIE	-I	2880
Query_10011	2765	LLLRILCSAVYSLPSGGQDIKa	-----EVRGLGLEAALRQTHNG[1]	-----SFVGCABALCLEDPETW	-I	2821
Query_10012	2743	HLFVLLFDGLLGEL-----YD-	EKEYVTQKVKAMGSI VSTTKDY	T---TMVMGAVMEIALCHSRYI	TL	2802
Query_10013	2813	MVFESLFYCVLHELDPILSER--	DAANMKAEYKNSLVNGLLQATRKG	Y---TPFTI AALHNLMSRADPAI	-V	2876
Query_10014	2733	VLFSFIFVSLISEISYLKDEQ--	DSYKILKELTQYLNAMLASSSEF	N---SNFIACILEIKCNPKLL	KL	2797
Query_10015	2781	LLFEVLAQGVINSAGTAMHHR--	EEQEWQNLHEALNGLLSSSYMC	C---PQMLRTVLHLMINND--V	RT	2843
Query_10016	2797	MLFQSL---ILSVCCYMKSNQv	tEYSEFCRSLSSALNEILKKSVMH	H---PQLVAFCECLNCDVKEM	DI	2860
Query_10017	2829	LLFCSLFNSIFSQLDHLLTEE-	DAATKGVIEGNINTILDCSTQY	F---PPFI GSLQDICYENS-L	KI	2892
Query_10018	2802	QNYRYVNLV-NDLVENFNDE--	DAETKAEIKASLSNIVKSKHY	F---PPLIACVQDICYQKEL-	SI	2864
Query_10019	2849	QLFSALFRGIFSKIDEKLPFR--	EAAVSTDLRSAMHMLSSSTQF	F---PPFTSSIQDICFHEKCLR	SV	2913
Query_10020	2679	LLFDELFSIVA-----TNEf	EKSGIVDELTEKIEGTLTSLKAS	Sv----YVGCIVSYAVASMRH[8]-L		2744
Query_10021	3000	LLCAQLSQSMVDWAASALSED--	DVADMCRLLAAAWGLIPTSFPF[13]	AtndPALIGCALDSLCAVASSA[9]DV[7]		3096
Query_10022	2874	EVFVEIFSELYKQNDQQQRE1-	-----LGGQVRNINLESILY	Dygc---INCIHRVAMELLKTD[2]TI		2933

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	2859	QPGVSVSTVCNMSDQQLGILLLE	EGLLHSP	EDQS-P-ASKRRSRDRSSR--ITQDTEP	WLELAKLYKS	2922
Query_10002	2930	VPSTISTASLTSQQMHIGIRLLE	ENLIQKP	WQEE-K-SAKRARFV-SRE--PSPEIST	WIELARLYKA	2992
Query_10003	2915	DASVIQQTGVISLQQPIGIVLLE	EQLVQKA	GVEE-P-RSKRARTS-SSP--VSAEATAT	WIELARLYKS	2977
Query_10004	2872	DPAAVASAGCLASLQQPVGIRLLE	EALLRLK	P-AE-L-PAKRVRGKARLP--PD--VLR	WVELAKLYRS	2932
Query_10005	2888	DSTSVRSTSLASLQQPMGILLLE	KSLIHST	TNEE-P-PTKKARGRQELP--PPTDVAR	WIELATLYRS	2951
Query_10006	2865	SFSNLSTCAIVSSLOPLGIVVLE	EQLIQTD	STEQ-R-GNKRKRTE-GTT--VSTDVNL	WIEMARLYKS	2900
Query_10007	2838	EPASAVDTASLISMQQPLGIVLLE	EQLIRDG	SIER-R-PAKRARTS-ATV--PSQETST	WIELARLYKS	2927
Query_10008	2769	EPTDIANVAIGSQDPCGIVLLE	RRIVEFG	VDISAV-TCQEKRLKQVNE-----DTRL	WVELARLYKQ	2830
Query_10009	2938	SPASIASASNASGNHFMGIIIVIE [ 3 ]	ESLLPTK [ 1 ]	-----SSKQYAQNQATIS----- [ 2 ]	WDHLRELYHA	2996
Query_10010	2881	DPNMIRKACSLSTESLGIILIE	RQLLRTS [ 1 ]	---ERE---SKRLRRTSSSLV---ASSKR [ 1 ]	WIELSHLYKS	2942
Query_10011	2822	SPKIMGSARKSTNYHSGIMLLE	KQIMNET [ 1 ]	PEPEPNvSSKRQKGGQAn-----RL [ 7 ]	WLELAELYKA	2891
Query_10012	2803	QAEDITTVSRFSGLLALGCLSLE	EILIQKD	TVSEAG-SSKRSKTFEP-----STEEIL	WLSLADLYKH	2864
Query_10013	2877	SPMTVSLTSAAGRSFHTGALLLE	STLAAAE	EEAAMQ-PSKRSMRRARnglTDDTEIER	WIYLGRLFPKE	2943
Query_10014	2798	DPGLVTSACLNSFRSLGIIILE	EYIINFE	EDSDAP-LLEKKVRLNPKDS--SSEASIL	WIELAKLYRS	2862
Query_10015	2844	NPESLTAACLGSRLEALGILVLE	RQALLSK	RGEEPH-PARKKMRKDEAE--LQODDSL	WLSIAEMYKN	2908
Query_10016	2861	DYEQICTASQASHKQSLGILLLE	DHLICYN	NFKSSN-PAKRKIKLNKPLI--SDELSI-	WLKMAEYKS	2924
Query_10017	2893	KPPSITTSSLASKQEFPTGIMLLE	KQLL-LN	VDSS-K-SAKRSRTN--K--SNEVETP	WIELARLYKS	2952
Query_10018	2865	QSSFASVASLASSQPIGILLLE	KQIMHLD	TEEP-P-S-KRIKTSASA---SEA-TSN	WIELSKLYRS	2924
Query_10019	2914	DSASVSTACTLNSLQQPIGIMLLE	KQLLEIE	GKSQ-N-SRKRKAKTSTVPP--SEV-TTC	WIELSKLYRS	2976
Query_10020	2745	RPKTI GDAGLASGNHVSQLALE	ELLIYET	QLMAGNm--KLSNESLTEC-----	WDLHLKLLST	2802
Query_10021	3097	DYGLVAAAASAAHLPTL GALALE [ 11 ]	QQQRDE [ 11 ]	TTQLPTAAKRRRMEAGSQe-SNDDNSA [ 14 ]	WLALARLYRD	3199
Query_10022	2934	BANI I QRTGQHSMSFTALALLE [ 11 ]	EQQQAAM	GKQGNKk--RKIQDMSGQN-----QKNGT [ 14 ]	WPNLIDLIDYQ	3020
Query_10001	2923	ISDYDVLRGIF---S-SQIGS-KPATRLAIEAE	T-RGDYRSALQLYNES---LAKEE-WDDGehPTEAEI	NMWD	2986	
Query_10002	2993	LGDFDAVQGVF---S-ACIGT-HEITKTALAEAE	A-RGDYSTALKLYNQA---F-EMD-WDEGA-VSQVEE	DLWD	3054	
Query_10003	2978	LGDFDVLQGF---S-RKVG-TQVTYEALAEAE	A-RGDFKKAVELYNQA---MQTE-WDPDQ-PLAVEE	DLWD	3040	
Query_10004	2933	IGEYDVLRGIF---T-SEIGT-KQITQSALLAE	A-RSDYSEAAKQYDEA---LNKQD-WVDGE-PTEAEK	DFWE	2995	
Query_10005	2952	IGDYDVLRGIF---S-EKIGT-KPLTHTALLAE	A-KSDYAEAVRLYDQA---LNTEE-WMDGN-PMETEK	DFWE	3014	
Query_10006	2901	VENYDVLHGIF---S-GKIDT-KPITQEAAMAE	S-KGDYQTANKLYDEA---LSIQD-WAGED-PLAEAI	DFWD	2963	
Query_10007	2928	LGDFDVLRGIF---S-SKIGT-KKITQDAALEAE	S-RRDYKTASHLYREA---RQCDS-WTDGE-PDGLEQ	ELWV	2990	
Query_10008	2831	CGELENLKA V F---T-KHLLT-TEKTKLAEME	A-AGDFEGAATYIGET---FSNKQ-SKSGDI VEKABS	NLLY	2894	
Query_10009	2997	LKEEDIVLGLLE---RQMG-LPYTKRALEAE	-L SNDWVA VLKAYDDGmaaleSESG---LAI tPSARET	SLWE	3061	
Query_10010	2943	IDGHVYNSIYE---S-KIVV-SELTKDALNAE	-LIGDYALACKHYLEA---LKSEA-----GVDDEIV [ 1 ]	LWE	3001	
Query_10011	2892	LGENDIVLGLFg---T-HVAK-HEKTRRALEIQ	-LGGDLRQALDLYDQV ia--IYEDGSLSGD-ITKSEI	DIWY	2956	
Query_10012	2865	MDEWVVRGIF---H-KYVNC-GENVHKALKEE	T-SGHWDKARMEYEA---LSSMD---DQD-GLS---	DYYY	2922	
Query_10013	2944	IGDYDVLHGIF---S-NQVLS-FEATQARA L E	A-ARGFDEAAEVYKEC---MFKEEdEEDDD-VPACEV	DLWE	3007	
Query_10014	2863	MNDYDSIKGIF---LvKSDLV-CEHTERGIDFE	A-VNDYYSARNEYGAA---LDSE---LLEN-VSQSEQ	DLWL	2924	
Query_10015	2909	LNMWVVRGVV---Q-GKLGAIKEETQKALEAE	A-TNNHVRAFLLYRQS---LDTTD---WEEE-PLAAEN	RLWE	2971	
Query_10016	2925	LSDYDSVHGIF---S-NSLNI-QEETKQALEFE	S-RGEYLLAVKIYSKL---YDQTE-WSAEE-PNQAEI	DLWD	2987	
Query_10017	2953	LEDYDVLHGIF---T-NLDDT-KDVTRKALEAE	E-RGDYLTALNLYKKA---IDF-D-WPQGSPPQVEE	DLWE	3015	
Query_10018	2925	IGYDVLRGIF---S-GHLGT-KDITQRALEAE	S-RHDYAAAVKYYNQA---MDCSS-WPDGE-PQAEAE	DLWE	2987	
Query_10019	2977	VGDYDVLRGIF---T-GHIGT-KDITREAMEAE	A-RGDYSQALKLYNEA---IGTDDvWIKEE-LRQEEE	DLWD	3040	
Query_10020	2803	LHKRDFLVALSLacS-----T-NADSKRALQEQ	L-SGDLPLAIAAYKKAestLRLQMeLDDGdFVIARN [ 4 ]	CRWQ	2873	
Query_10021	3200	MGEPPDQLGALrrcRsrHNGQ--EGVDNDDDE [ 7 ]	ClSGDVQLALATFDRL---FRQAdesGTESaSAAADs [ 5 ]	INACQ	3280	
Query_10022	3021	IGNDDALQGIWtflA-DEEGI---IFKVKDDQE	KLSQDQERTVNLIKEAhc-LRTKG--SIDKgLKVLEQ [ 1 ]	LND	3087	
Query_10001	2987	SSRIECAEN--LTQWRDVELTSTSDL-E-G	SS---LSAVWDSTYHQ---EVYLPPLMRSKIKLILN---GD-TSQDN	3049		
Query_10002	3055	ESRLNCCNQ--LTQWKELQNYSTVNI-D-D	NSPHDLGKIWSDTFYQ---ELYLPMMRSMLKQLME---GG-EDE-T	3119		
Query_10003	3041	DSRLQCYDH--LTQWKNLVEVSTVNI-D-E	DSPPNLDKMDDTYFQ---EHYLPPTIRSKLKLLE---GE-MDQ-S	3105		
Query_10004	2996	LASLDCYNH--LAEWKSLCYCSTASI-D-S	ENPPDLNKI WSEPFYQ---ETYLPMYIRSKLKL L LQ---GE--ADQS	3060		
Query_10005	3015	IAALECYNH--L TEWKAL EYCVSTVNI-D-D	KTPADIENTI WSDTFYQ---ETYL SYMIRSKVKLLQS---GE--VDQS	3079		
Query_10006	2964	DSRMQCLDH--LTQWKDLETVCLTAI-D-N	TELPNLK V WEDTLYQ---EHYLPYIIRSKLKL L L L---GG--DQQS	3028		
Query_10007	2991	DAYLLCVDQ--LMSWEELGTEAMDCI-E-G	LEETGLNKI WEDAYYQ---EHYLPYIIRSKVKLL L LQ---GN-EDQQ	3056		
Query_10008	2895	EARTRCLEM--LNDWKLETWSTTFS--as [ 1 ]	-KTAKL NQIWETDYLI---NMHLPVFLRSK L K IAL T---SDsAALEE	2961		
Query_10009	3062	NGRLQCYES--LREWSNLKANFIAYY--p [ 2 ]	-----QSVFEEQDEA-tRHEMMSHFLOYSLKVKENw-----PD	3121		
Query_10010	3002	ERRFDCFEK--LCQWDDLAETVLVDI---	--EDNLSNLWEDELQD i---YLQHFVHSFPKLSLSddnSE---YRQR	3064		
Query_10011	2957	NQRLACLAD--LNEWVLMGEIEYESmEgD	NSEPDFKRLWEPTKQE i---YLG YFIRGAVKEAKYh---GL-----	3019		
Query_10012	2923	EALYKCLAH--LSDWKGLNQQVRRQV----	--ENNLDLQWTDWLQ---EKL L P Q I L L S E I Q M M L D---kG---VDSK	2983		
Query_10013	3008	ETLAECSLR--LCDWGELENTLTICI-Np [ 9 ]	IAEVIDANNI WQPRLI---QRHLP H I T V R L E L G I S a-----VVCAC	3082		
Query_10014	2925	ESLLKCCNE--LTDWKAM-----C-EwS	MNGNTLDQLFEDAYS I---ENVPYAFRSK L K L I L Q e d e N e q K K H K D	2988		
Query_10015	2972	EWYSNCAAI--LGQWSELENFVEKRFLQdD	DGEVLDNRVLLPKPT---ATALPALVHSK L M N I L D---gSE--TDGN	3039		
Query_10016	2988	NAI L Q C Y D Y --L T Q W D L L Q L S I I N R L -D e N	QKQP N L N K I W E E P Y Y Q ---E N Y L P F L I K S K L K L L I ---E n R n D Q S	3053		
Query_10017	3016	DSLNLAYSH--LTKWKDLEKASTCNI-G-E	ETNSKLEQLWKDSYCYQ---EHYLPVLSKSLK L W C T ---D D---SDHG	3080		
Query_10018	2988	ISRLCCLYN--LTKWDKLEYLSLVNV-D-N	NDPPSFDKI WEDNY Y Q ---E N Y L P L F L K S K A K V M C E ---N D L A G S D D	3054		
Query_10019	3041	VSRLQCLAD--LTKWKELEKCSVESI-D-D	KNPPDLK V W S D T F Y Q ---E H Y L P H I T S K L K L Q C Q ---G---SGDMT	3105		
Query_10020	2874	--RLNCMET--LNSWHALNSEITEAA---N [ 1 ]	---DTNLYWKORPPY---LEQGVGHYLRSC L G L L E A s p s N D ---L L E S	2936		
Query_10021	3281	LARLDCLTR--LADWSRLSGVSKLL1TpA [ 8 ]	-----ADRLWSRVGVQnvcRNALSSYMLSSRTLRVCDkpeSAiGSDSE	3358		
Query_10022	3088	KSLTDCLEDhiLCELENKQLDSKAEL1Kwd [ 2 ]	---AKDMSEK Y H S K Q L S e i p F K H A E K M I R A Q I R T D P E Y w-----N N L T K	3159		

Supplementary Figure 5: COBAL analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	3050	LLSF---IEESLKT <sup>1</sup> PVHKAFIEG <sup>2</sup> SFSE <sup>3</sup> ELALVCIH	QELYDKAKHYANMSLQAFIQD <sup>4</sup> WANLSTTRTSTRA--TKLQKL	3121
Query_10002	3120	LLQF---IDASMSDAD-----KRAFLEARYSE	QEDYDRAMFYLANCRQSF <sup>5</sup> LHDW <sup>6</sup> SGLG <sup>7</sup> PLMISSRR--AKLQCI	3191
Query_10003	3106	LLNF---VDSAMQK <sup>8</sup> PKDRKALLESRYCE <sup>9</sup> ELAML <sup>10</sup> YL	QDDYGRAKYV <sup>11</sup> GNCTDA <sup>12</sup> FLQEW <sup>13</sup> SGMDRLMTSSRA--AKLQSV	3177
Query_10004	3061	LLTF---IDKAMHGELQKAI <sup>14</sup> LELHYSQ <sup>15</sup> ELSLYL	QDDVDRAKYI <sup>16</sup> QNGIQS <sup>17</sup> FMQNYSSIDVLLHQ <sup>18</sup> SRL--TKLQSV	3132
Query_10005	3080	LLTF---VDRAMKVEQ <sup>19</sup> KAVLE <sup>20</sup> TYYSQ <sup>21</sup> ELSLYL	QDDVDRAKYI <sup>22</sup> INNCMEV <sup>23</sup> FMQNYSSIDVLLHQ <sup>24</sup> SRL--TKLQSV	3151
Query_10006	3029	LLTF---VDTSMKTTEY <sup>25</sup> RAVLESRY <sup>26</sup> MELSLMYLW	QDNYDRARHYANLALQKFLQ <sup>27</sup> DWSSTDTLIVASRT--SQLQTL	3100
Query_10007	3057	MLDF---VDNMMNVEEQ <sup>28</sup> KIFLEARYSEELALYL <sup>29</sup> FL	QEDYARARHYTSIV <sup>30</sup> KTF <sup>31</sup> LQD <sup>32</sup> WGSIDTLM <sup>33</sup> TTRI--GKLOIL	3128
Query_10008	2962	LSAF---VEQNTVPK <sup>34</sup> NREILEN <sup>35</sup> RHSG <sup>36</sup> LVLN <sup>37</sup> LIL	REDFTKSQY <sup>38</sup> FLNVA <sup>39</sup> IDS <sup>40</sup> FLKCY <sup>41</sup> PTLRVCDV <sup>42</sup> TARL--KLVDL	3033
Query_10009	3122	LYQFm---GSLTPSQ <sup>43</sup> HQYMESS <sup>44</sup> FPGL <sup>45</sup> EALF <sup>46</sup> IT	RSDFNRSQY <sup>47</sup> YIQ <sup>48</sup> KFYRL <sup>49</sup> FRD <sup>50</sup> WASTHPLALASRHf--ILQPI	3191
Query_10010	3065	FFDF---IDGAMNNOY <sup>51</sup> QKSLT <sup>52</sup> TQY <sup>53</sup> PELEALS <sup>54</sup> SVT	KKDLEQGRFY <sup>55</sup> V <sup>56</sup> RKAFD <sup>57</sup> N <sup>58</sup> FLF <sup>59</sup> AWSTLHPLAIG <sup>60</sup> TRL--GKLSL	3136
Query_10011	3020	LTEF---VQSSSEILPK <sup>61</sup> RQFL <sup>62</sup> LN <sup>63</sup> EFST <sup>64</sup> QLATLAAI	NKEWDRARFY <sup>65</sup> LQ <sup>66</sup> CHLI <sup>67</sup> FR <sup>68</sup> QWSALHPFALG <sup>69</sup> SRH--QVRVQL	3091
Query_10012	2984	REQFlqqVDRWLRDRD <sup>70</sup> KSDY <sup>71</sup> LSSH <sup>72</sup> YCEELAAIFLI	QDKLDQAQH <sup>73</sup> QVEHSL <sup>74</sup> TM <sup>75</sup> FLDSWAALNPSSHN <sup>76</sup> LRT--KQLLDL	3058
Query_10013	3083	LQAF---VEAARPHDDRWAH <sup>77</sup> LQYH <sup>78</sup> FAMELAML <sup>79</sup> SVL	REDFDLAR <sup>80</sup> YHL <sup>81</sup> SQLQR <sup>82</sup> FLDD <sup>83</sup> WT <sup>84</sup> LSPLMT <sup>85</sup> TERR--ASLQRL	3154
Query_10014	2989	LIRF---INGL---EVEKKY <sup>86</sup> LEQ <sup>87</sup> SFCQ <sup>88</sup> EALF <sup>89</sup> LSLH	QNDSAANYYAHLA <sup>90</sup> IQ <sup>91</sup> KYL <sup>92</sup> TEW <sup>93</sup> SSTNKTI <sup>94</sup> KONQI--NKLQSL	3058
Query_10015	3040	LIAF---IDQSMGDHRH <sup>95</sup> RSML <sup>96</sup> EGL <sup>97</sup> PLQL <sup>98</sup> GVLA <sup>99</sup> AH	QEKFEQANVY <sup>100</sup> ANTAT <sup>101</sup> SM <sup>102</sup> TL <sup>103</sup> LSQ <sup>104</sup> SVL <sup>105</sup> TPN <sup>106</sup> PLL--DTLQNV	3111
Query_10016	3054	LLCF---IDSALKEEN <sup>107</sup> KKYLEEN <sup>108</sup> FSEELALLYII	QENYDRAQYYS <sup>109</sup> QLC <sup>110</sup> LQ <sup>111</sup> DF <sup>112</sup> LKEW <sup>113</sup> S <sup>114</sup> SLN <sup>115</sup> VLM <sup>116</sup> PI <sup>117</sup> IRK--TKLQSL	3125
Query_10017	3081	LASF---LDKLSLSEAN <sup>118</sup> RSVLESH <sup>119</sup> YQ <sup>120</sup> E <sup>121</sup> FL <sup>122</sup> SVL	RDDYDRAYYL <sup>123</sup> SL <sup>124</sup> SLKAK <sup>125</sup> Q <sup>126</sup> QI <sup>127</sup> VS <sup>128</sup> VDG <sup>129</sup> ISS <sup>130</sup> TGA <sup>131</sup> vs <sup>132</sup> TVQ <sup>133</sup> SL	3154
Query_10018	3055	FIQY---LNKSLANDE <sup>134</sup> HLSVLE <sup>135</sup> NRH <sup>136</sup> C <sup>137</sup> DCLAL <sup>138</sup> HVL	LEKYDRAR <sup>139</sup> Y <sup>140</sup> LD <sup>141</sup> FS <sup>142</sup> I <sup>143</sup> Q <sup>144</sup> FLQ <sup>145</sup> D <sup>146</sup> W <sup>147</sup> SALD <sup>148</sup> STLL <sup>149</sup> YSRA--SKLQSL	3126
Query_10019	3106	LN <sup>150</sup> EF---VAKAMK <sup>151</sup> T <sup>152</sup> SE <sup>153</sup> RAAL <sup>154</sup> RS <sup>155</sup> Y <sup>156</sup> SD <sup>157</sup> SLAL <sup>158</sup> MYL	QDNCD <sup>159</sup> S <sup>160</sup> Y <sup>161</sup> AA <sup>162</sup> L <sup>163</sup> SE <sup>164</sup> QS <sup>165</sup> FL <sup>166</sup> SD <sup>167</sup> WAG <sup>168</sup> LD <sup>169</sup> SN <sup>170</sup> LR <sup>171</sup> SRS--IKLQSL	3177
Query_10020	2937	LQNF---IGDAV <sup>172</sup> RS <sup>173</sup> AKV <sup>174</sup> KL <sup>175</sup> IQ <sup>176</sup> SK <sup>177</sup> F <sup>178</sup> VEL <sup>179</sup> CLAY <sup>180</sup> LA	[2]--DKNQ <sup>181</sup> TR <sup>182</sup> VC <sup>183</sup> VEN <sup>184</sup> F <sup>185</sup> Y <sup>186</sup> LN <sup>187</sup> FL <sup>188</sup> K <sup>189</sup> I <sup>190</sup> WR <sup>191</sup> Q <sup>192</sup> T <sup>193</sup> SP <sup>194</sup> MAT <sup>195</sup> IS <sup>196</sup> RM--ELLQSL	3008
Query_10021	3359	SAEF <sup>197</sup> m <sup>198</sup> d <sup>199</sup> f <sup>200</sup> LR <sup>201</sup> CT <sup>202</sup> TD <sup>203</sup> ND <sup>204</sup> SV <sup>205</sup> RT <sup>206</sup> Q <sup>207</sup> PE <sup>208</sup> SE <sup>209</sup> HA <sup>210</sup> LLAS <sup>211</sup> VLI <sup>212</sup> K	[2]--KYHN <sup>213</sup> AMAY <sup>214</sup> V <sup>215</sup> DS <sup>216</sup> AL <sup>217</sup> DR <sup>218</sup> V <sup>219</sup> L <sup>220</sup> NN <sup>221</sup> V <sup>222</sup> SS <sup>223</sup> PA <sup>224</sup> LQ <sup>225</sup> RQ <sup>226</sup> F <sup>227</sup> HL--LNMASV	3433
Query_10022	3160	QIAFw-----Q <sup>228</sup> RD <sup>229</sup> V <sup>230</sup> SKAY <sup>231</sup> LD <sup>232</sup> KN <sup>233</sup> F <sup>234</sup> YEL <sup>235</sup> AN <sup>236</sup> SL <sup>237</sup> T	[1]--KD <sup>238</sup> Y <sup>239</sup> DRAR <sup>240</sup> F <sup>241</sup> YL <sup>242</sup> DK <sup>243</sup> ET <sup>244</sup> N <sup>245</sup> EL <sup>246</sup> LS <sup>247</sup> Q <sup>248</sup> W <sup>249</sup> KN <sup>250</sup> L <sup>251</sup> TK <sup>252</sup> LS <sup>253</sup> QVA <sup>254</sup> QHf--LVQRI	3228
Query_10001	3122	QBFTEMQEFLSVIKTTS--GEDFPRQ--IDRTLHKWEARSPDRLLDSVVIW <sup>255</sup> DNII <sup>256</sup> LN <sup>257</sup> RNV <sup>258</sup> YMD <sup>259</sup> HFAEHL----	----	3186
Query_10002	3192	QRLTEKQNF <sup>260</sup> LKLA <sup>261</sup> KIPD--LQEAIGCq <sup>262</sup> ks <sup>263</sup> LE <sup>264</sup> LLDR <sup>265</sup> WG <sup>266</sup> CH <sup>267</sup> LD <sup>268</sup> SK <sup>269</sup> KDP <sup>270</sup> INI <sup>271</sup> W <sup>272</sup> DD <sup>273</sup> V <sup>274</sup> SN <sup>275</sup> RL <sup>276</sup> LY <sup>277</sup> MD <sup>278</sup> KL <sup>279</sup> GSR <sup>280</sup> FIRSE	SQQA	3266
Query_10003	3178	QKLFTEMQEFLLELISNDR--NFETPW--p <sup>281</sup> V <sup>282</sup> TR <sup>283</sup> LE <sup>284</sup> E <sup>285</sup> K <sup>286</sup> WS <sup>287</sup> SR <sup>288</sup> Y <sup>289</sup> PD <sup>290</sup> N <sup>291</sup> RM <sup>292</sup> DP <sup>293</sup> IS <sup>294</sup> V <sup>295</sup> W <sup>296</sup> DDI <sup>297</sup> AN <sup>298</sup> CL <sup>299</sup> Y <sup>300</sup> LD <sup>301</sup> KL <sup>302</sup> ND <sup>303</sup> KE <sup>304</sup> FQMA	ARGV	3250
Query_10004	3133	QALTEIQEFISFISKQG--NLSSQVP--LKRLLNTW <sup>305</sup> TN <sup>306</sup> RY <sup>307</sup> PD <sup>308</sup> AK <sup>309</sup> MD <sup>310</sup> PM <sup>311</sup> NI <sup>312</sup> W <sup>313</sup> DDI <sup>314</sup> T <sup>315</sup> N <sup>316</sup> RC <sup>317</sup> F <sup>318</sup> FL <sup>319</sup> SK <sup>320</sup> IE <sup>321</sup> E <sup>322</sup> K <sup>323</sup> L <sup>324</sup> PLP	[12]P <sup>325</sup> SDR	3217
Query_10005	3152	QAVTEMQDFLHFI <sup>326</sup> SK <sup>327</sup> EY--NFISEVP--LR <sup>328</sup> KL <sup>329</sup> RT <sup>330</sup> W <sup>331</sup> T <sup>332</sup> GR <sup>333</sup> Y <sup>334</sup> PD <sup>335</sup> AK <sup>336</sup> MD <sup>337</sup> PM <sup>338</sup> SI <sup>339</sup> W <sup>340</sup> DD <sup>341</sup> VI <sup>342</sup> T <sup>343</sup> N <sup>344</sup> RC <sup>345</sup> F <sup>346</sup> FL <sup>347</sup> DK <sup>348</sup> I <sup>349</sup> Q <sup>350</sup> ER <sup>351</sup> LV <sup>352</sup> SQ	[1]V <sup>353</sup> ETR	3236
Query_10006	3101	QPLVELHEFLTYI <sup>354</sup> SNEN--DITHT--T--AN <sup>355</sup> QL <sup>356</sup> LL <sup>357</sup> AK <sup>358</sup> WR <sup>359</sup> Q <sup>360</sup> PH <sup>361</sup> KL <sup>362</sup> DP <sup>363</sup> VNI <sup>364</sup> W <sup>365</sup> DD <sup>366</sup> VI <sup>367</sup> S <sup>368</sup> Q <sup>369</sup> RV <sup>370</sup> Y <sup>371</sup> LD <sup>372</sup> HS <sup>373</sup> QCL--AG	KNKE	3170
Query_10007	3129	QKLFTEMQEFLDFVSNED--TFQDSQA--ALS <sup>374</sup> LV <sup>375</sup> TR <sup>376</sup> W <sup>377</sup> Q <sup>378</sup> TR <sup>379</sup> SP <sup>380</sup> PI <sup>381</sup> VE <sup>382</sup> PC <sup>383</sup> NI <sup>384</sup> W <sup>385</sup> DD <sup>386</sup> IV <sup>387</sup> T <sup>388</sup> N <sup>389</sup> RC <sup>390</sup> LY <sup>391</sup> VD <sup>392</sup> QL <sup>393</sup> CHKL--SS	VQ--	3197
Query_10008	3034	L <sup>394</sup> PV <sup>395</sup> SE <sup>396</sup> TAK <sup>397</sup> F <sup>398</sup> L <sup>399</sup> H <sup>400</sup> L <sup>401</sup> H <sup>402</sup> ASE--HLKLESV--V <sup>403</sup> RL <sup>404</sup> CR <sup>405</sup> F <sup>406</sup> W <sup>407</sup> LR <sup>408</sup> Y <sup>409</sup> GL <sup>410</sup> D <sup>411</sup> VE <sup>412</sup> T <sup>413</sup> LT <sup>414</sup> T <sup>415</sup> ADD <sup>416</sup> V <sup>417</sup> V <sup>418</sup> NR <sup>419</sup> KT <sup>420</sup> F <sup>421</sup> LEM <sup>422</sup> FG <sup>423</sup> K <sup>424</sup> GL <sup>425</sup> V <sup>426</sup> DSG	[1]----	3102
Query_10009	3192	QKVVEVEEFLGLVTNDK--RAIEVdk <sup>427</sup> ME <sup>428</sup> KL <sup>429</sup> VL <sup>430</sup> WK <sup>431</sup> TR <sup>432</sup> Y <sup>433</sup> PSS <sup>434</sup> F--DD <sup>435</sup> IL <sup>436</sup> V <sup>437</sup> W <sup>438</sup> DD <sup>439</sup> LV <sup>440</sup> S <sup>441</sup> FR <sup>442</sup> S <sup>443</sup> V <sup>444</sup> LE <sup>445</sup> KI <sup>446</sup> YER <sup>447</sup> FIG <sup>448</sup> Q	[1]----	3260
Query_10010	3137	Q <sup>449</sup> VI <sup>450</sup> VE <sup>451</sup> E <sup>452</sup> EF <sup>453</sup> LN <sup>454</sup> L <sup>455</sup> Q <sup>456</sup> SS-----K <sup>457</sup> d <sup>458</sup> q <sup>459</sup> I <sup>460</sup> W <sup>461</sup> D <sup>462</sup> S <sup>463</sup> M <sup>464</sup> I <sup>465</sup> SH <sup>466</sup> ERR <sup>467</sup> F <sup>468</sup> PS <sup>469</sup> KT <sup>470</sup> LD <sup>471</sup> SS <sup>472</sup> NT <sup>473</sup> W <sup>474</sup> DD <sup>475</sup> IV <sup>476</sup> T <sup>477</sup> TR <sup>478</sup> QS <sup>479</sup> IL <sup>480</sup> KE <sup>481</sup> MA <sup>482</sup> ER <sup>483</sup> IK <sup>484</sup> LN <sup>485</sup> R	----	3202
Query_10011	3092	Q <sup>486</sup> KI <sup>487</sup> VE <sup>488</sup> LE <sup>489</sup> EE <sup>490</sup> FF <sup>491</sup> TI <sup>492</sup> SK <sup>493</sup> PE-----A <sup>494</sup> ge <sup>495</sup> LS <sup>496</sup> RL <sup>497</sup> SG <sup>498</sup> WH <sup>499</sup> K <sup>500</sup> WR <sup>501</sup> PS <sup>502</sup> PN <sup>503</sup> F <sup>504</sup> DD <sup>505</sup> VEA <sup>506</sup> W <sup>507</sup> DD <sup>508</sup> IA <sup>509</sup> NR <sup>510</sup> LL <sup>511</sup> FL <sup>512</sup> KY <sup>513</sup> Q <sup>514</sup> EH <sup>515</sup> FN <sup>516</sup> NP	--QRN	3159
Query_10012	3059	Q <sup>517</sup> KT <sup>518</sup> A <sup>519</sup> EL <sup>520</sup> K <sup>521</sup> AC <sup>522</sup> L <sup>523</sup> AF <sup>524</sup> AT <sup>525</sup> NGE--HS <sup>526</sup> Q <sup>527</sup> ELL--AK <sup>528</sup> EL <sup>529</sup> V <sup>530</sup> K <sup>531</sup> RV <sup>532</sup> KN <sup>533</sup> V <sup>534</sup> PS <sup>535</sup> AF--DS <sup>536</sup> LL <sup>537</sup> R <sup>538</sup> W <sup>539</sup> DI <sup>540</sup> RAM <sup>541</sup> Y <sup>542</sup> R <sup>543</sup> KK <sup>544</sup> F <sup>545</sup> T <sup>546</sup> NR <sup>547</sup> ID <sup>548</sup> Q <sup>549</sup> ML--EN	----	3124
Query_10013	3155	Q <sup>550</sup> CM <sup>551</sup> VE <sup>552</sup> L <sup>553</sup> QE <sup>554</sup> FL <sup>555</sup> DF <sup>556</sup> V <sup>557</sup> Q <sup>558</sup> ET--NL <sup>559</sup> AS <sup>560</sup> ETS--V <sup>561</sup> N <sup>562</sup> LL <sup>563</sup> SS <sup>564</sup> W <sup>565</sup> SK <sup>566</sup> RT <sup>567</sup> PT <sup>568</sup> Y <sup>569</sup> K <sup>570</sup> DP <sup>571</sup> I <sup>572</sup> H <sup>573</sup> V <sup>574</sup> W <sup>575</sup> DE <sup>576</sup> VT <sup>577</sup> I <sup>578</sup> FR <sup>579</sup> RV <sup>580</sup> F <sup>581</sup> SK <sup>582</sup> I <sup>583</sup> H <sup>584</sup> CG <sup>585</sup> Y <sup>586</sup> R <sup>587</sup> PR	[1]GAKA	3228
Query_10014	3059	Q <sup>588</sup> SIV <sup>589</sup> EL <sup>590</sup> N <sup>591</sup> EF <sup>592</sup> L <sup>593</sup> K <sup>594</sup> F <sup>595</sup> INS <sup>596</sup> NQ--A <sup>597</sup> FT <sup>598</sup> V <sup>599</sup> N <sup>600</sup> L <sup>601</sup> ak <sup>602</sup> ID <sup>603</sup> N <sup>604</sup> L <sup>605</sup> IV <sup>606</sup> W <sup>607</sup> T <sup>608</sup> DS <sup>609</sup> MP <sup>610</sup> N <sup>611</sup> M <sup>612</sup> Q <sup>613</sup> T <sup>614</sup> DP <sup>615</sup> PCT <sup>616</sup> W <sup>617</sup> DD <sup>618</sup> VI <sup>619</sup> T <sup>620</sup> NR <sup>621</sup> S <sup>622</sup> IF <sup>623</sup> FE <sup>624</sup> IE <sup>625</sup> E <sup>626</sup> K <sup>627</sup> YL--	[10]STTF	3240
Query_10015	3112	Q <sup>628</sup> LL <sup>629</sup> TE <sup>630</sup> LS <sup>631</sup> DC <sup>632</sup> L <sup>633</sup> NN <sup>634</sup> V <sup>635</sup> K <sup>636</sup> YAD--R <sup>637</sup> DF <sup>638</sup> YSSK--V <sup>639</sup> Q <sup>640</sup> T <sup>641</sup> V <sup>642</sup> TS <sup>643</sup> WK <sup>644</sup> K <sup>645</sup> CR <sup>646</sup> ID--DN <sup>647</sup> PL <sup>648</sup> M <sup>649</sup> T <sup>650</sup> Q <sup>651</sup> CL <sup>652</sup> AS <sup>653</sup> Y <sup>654</sup> RD <sup>655</sup> LY <sup>656</sup> L <sup>657</sup> HL <sup>658</sup> ME <sup>659</sup> KK <sup>660</sup> L <sup>661</sup> PEDS	----	3179
Query_10016	3126	Q <sup>662</sup> K <sup>663</sup> H <sup>664</sup> EL <sup>665</sup> QE <sup>666</sup> FL <sup>667</sup> N <sup>668</sup> F <sup>669</sup> ID <sup>670</sup> TYR--NID--E <sup>671</sup> yh <sup>672</sup> F <sup>673</sup> KN <sup>674</sup> L <sup>675</sup> TR <sup>676</sup> WL <sup>677</sup> K <sup>678</sup> RY <sup>679</sup> PD <sup>680</sup> IF--DS <sup>681</sup> VNI <sup>682</sup> W <sup>683</sup> DDI <sup>684</sup> IM <sup>685</sup> NR <sup>686</sup> SL <sup>687</sup> Y <sup>688</sup> ME <sup>689</sup> KI <sup>690</sup> V <sup>691</sup> SK <sup>692</sup> LQ <sup>693</sup> ETN	[11]MDTS	3207
Query_10017	3155	Q <sup>694</sup> RL <sup>695</sup> TE <sup>696</sup> M <sup>697</sup> NE <sup>698</sup> FI <sup>699</sup> TF <sup>700</sup> M <sup>701</sup> KS <sup>702</sup> ND--NF <sup>703</sup> SS <sup>704</sup> GP--V <sup>705</sup> N <sup>706</sup> LL <sup>707</sup> R <sup>708</sup> K <sup>709</sup> WR <sup>710</sup> K <sup>711</sup> FP <sup>712</sup> PK <sup>713</sup> LD <sup>714</sup> P <sup>715</sup> VI <sup>716</sup> W <sup>717</sup> DDI <sup>718</sup> I <sup>719</sup> T <sup>720</sup> FR <sup>721</sup> W <sup>722</sup> M <sup>723</sup> L <sup>724</sup> R <sup>725</sup> KI <sup>726</sup> ST <sup>727</sup> KL <sup>728</sup> Q <sup>729</sup> SV	[4]CEYL	3231
Query_10018	3127	Q <sup>730</sup> TL <sup>731</sup> TE <sup>732</sup> M <sup>733</sup> QE <sup>734</sup> FL <sup>735</sup> Q <sup>736</sup> F <sup>737</sup> M <sup>738</sup> DD <sup>739</sup> GR--NF <sup>740</sup> ES <sup>741</sup> VP--V <sup>742</sup> LS <sup>743</sup> LI <sup>744</sup> SK <sup>745</sup> W <sup>746</sup> STR <sup>747</sup> F <sup>748</sup> P <sup>749</sup> NE <sup>750</sup> VID <sup>751</sup> PI <sup>752</sup> N <sup>753</sup> V <sup>754</sup> W <sup>755</sup> DD <sup>756</sup> VI <sup>757</sup> N <sup>758</sup> RC <sup>759</sup> M <sup>760</sup> Y <sup>761</sup> LR <sup>762</sup> KL <sup>763</sup> LL <sup>764</sup> Q <sup>765</sup> RV <sup>766</sup> IQ	[13]GEDR	3212
Query_10019	3178	Q <sup>767</sup> KI <sup>768</sup> TE <sup>769</sup> M <sup>770</sup> QE <sup>771</sup> FL <sup>772</sup> E <sup>773</sup> IA <sup>774</sup> EEG--NF <sup>775</sup> SS <sup>776</sup> MP--V <sup>777</sup> LS <sup>778</sup> LL <sup>779</sup> N <sup>780</sup> K <sup>781</sup> W <sup>782</sup> Y <sup>783</sup> RL <sup>784</sup> PH <sup>785</sup> IL <sup>786</sup> HP <sup>787</sup> IG <sup>788</sup> V <sup>789</sup> W <sup>790</sup> DD <sup>791</sup> IV <sup>792</sup> T <sup>793</sup> NR <sup>794</sup> VL <sup>795</sup> F <sup>796</sup> MS <sup>797</sup> K <sup>798</sup> FL <sup>799</sup> Q <sup>800</sup> K <sup>801</sup> FEAM-	[1]AEER	3250
Query_10020	3009	S <sup>802</sup> LIV <sup>803</sup> IED <sup>804</sup> EAL <sup>805</sup> I <sup>806</sup> RL <sup>807</sup> NEDG--SD <sup>808</sup> ICK <sup>809</sup> N <sup>810</sup> eh <sup>811</sup> FG--F <sup>812</sup> VE <sup>813</sup> N <sup>814</sup> W <sup>815</sup> K <sup>816</sup> Q <sup>817</sup> P <sup>818</sup> PT <sup>819</sup> GE <sup>820</sup> EG <sup>821</sup> M <sup>822</sup> VL <sup>823</sup> WS <sup>824</sup> Q <sup>825</sup> HT <sup>826</sup> M <sup>827</sup> Q <sup>828</sup> NAI <sup>829</sup> T <sup>830</sup> SL <sup>831</sup> ML <sup>832</sup> NR <sup>833</sup> AHD <sup>834</sup> Q	[1]----	3079
Query_10021	3434	H <sup>835</sup> RL <sup>836</sup> TE <sup>837</sup> LES <sup>838</sup> L <sup>839</sup> T <sup>840</sup> L <sup>841</sup> T <sup>842</sup> GD <sup>843</sup> R--RR <sup>844</sup> Q <sup>845</sup> V <sup>846</sup> vs <sup>847</sup> TAD <sup>848</sup> L <sup>849</sup> AK <sup>850</sup> WR <sup>851</sup> RR <sup>852</sup> PD <sup>853</sup> F <sup>854</sup> IR <sup>855</sup> DAC <sup>856</sup> D <sup>857</sup> S <sup>858</sup> ERT <sup>859</sup> V <sup>860</sup> AV <sup>861</sup> RL <sup>862</sup> LC <sup>863</sup> ANK <sup>864</sup> M	----	3496
Query_10022	3229	Q <sup>865</sup> KI <sup>866</sup> YEM <sup>867</sup> REF <sup>868</sup> L <sup>869</sup> QS <sup>870</sup> IK <sup>871</sup> Q <sup>872</sup> EGL <sup>873</sup> Q <sup>874</sup> TK <sup>875</sup> S <sup>876</sup> D <sup>877</sup> IV <sup>878</sup> p <sup>879</sup> q <sup>880</sup> V <sup>881</sup> IK <sup>882</sup> SI <sup>883</sup> HK <sup>884</sup> W <sup>885</sup> MAR <sup>886</sup> V <sup>887</sup> PS <sup>888</sup> YAF <sup>889</sup> DT <sup>890</sup> IV <sup>891</sup> V <sup>892</sup> W <sup>893</sup> DD <sup>894</sup> IL <sup>895</sup> T <sup>896</sup> ARN <sup>897</sup> LY <sup>898</sup> ID <sup>899</sup> LY <sup>900</sup> K <sup>901</sup> F <sup>902</sup> RL <sup>903</sup> K <sup>904</sup> DEF	[1]----	3301
Query_10001	3187	TSIDAKIDPPG-- --KQRVV	MSLAQVDAASLQGNHS <sup>905</sup> VARK <sup>906</sup> LK <sup>907</sup> LSA----KD--LVK <sup>908</sup> Vv--dS---EELS <sup>909</sup> VL <sup>910</sup> FT <sup>911</sup> H <sup>912</sup> SH <sup>913</sup> A	3248
Query_10002	3267	GEEETALEK---[1]ILRERLW	LELKMADSA <sup>914</sup> CQ <sup>915</sup> GN <sup>916</sup> FS <sup>917</sup> V <sup>918</sup> AQ <sup>919</sup> K <sup>920</sup> HL <sup>921</sup> KAT----HG--SLSA---H---EDL <sup>92</sup>	

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	3249	NANKRKC	aSG-SEPSERLNLLVEVKKR-LDKCS	ESSTLKLN	--AALLCDQHLLQGE--	HLVTMATILQNVKDI	3316																								
Query_10002	3328	QMNQQKI	-LT-IELKEQVMTLLTTFDP-LAKLS	NTPALTDG	--YCLTRRHHTLTSKSY	DLVAIA-IQKGGADIL	3395																								
Query_10003	3315	RLHKKKA	-SL-LTGADKLNNVLTTFEQ-LDKFS	SCRVFEED	--LAVAKQHHTLTSLSY	EIVAHALIQEGGA-IL	3382																								
Query_10004	3282	RLSHCRS	-RS-QGCSEQVLTVLKTVSL-LDENN	VSSYLSKN	--ILAFRDQNILLTGTTY	RIIANA-LSSSEPACLA	3349																								
Query_10005	3301	RLSHRRS	-QR-QARPEQVTVLKS IPL-LVEIN	VD-YLNTH	--TRILRDQNTLLGSSY	CILADA-LNKPEPTCLQ	3367																								
Query_10006	3233	STHQKKI	-IS-NDWSE--DGINGLLTT-IGKFN	DSDVLSKK	--PDLYRRHHVLTATSY	ELLAESCLAELPQSL	3299																								
Query_10007	3259	KLHNKITE	VT-MDASATLSPVLGALGL-LERRR	DHKLVKES	--NVLTRRHVLVQGNIF	DSVTSLSLSDPDFSSV	3328																								
Query_10008	3165	AVKSRTA	-----YLMQALNLFDL--KRIS [ 2 ]	EVNVTEGS	-----SADQFALIGMLF [ 2 ]	--VAETVVASSENTLA	3224																								
Query_10009	3326	KIYCLKAK	NS-TTQVETLDRFIKALKF-VESKS	DEDSIKNN	p--DNLQRYKRLHGNIQ [ 5 ]	-----LDQKLGSLN	3393																								
Query_10010	3255	KLNLREA	qNT--IDTEKKAELLLALSeFKKVK	--SSDRDDi	-----KMRIIESETY [ 3 ]	-----	3306																								
Query_10011	3227	RLQCLRA	eRD-----ESANNRIKKLQQLVYIE [ 1 ]	JERRKLSKH	p-----	EFLPEFQMEQADVNR	3280																								
Query_10012	3179	KAKWLQA	qLK-NTTDEKLLGLDWS-----	----VIDD	--QRFE-QYPMLHAETL [ 4 ]	VMAAGILQQLEQQPGV	3241																								
Query_10013	3310	AYHLHEF	----NLGVREQKHFRFARDLDRLN [ 4 ]	DLPPSKS	ahSLYKAEVLLSLGKIV [ 9 ]	KVLWDHIIAEVDAPDA	3392																								
Query_10014	3219	NTHLNKC	-KQ-ANSKESLESFF--ESM-VLREI [ 5 ]	ISDTFKNN	gkNYLYQEQRVLHKSFS	KFLIESLMDHMNQTFD	3292																								
Query_10015	3236	ETHLLRG	gGK-QT--DRLKYLVEAWAKYLRGVS	SMSILDQN	--PRVEIQYKLKES-- [ 3 ]	LEICRAIQMDGDDWTE	3304																								
Query_10016	3278	D--LKKA	-KN-QNYPEQLDTLLCAHKQ-LAKLD [ 5 ]	KLDVSTMN	--HYLQSE-----	--IIEMAADIVKYNDV	3339																								
Query_10017	3300	KVRLDRI	-HQ-DPSPDMEAVISLTCQ-MDNFK [ 5 ]	SGKGAISS	--KPFISQHHVLLGKLL	NKMAEFATPEDGMLA	3373																								
Query_10018	3279	KVHCDKS	-LL-TPPDERLKLTLTANSQ-FEKVK	DLSVILEN	--TTLMREHLLLRSLFL	DMTSKI-LAESDASIL	3346																								
Query_10019	3314	EVHFRKI	-QT-LTAVGAVETALLVAGQ-LDKFS	ASKILBIE	--PLRGIQHHLRSKTF	DILCKA-MLDGGVVIA	3381																								
Query_10020	3130	KLNVIRMI	eVF-----VSHVLKLI---DRQE [ 2 ]	SSDSTSik	----LITKYETATKLF [ 3 ]	EIMHMMKIATARDQVA	3193																								
Query_10021	3570	KAWTTA	ALSLd-----TDRnFDAMS ---VAEKv	--AKVRRQLISPSPTW [ 1 ]	ELLAAGLCDRASALL	3625																									
Query_10022	3372	DTMNLNF	aDKg----DKLQIKVVFdskMQESF [ 1 ]	EDQQFOED	l--RTEIKCILLKQKLE [ 4 ]	NIIFQGVSAEKEIKHL	3444																								
Query_10001	3317	DA-IDQK	LLEKITESVGPALSHQE-IVA	SLHEAA	YNSFKSGVGVA	A-----SQPESVVCVNKAYMSLVHFCDQ	3382																								
Query_10002	3396	TGlrNV	GKLDKLLGLVGQNA	TRPEQ-ICST	LYSKAYS	ELKQAVSLA	DKDDAV---TSQESRSMTEAYMALVSFCDR	3467																							
Query_10003	3383	TglkD	TRHAQLLTAAGVSPD	TSVDQ-TISS	LLQKGHACLQKAVKLA	SDEEKARQD	TSLDKSGIVDAFAMAMVGFCDK	3457																							
Query_10004	3350	EI--EED	KARRILEL	SGSSSEDSK-VIAG	LYQRAF	QHLSBAVQAA	EEEAQPPSWSCGPAAGVIDAYMTLADFCDQ	3422																							
Query_10005	3368	RI--KED	QAERVIEHSGSSSEIEIQE-VVIG	LHRKTL	HLHLERAVRKA	EEEVQLSAGEHVD	IAGVIEAYVNLANFCDR	3440																							
Query_10006	3300	SS-kTC	DKLKNLLKTDNLDK	TQVN--IVS	LLQKSHQELKSAIS--	--DSVVTTRSRSFTE-ADAQI	LALSKFCDK	3367																							
Query_10007	3329	SP-rDI	QKLVAAASCSMQNADMQE--IING	LVDRG	YQYKEAV---	---DIYEEDSSNDKST-VE	THMALAKYCDK	3394																							
Query_10008	3225	EF-----	CLSRWISKTD	TADISELVMT	MCYQAFRYLKLSCQYS	-----SQLNIESPKLTPLEPL	MELSSFCDR	3285																							
Query_10009	3394	VM--ENL	KKNGLTSLSLPT	TRLSS-----ELF	SLAYQSYSQSIQLY [ 1 ]	-----SNSNNNDTHKS	KSAYLQFGNFCDT	3459																							
Query_10010	3307	-----	VSELDKRV	PNRLLKIKHNK	NYFIERGYNLFNNATKEI	-----SLQKPSMHAKILV	KFAKFCDY	3366																							
Query_10011	3281	VA-----	WATLSSQSGEGATSQDITAC	PALGSAFK	YJSTALDND [ 9 ]	-----SLSPHD	TAKALLKPLFCDE	3349																							
Query_10012	3242	LTKlP	AAEVKKLADRIGAPDCSQDT-LM	QNLQYGLSSLKTAAEVV	TP---RG---NSE	QAVAKAYLQVAKYCR	A	3309																							
Query_10013	3393	TRagE	KCIVRAFVLLETAQ	RAPDpKALPVG	KRRAGQFA [ 21 ]	KPLDVLFP	SATAYGSETLNAFVRLAEFCAD	3489																							
Query_10014	3293	DS----	GLQGIITE-----K	HHVLT--LSD	YVKNDDINDFEKTISS	LMKHGVDNLLKIKD	SLKGS--LELATYCDF	3355																							
Query_10015	3305	DnqyM	QVLSKFFENANGKES	WQREL-LSCS	YKSLQLAVNCAEGLV	KSDALNQD--DSK	SDSAHLALVKYCE	3376																							
Query_10016	3340	HA----	EKIQLTDFFSIQSKS	LE--IDL	FCKN---YEYK	CRALD	IIQEVEQKMIEDFQATIGETYP	KFANYCNK	3406																						
Query_10017	3374	LSaL	SDNGNTKLSNLIGGTP	STSKaISEGL	WLQALSFNAGVRS	TTLNPL---TKDP	VPTVQSL	ELFGCDN	3445																						
Query_10018	3347	DL--L	NEKESLYQLANVNQ	HDSKEIAAKL	TSACALHQAVKAS	KEERG--G-----	SSYIEAMMMVNFCD	D	3412																						
Query_10019	3382	TL--E	SQAKNALVKLCRIDETES	QkVVSD	LMTRS	FVLSQSAIKAA	KNAEG--KCNTTQTSG	TVDAMIAMVNFCDN	3453																						
Query_10020	3194	VgyLE	AKTFAKAAAFYAFND	LDLDDL--SKK	FYRS	SLDRFKDSCRR	--DAVLSKQSQDDAP	PAFLRCLRTYIEFLND	3266																						
Query_10021	3626	RI-----	QLASSPSTGQAN	MRaEAEAFK	FKSALSNLQSAASKA [ 9 ]	-----SSSQ	CHLELIRFCDS	3688																							
Query_10022	3445	QR-----	YQkDCT	YLD	RSFKTL	DQVIKKV	---QTA	AQFDDQMKMTMAKTYFKFAQFSDK	3496																						
Query_10001	3383	V-LRQ	RENDEISG	TGC-E	SEEDFP	LALVTSLLSAMKNSK--DARQL	FPRL	LQIIE	RFP--QT	RE	VFIRKSR	3448																			
Query_10002	3468	M-LRL	KESDEV	PMTYL-P	QTDNF	SQT	VVYHTL	KAMHYS	SQ--EARQR	FPRL	LQ	VEACP--DT	MQ	LMIKEVS	3533																
Query_10003	3458	I-LRH	QEDDE-EG	TCL-S	SIPGL	AAAVVSNL	QALRHDS	P--EARQR	FPRL	LQ	VEI	YP--DT	MD	LFIEKVS	3522																
Query_10004	3423	Q-LR	KEEENAS	VIDS	A--ELQ	AYPALVVEKML	KALKLNSN--EARLK	FPRL	LQIIE	RYPE-ET	LS	LMTKEIS	3488																		
Query_10005	3441	C-LR	KEES	A	INFS--EQQT	FPAC	IKAMK	YNSK--EARLK	FPRL	LQIIE	VKLYSA-ET	LD	LMKKEVS	3506																	
Query_10006	3368	F-LRM	KENE	EHIKE-K	VIENY	GETI	IICLL	KAMK	LSM--EAIER	FPRL	LQ	FVELYP--DT	MP	TFIKMVS	3432																
Query_10007	3395	Y-LKL	RE-ED-DS	LTG-E	SIDS	YPEAVV	TSLLR	CIELDH	M--DARQR	FPRL	LQ	VEYYP--DT	MQ	GFINKVS	3458																
Query_10008	3286	I-LR	KNGT	SLVNS	ECAC	-----VM	RSILKSMKNGSH-LARQL	FPRL	LSVME	KYPE--S	EL	EFQAEID	3345																		
Query_10009	3460	E-L	QRRG	D	S	EDSNL	RVL [ 1 ]	-----VVNS	INNAIKQ	IP-GAIEK	FPRL	LEIISQ	YEK--A [ 2 ]	E	-FVERTK	3519															
Query_10010	3367	F-LR	KLDF	TEEQ	TQIFD [ 1 ]	--AQYAS	NVESILQAMENGSK-EATEL	FPRL	LQIIE	KYHN--T	Q [ 2 ]	-FK	KEVK	3432																	
Query_10011	3350	L-LK	LSQ	PEERS	QMTQH [ 6 ]	QASD	LSLLVVRHMLKALALD	ST1QAQ	HAVARLLS	LVGFSSA--T [ 2 ]	-E [ 2 ]	RFLDN	VNS	3426																	
Query_10012	3310	-----	HSNTS	REDN--	----FG	TDVTSILRAMRYGSK--DARQL	FPCL	LQ	LPGLDT--SL	SE	LFHKES	3364																			
Query_10013	3490	V---L	QRIAMP	KDSMGR [ 1 ]	----KQV	FVTALAMRSGNV-LMRQ	FPQL	NLNG	RPTds	SDH	LQ	RVLAS	AT	3552																	
Query_10014	3356	Y-LR	IAEND	NDFE	KLTL [ 4 ]	IRSQ	FPFIVTEQLYSIRSGSL-EARTR	FPRL	LQIIE	LYPD-EC	LE	KFVSSSE	3427																		
Query_10015	3377	CvIK	WDH	V	D	SQ-----YSK	SLMVSVLRAMANGSR-DAH	FH	PRLIN	LMNE	DA--SL	LS	SFKEEYE	3435																	
Query_10016	3407	I-LRM	QEE	----E	EKTL [ 2 ]	SFET	YPSVLTNSLLNALKNGNL-NAIKL	FPRL	LQIIE	VYPD-T-	IP	EFI	AKSK	3470																	
Query_10017	3446	A-LR	QED	H	DEQ	STGL--DTKS	PPSIIINH	TLTAMSLGSS-KAREL	FPRL	LQ	VEAHP--DT	MD	LMINKSS	3510																	
Query_10018	3413	I-LR	HED	G	RANQ	PKGLI	DTKI	FPFVIRY	TLEAMAYNSM-DARQR	FPRL	LQIIE	RNP--NT	ID	GFMRKSS	3479																
Query_10019	3454	C-LR	KEV	S	V	APVE--V	DTK	LPFGIVVYMLKSMCYSSK-EARQ	K	FPRL	LQIIE	VHKYP--DT	LE	AFVKKAA	3517																
Query_10020	3267	LLY	KRQ	LEK	L	ANF	ADSKV ---L	LKIL	LVNR	GMAGDR-EC	ANY	FPGL	CDVVS	PYPD-IV	AE [ 5 ]	----LT	3331														
Query_10021	3689	CILE	AAP	N	Q	P	Q	FRAAL [ 3 ]	-----AV	SVTSAM	RCSQ--EASIR	FPRL	CLL	GESAD	DEA [ 4 ]	AD [ 5 ]	RFSRQMA	3762													
Query_10022	3497	I-LR	K	M	E	N	D	Q	C	R	V	I	T	N	S [ 6 ]	SEBQ	VARD	IKYGL	LSLN	NHEYQ--NASDI	IPRL	L	I	T	S	KYK	--H	VQ [ 2 ]	-FI	AYSK	3570

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	3449	D---IPAWMFLGWSQICAILDKPE--SSAIHEILLNIADIYPNALVFPFKLSNHNYKF-TSDAVSSKSKSVCDRHLQML	3522
Query_10002	3534	D---VPCWFMFIGWISQMVALLDKPE--GPAVHGILTDIAYEYPQALCFPFKISSSEFFTF-DESNTRGRKNASAAARKIGAIL	3607
Query_10003	3523	H---VPSWLCISWLGQMVALLDKPE--AVAVHGILTRIAQEYYPQALVYPLKISSENFEF-EDSVTGRNRQEAFTSLKKTLL	3596
Query_10004	3489	S---VPCWQFISWISHMVALDKDQ--AVAVHVEEITDNYPOAIVYFPFKISSSEYSF-KDTSTGHKNKEFVARIKSKL	3562
Query_10005	3507	S---VPCWQFISWQMMALDKKE--AVVVOHIEEELADSPQAIYVYPLMISSESYTF-EQTARGNTKNEFVERLKNKL	3580
Query_10006	3433	Enr1FPLWMLLWLNQMMALDKTE--SPAVQPIILQIGTEYPQAIAYPPFRISCEGFEF-EDSAQGGKKNKAVKKLASIL	3509
Query_10007	3459	---VIPCWFMFTWIGQLVALDKPE--APAVHILIRAREYYPQALVYFPFKVSKYCYTF-ERSSEGENQAVVERLDPSL	3532
Query_10008	3346	S---VPSWMLFSLWMMQMLSAKLR--SNFKYSILERADDPQALYPLMLK-----SATGDVGETDFYRLQKLL	3410
Query_10009	3520	Q---IPCWQFIRWQSQMFPYLDTPQ--GPYIPLIQEIASNYPOAIVYFPFKISSSEQFGP-----TAKKLTAKLEQTL	3586
Query_10010	3433	Sie--SVWKFIRWIPQIVALDQSI--AHCVPFILFKLAEPFGSLYPLTISSEYKFDENNPLTRENKAKIQELKQMI	3508
Query_10011	3427	S-----WVPVAWIPQMLANIEEDS--GRVLIKVLESIAKSPQALYFPFVSEDTDFGA-----IGRQRTQHLKILL	3490
Query_10012	3365	A---VPVWMLGWNQQLACLGPV--GPVLHDIILRLAETYPQSLIYAFQLSREKYNL-E-SGHVSL-KALVDQLASLL	3436
Query_10013	3553	TS---VPCWMLNWIQMLAIDLDAPE--APVVOQLKTIASSYPQALCFPVMLESEKLFK-SGAASSTQRTFFTNLVRMVT	3627
Query_10014	3428	D---IPSWMLFSLWQMTALDKPQ--AKAIFQIENISIEYPQALVYFPFRMSTENFKE-SNELN--EQSDFVQKISRHL	3499
Query_10015	3436	K---VPIWMLLWLSHILYVDPKVP--GPAQPVVERLAAYEYPAQVYFPFGISKQOYNF-S-SHEGKLAKAMCDKVEALL	3508
Query_10016	3471	V---VPCWMLFSLWINQLLTLDKPE--ATAVYIIEIGTNPALVYFPFKLSYEMKVF-PDTLQGGKKNKEFVMKLSL	3544
Query_10017	3511	T---IPCWFMFIGWINQMIALVDKPE--SRAVHDILIRADTYPQSLYAFRMSSEGYKF-GHN--NSEGKELFELKKEKL	3582
Query_10018	3480	A---IPSWMFIGWISQMMALDKPE--ARAVQELIFRIDEYPAQIYAFKVSNEQYTF-EETSVGRQNTAVSKLSKLL	3553
Query_10019	3518	D---VPCWFMFIGWINQMAIDLKPE--GKAVHGIVSEIAKRYPQALWYFPFKISSSEHFVF-ENSEDGKANEEAVKRRLSSL	3591
Query_10020	3332	D---VPLWTCLOWSAQLMALNGDI--GMTIVAILKLAERYPVALYFDFMVTCSRSD-----IFKVDLRRLVLL	3398
Query_10021	3763	S---PVMWLSLWPLQVSLDKPDqgGSVLPLESVCDIYPALVYFPFLARANYRF---ETGDSGARQAARLRQRL	3836
Query_10022	3571	Et---PTWYFLRWIQLIVAVVNRPE--SSVIEEKVIQIAKKYPAQALYFPFKVSESNVKV-NVLDSDVQKTLFQKLSKY	3644
Query_10001	3523	SK---NPLIDRLIFSELEMMHEPAVLIKDWIMDM---KNAKPGGSSqSLSMAYVKERYQELYKELL-- [ 4 ] EDGSLYSDS	3595
Query_10002	3608	ES--i-PLLDDFICALQELNNEPEFVKDADDV-mKPLLNv----aKRDKNKIRQACHEIYMNLINH [ 7 ] GAGSSQDSQ	3682
Query_10003	3597	SV--iYSLVDKLLIGGLEQLSHDPQLFKDWGDDAdLKRLLKVPDKKrtKEDKVTIKRLLYDEMYSLLDF [ 7 ] SMSSLFQSG	3678
Query_10004	3563	DQ---GGVIGDFINALDQLSNPPELLFKDWSNDV-raELAKTPV-----NKKNIKMYERMYAAL--- GDPKAPGL-	3625
Query_10005	3581	DK---GGMVVFVINSLEQLSNPDMFLFKDWDVNF-rSKLEKKE-----NKTQLKNLYSEMYSNL--- ANAGAPGI-	3642
Query_10006	3510	Tedkv-PLVNFKFISSALDQFGQPPEMLFKDWDTDI-l-KMMLQSK---NRNKSIMEKYKEIYNMLIEY [ 7 ] STQLSYMTQ-	3585
Query_10007	3533	S--ki-PLVETFINALEQFGQPNALFKDWTGEM--AKLLHVK---QRNTDAVQRCTQMYSSLM-M [ 7 ] VSDQIFDTE	3605
Query_10008	3411	RKt--KIVKEFIDALNLLSHDPDIHAIDELESI---ILNISNNS---NEARKEAIKAVYSFVTF-----	3465
Query_10009	3587	KN---PLIDTLISEFERLTHPEHRFKDYMETV--KTMKSTPIN----TSEIQKLGVEIYNDCFNP [ 4 ] -----	3647
Query_10010	3509	Kcd---VDTFVTELHRLTDSEIHFNCWINPV--YNLVKNRYSG--VDKTAIEIKASFQKMSFLKLP-----	3568
Query_10011	3491	RN---PLEGLVRALEDLTFPQQRLKDGLLQI--LNLFLPTN---DSERAQLVFNIEYQDCLDT-----QSL	3548
Query_10012	3437	FT---DPVIDIFLRAFSCLGQPRKMLYHFVEKL--EQLLNAK---RPNCDSLHHMVDTMKELFP- [ 1 ] SYGKIHGAL	3503
Query_10013	3628	TDh---LRRFVELKCLTNPVMAVMDWLEEQ---DQAIK--KRDQDALTAFKHLKSELLV- ----TSSTT	3685
Query_10014	3500	NT--qLPLVEQFITALHLNTPKLFNDLIMDI--NPNLK----K-----NDVIMKLFKDLNNLID- ----LEGK	3557
Query_10015	3509	SS---HKLLHQFVSAMSLMVVPSACKDAIQNV--LQEKDKQ---KME-EGLESEIVSEFLKVQNR- [ 1 ] SKGNAEEKG	3574
Query_10016	3545	S---qLKLVDVFNALDLSVPELLYKEYEIKETI--LQMKKDKGDK---NEIV-KLYKLLCHELNFN- ----YNPR	3604
Query_10017	3583	QS---QLMDEFIAALEQLNSPELAWKDWYDNB-lKPLVGSVTPV--KRDKNAIKRVWEQMYSQLL-- [ 6 ] GDKSLYGGG	3626
Query_10018	3554	-Q---LPLIESEFVRALEQLSQPDLVFKDWDV-E-vKGLKKS--G---SIAESIRICYQLKRLDGLLVD [ 7 ] ASSLTSSEL	3656
Query_10019	3592	-E---HPLLEDFFIAALEMLTNPVEMVFRDWCCEP-mKQLLEDQIKR--NDSQAVKQCFGAMFQ--LLMD [ 7 ] GSQ-TRGT-	3664
Query_10020	3399	AN---PVMKFFVAALGLIHHPELRLKEGLRKT--ANLLDENRTS-----DARQTVLELWRDCEFSF-----	3453
Query_10021	3837	RR1--PLVEEFIAQMGVCLPACTALDWLSSA-----KRSIQG---TERLRELHAEIADQLFFD-----	3890
Query_10022	3645	Ekt--FRNLNAWAEALDCLVYPEHRFKYWLQI--NDIVSDDNMR--GIQKAKLKDIHIMMAYDLTCT-----Q	3707
Query_10001	3596	ADC--- [ 1 ] MGPYRSDFVKQFISQFDEYFGKMGSKIPGL---TFMKVLSWLS [ 2 ] SEKVKKAFSNKQPRK-LKEYSQWLA	3665
Query_10002	3683	SSP--- [ 3 ] FGNFRRRFAQEWQKFEAAGFKDGTVIDMQLKTFSE----- [ 7 ] EMAAKRRD---PPGN-LKEYSTWFS	3753
Query_10003	3679	SSDLES [ 4 ] FGPFRRRFAQFAKECEKMFQKDGSKLLDMKEKEPFMKLSGLQ [ 8 ] KAFPPEKANCYAPGL-LKDYCPWLS	3763
Query_10004	3626	-----GAFRRRFIQTFGKEFDKHFQGGKSLLRMLKLSDFNDIT--- [ 1 ] MLLKMNKDSKPPGN-LKECSPWMS	3688
Query_10005	3643	-----GAFRRRFVTGFKKKIDQHFGKDGGLLLNMLNLFKFNVAV--- [ 1 ] QIM--ETQKPEPPGN-LKEYSPWMS	3703
Query_10006	3586	TTN--- [ 2 ] MGEYRKKFADTFKKDFDIAFGKNGEKLASMSIK-----KGRGKLVPPLS-LHEYCQWLA	3643
Query_10007	3606	PDS--- [ 2 ] IGQYRQKFAKAYCKDLSAFQKDGKTKIMTSEEAADGIFKRM [ 2 ] SKGKTGHPNLVPEK-LGDYSKWMS	3679
Query_10008	3466	ADIRRD [ 1 ] LGSFWKCYFKKYLPMQEIFSVLDLCAFDNKAENYLRLCSSFL SELRTPYIPECCPN-LNEYSPWLA	3539
Query_10009	3648	-----GDYNLKFKEWDHHSYQHFSGKDGSLAKMDVKKLVELVADMS [ 1 ] -KMAKTMKPSSTATMkLDFSNWL	3714
Query_10010	3569	----- [ 3 ] HGIFAKEFVNRHSEYLFDICRDRGSLILARMSPNDFKSKIYDYY [ 1 ] KNIKRQDRKAEQGVNLLKYSYAWLA	3640
Query_10011	3549	KDPVRQ TGEYNLKFARDWASLIVKELGLNGSKLRSMDAKQFPSSVISGLF -EKLSKNDISAEKIS-LGSLSKWLA	3620
Query_10012	3504	FKDIQE ---YRRLKLEK--AESNIKCGM---KVKLKDICFKEKLYI- IQDIK-----PVQ-LKDYSPWLS	3560
Query_10013	3686	ASDREV -GSAWKAAKAKARNFMDQFGKDGSKLLRMDAKAYAKVKVVK -AKLREYDTPKPPRT-LKEYSPWLS	3756
Query_10014	3558	IDFCGR ---LWTDPSNFIKPLFMQFGQFRLLPQLSELEIKNLQNIC SQVNEFCQRKPPSGN-LGDYSAWLK	3627
Query_10015	3575	EAYTRQ ---DRAKQOM--AEAFKAFGENLKKAKSLIRDIRQRELGSMD [ 1 ] VLQLSPEEMRKLPRQ-LKAYSWLA	3643
Query_10016	3605	DKMKER --KLHLKVAKVCSSVVFSEFGKDGKNLISMPKLEASVKYETLL KKITTMSEMNK-EAK-LRDFSDWLS	3674
Query_10017	3657	PSLQK- -GARQKTFQAAYAKNIEDKFGKGGAKIISMSGTQLASAFNEVY NKMRESLNKFSSSAL-LKDYSPWLD	3727
Query_10018	3627	IEF--- -GSHKRRFIQHFQKVIQYFGKDGSKLQSMKASDFEKIASRLI [ 1 ] EISGKIKNR---PEN-LKEYCPWLS	3693
Query_10019	3665	--ES--- -GPFRTKFAQEFANEVQRIFGKKEGKLVGVKMTDFSKWC-RDI [ 1 ] QKMKVFRNRTPPPGN-LKEYSPWLS	3732
Query_10020	3454	--QRQL [ 4 ] IGRYNRWRSRVAKRDVEKIMKDGKSLAKSLRDTVSAAREWINNHF SVTPGRYGITKDMKAhLGDFAEWLQ	3529
Query_10021	3891	PDGPPA [ 2 ] HGRLRREFAAHRHREFLALP TADRLVGRSEAEALQRAVQKLV KRVDRDERSSSVATARLADPSTWLA	3966
Query_10022	3708	KDLLGD [ 1 ] VGSYRKFALWQVFAKEFGSNLNRVDRMSNSEIKSVQNLN NKLNQSIGNLQGYEKLSTPSEWLD	3782

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	3666	EFQ	AGNHEA--S--LEIPGQY	TGE-VCPLPEYHAKISSFDSQVTVLRSRLMPKISIVIRGNDEKEYKFLVKGGED	3734
Query_10002	3754	DFQ	QLSYDR--E--LEIPGQY	GGK-SKPLPEYHVKIAGFDEQVKVLSMRKPKRIVIRGNDEHDYSFLVKGGED	3822
Query_10003	3764	DFN	PLDYDR--E--LEIPGQY	TGR-VKPLMQYHAKIAGFDQRVKVMSLRKPKRVTIRGDDERDHPFLVKGGED	3832
Query_10004	3689	DFK	VEFLRN--E--LEIPGQY	DGR-GKPLPEYHVRIAGFDERVTVMASLRKPKRIIRGHDEREHPFLVKGGED	3757
Query_10005	3704	NFK	PQFLGN--E--LEIPGQY	DGK-CKPLPEYHAKITGFDERVKVMSLRKPKRILIRGNDEKEYPFLVKGGED	3772
Query_10006	3644	DFN	PNRDSR--E--LEIPGQY	DGL-SKPLPEYHVKVAGFDERVLMVSSLRRPKRITIRGNDEKDYHFLVKGGED	3712
Query_10007	3680	GLA	AQDMDL--S--LEIPGQY	TGL-SKPMPEYHVKIAGFDSRVLVLSMRKPKRITINGTDMKDHPFLVKGGED	3748
Query_10008	3540	GFR	PHPMDV--F--LEIPGQY	SGD-RRPPEYHAKIAGFADIEIRMSLRKPKKICIQGDDGRDYKFLVKAGED	3608
Query_10009	3715	EFD[1]	SNQLTE----LEIPGQY[1]-GT	GT-SKPLPEYHAKISSFDPNLLVMGSLRKPRIKIHGSDQDYFLVKGGED	3783
Query_10010	3641	EYS	SSNVEK----eLEIPGQY	DGLLSIPDPSKHKIASFDPKILVLSLRKPKRITILGTGNEYLFLVKGGED	3710
Query_10011	3621	DFD[8]	SRDPDK--nEmqLEIPGQY	NGS-SIPNLSTHVQLLGFQVLIKLSKQRPKILTRGSDENDYKFFVVKGGED	3700
Query_10012	3561	EFQ	GGKHCFF--Le-LEIPGQY	TGE-SKPLIQHVVKIAGFKQTVVPMKTLRPLQVTFIGSDAKEHQYLVKFGED	3630
Query_10013	3757	KYR	SPGAVF--D--LEIPGQY	GCT-EMPIPEEHAKIEGFDEKLIMASIRRPMLTFRGSDEREYRFLVKAGED	3825
Query_10014	3628	GFK	R-NITK--D--LEIPGQY	NGM-EKPMPEYHVKIESFDERILEMSLRKPKCITIRGDDQREYKFLVKCGED	3695
Query_10015	3644	NFQ	ASKYSE--V--LEIPGQY	TGM-SKPLPEYHVKISSFDENVLLMSMRVPMRIIRGDDEKEYKFLVKWGED	3712
Query_10016	3675	EFH	SIKYSE--D--LEIPGQY	TGK-RKPIPEHHVKICNFDQKVDILSSLTKPRCLTIRGNDQKDYKFLVKSGED	3743
Query_10017	3728	TFR	QEDYLO--V--LEVPGLY	DGR-SKPLPEYHAKITRFDKVLILSSLRKPKRLLTIGDDTREYMFVKGGED	3796
Query_10018	3694	NFQ	AANFTY--S--LEVPGLY	TGR-SKPLPEYHLKIASFGERLLVLSIRKPKRLTIVGQDDEKEYLFLVKCGED	3762
Query_10019	3733	RFQ	SNNFTH--T--LEIPGQY	TGK-TKPLPEYHVKIVGFDEKQVLSIRKPKCLTIRGDDEKDHMFVKGGED	3801
Query_10020	3530	EFD	HSSCIL-----ELPGQY[4]-GP	GP--PDPATHVRLSFDMSLGVLSAQKPKKLTACSDKKNYFLVKGGED	3597
Query_10021	3967	EFA	ATEHGD--Ay-LEVPGLY[6]TAS	ASLAKPQPELHTRIAFGDTRVVMSSLRPHRLTIRGNDQDRPFLVKAGED	4043
Query_10022	3783	SLD	FNDFYNpsEy-LEIPGQY	EGLkSEPLIEKNVKAISVKKTLILLGSIRRPKRITVHGSNEKDYHLLIKGGED	3855
Query_10001	3735	LRQDQRVQVFNLMNEIYANDANCHARR--	LQKTFQVIMPMTPRLGMIIEWL	KNSCTLKDFLRQCCSDV	SWEQI 3805
Query_10002	3823	LRLDQRVEQLFRIMNSILEQDSTCSQQG--	LQLVTYEVIPMTPRLGMIIEWA	QNTTTLKEFLHLAMTEQ	EGNAY 3893
Query_10003	3833	LRQDQRIQQLFDIMNDILTDPAACSQRQ--	LRIRTYGVIPMTRSRLVGLIEWM	QNTTPLAGFLENAMTTR	EQEAF 3903
Query_10004	3758	LRQDQRVEQLFQVMNGILAQDSACSQRA--	LQRTYSVVPMTSRLGLIEWL	ENTVTLKDLLLNTMSQE	EKAAY 3828
Query_10005	3773	LRQDQRIEQLFQVFMNIIILSRDAACSQRS--	LQKTYQVIPMTRSRLGLIEWL	DNTCTLKDFLSSNITEE	TQSRR 3843
Query_10006	3713	LRQDARIEQLFVIMNKVFDKDPVCRHRK--	LHLKTYQVIPMTPRVGLIEWM	NNTTPLKDFVYVYALTEQ	ERKYY 3783
Query_10007	3749	LRQDQRLEQLFYIMNRFVSDQDDCKPRQ--	YKLRTYQVIPMTPRVGLIEWV	NQTKTLKEVYVYGTMTQE	EK--- 3816
Query_10008	3609	LRQDERVQQLFDFLMMGILQKQHC SRLK--	ARIRTYRIVPLSIKLGLIEFL	PNVVPLQQFMEESLRK	QYQTI 3679
Query_10009	3784	LRLDQRIQQLFFVMNEILKRPATSKRG--	LHIKTYQVVPMTGKVGIIIEWL	NDTKPLKEIIEDQMAVH[3]ERSNV	3857
Query_10010	3711	LRLDQRVQQLFSLMNGILKDKSFCQRN--	ITQRTYKVVPIITGHLGIIIEWI	NNTTPMRYFIEREPRTS[3]----	3779
Query_10011	3701	LRLDQRIEQLFEAMNVLNRDAACSQRK--	LAIRTYSVIPVSKKCGLLQFV	ENTCVLQEVIRDGLSSK[9]-KPLQ	3779
Query_10012	3631	LRQDQRIEQLFNLMMKILDQDPTCRHRN--	LSITTYQVVPVSSHLGIIQVW[47]	GEAHMKYS-----	3734
Query_10013	3826	LRTDARMEQLFAVMNQALAADPACGERG--	LSIRTYVVPMTTRIGMLEVW	PNSCVFKEMINLGMTDA	EREDA 3896
Query_10014	3696	QRQDQRVEILLDIINLLKSDSKYQRN--	LSIKTYQVVPMTSKLALIEWL	PETKPLKVIENDLVNE	AQNNY 3766
Query_10015	3713	LRTDQRMEQMFLLMNSIYARSPLCTKASV	PSLDTYQVIPLSLRVGLIQVW	DSTQPLKDFINDSYREN	ETKCY 3785
Query_10016	3744	LRQDSRIEQLFVDMNDIYQCNALCFQNK--	LHLKTYQVIPLTTIKIGLIEWV	DNTNVLGSFLKDAMTDK	EKESF 3814
Query_10017	3797	LRLDQRIEQLFETMNVKLSDDPICSQR--	LNLRTYSVIPTPRVGLIEWV	RNTMPLKDFIQLSDPEK	KSLPA 3867
Query_10018	3763	LRLDQRIEQLFCVMNEILAQDSICSQD--	LRLRTYQVIPITQRLGLIEWM	KNTITLKEFVTDAMEDH	EKQHY 3833
Query_10019	3802	LRLDQRIEQLFGLMNDIMANDSACSQRG--	LRLRTYQVIPMTPRVGLIEWM	NNTKPLKGVVNSAMTKA	ENDHY 3872
Query_10020	3598	LRLDQRIEQLFVDMNQIFKADPRCRDQK--	LSITTYVVPMTQEIGLLEWI	DGTSTLKGVLVEAQLQID[1]RCVGL	3669
Query_10021	4044	LRQDQRIQALGLFNLSLQRDIQCLRLR--	LSLVTYQVVPVPLSDPLGLEWL	ERTQPYKDFLVHQMSAA	EQRSH 4114
Query_10022	3856	LRLDQRVQQLFQIINSIFKDDAQQRND--	LYIKTFKVVPTINRLGTFLEWV	DDTEPMKSLINREHMRN	EAGKD 3926
Query_10001	3806	QNK-NLPNNPANLCCQWIG-dsS-NTRQMI	ENYRKYQKYSRNETLNAFKKREAMVPWDM	LKRSYQLASSPESYFVLRG	3882
Query_10002	3894	NSR-HGPRYQHSEWVDKYAKgnQ-SNH--	GVRYGAVYRHANYTDTVKCFRSEALVP	WDLMRRAFIQLSASPEAFITLRS	3969
Query_10003	3904	LHQ-RLPLIDFDAYVHRLG-gtV-KNDP	WGKVVYGLYQKASRTDVTLAFRQVQSK	VTDIMRRGYLQMSASPEAFITLRS	3980
Query_10004	3829	LSdpRAPPCYKDWLTKM----S-GKH	DVGA-YMLMYKANRTEVTVSFRKRESK	VADLLKRAFVRMSASPEAFITLRS	3902
Query_10005	3844	QQEvLP---KYREWLOQM---A-GRK	DMFLRYPTMYLKSARTEVTAFRTHEQA	IPDGLLRRAFIKMSTPEAFITLRS	3915
Query_10006	3784	LGD-QGPGKHENWYKYF--da-KKRC	IKSMYQVYQNSRAETIKDFRMKENK	VPWDLFRRAFHQMSASPEAFVLRG	3859
Query_10007	3817	-AQ-QKAAILQRNFLSRYG--np-SD	PNICCKYWGMYERCSTRTEAVKSCQDK	ENSLPWDLRRAYLQMSASPEAFITLRS	3891
Query_10008	3680	AME--MFTEGPEKILLKAA-geG-SS	PLNHLTYWRSFQNI TPEAARCFNEV	LKIPDHLLRDLQLLNCVSPDVFCFLRR	3755
Query_10009	3858	SLsMEAYKIHNENWISNFGkyiKm	ANPPAGPLYQQMFMSDRADAIKKQEK	QHDKVPANLLQNGIWAISSPESYIFIRN	3937
Query_10010	3780	----DDAQSMHSRWVQYK-----	SYHNMFMENVNRDAVVQHFKDLQRE	ISNDMLKALYKLSASPEAHLIRS	3843
Query_10011	3780	ATSeTLKKEIQHEIYEWVrk-G-GGKS	QTEPCYNLYRKVGFEEVSLKLSALQ	QGLPWDTLRMGFSRLVTSAESYALRS	3857
Query_10012	3735	-----PEETS	AEPFRKKNMIPWDIMRRALQRLS	VSPESYIFILRQ	3773
Query_10013	3897	KKR--KATQLPQNHFQFAnrfs-KD	PSPLMKSSVMTKAKRSDISRSLERQ	ALVRPTLLRDALLHLSATPEALYTRRQ	3973
Query_10014	3767	AE-----FIYGNKDG---S-K-KT	QSEVGMVYAKYSRELICREFDKIQNS	IPWDLFRQFLRTLSDSEAYFVLRN	3833
Query_10015	3786	GEA-----RALYKNE-----	NYTVAKKAKRVDVIKAYEEVNVN	IPWDLVLRGLVLRMSSTEGFSLRS	3844
Query_10016	3815	RES-DPKMTFYKFLLSKTN---S-R	DKSIPLLYENIY-KIADTEVNKKFTE	CINIPWDLFRRAFLKFSASLEAFVILRY	3888
Query_10017	3868	ASK-----MNQKWLEQKF-----	GGSNYVNYEMYKKLKADBEVKKEY	RRIIVSAVPWDLRRSLISLASTSESFLLRG	3936
Query_10018	3834	LSSgNSPHQLYSKWLGKF---gs-AG	GPVYM-YQSMYRKAKRSEVENHFK	RENSEVPWDLRRGFVALSASSEAYMLRS	3908
Query_10019	3873	IGP-NGAGACYRNWDKPKF--gd-PR	NSPRT-YGEMFKKANRTEETERSF	TQKQALVPWDLRRRAFLQLAVSPEAFILRS	3947
Query_10020	3670	KSNnSQKLDLNFNTAAKAYes-F-L	LEQRGATYSAKVVASRSEEDVVDH	FAKVHAMIPDILLRQLLKLGS	TSVDFVLRD 3747
Query_10021	4115	SKQ---SLDYCKFVSNDHck-----	LYTSLDQQVVQRNFPQLLDCVPD	CLLARGLRQLSSWEAYAMRD	4176
Query_10022	3927	LNQ-SRALNERRRKWLKSSVpg-N- AKKGETDQHCILLCDEKIVIEGF	NKHKEAIPWYLLRNLGLENLCLT	SSAFLAIKN	4003

Supplementary Figure 5: COBALT analysis DNA-PKcs from representative examples from all major phyla (short list)

Query_10001	3883	NCATSHALNCISAYIMGIGDRHLSNHMVDLNGKMAIDFGHAFGSATQFLDIPPELIPFRMTRQLRNLNLLPL-REKGVIE	3961
Query_10002	3970	HFARSHAVICICQYILGIGDRHLSNFIISLTKGMMIGIDFGHAFGSATQFLVPELMPFRTRQIVNLMPLP-KKEGLIQ	4048
Query_10003	3981	HFTTTHATLCAIQYVLGIGDRHLSNFMIDLESGVMGIDFGHAFGSATQHLVPELMPFRTRQMLNLLPL-KESGLLY	4059
Query_10004	3903	HFASSHALICISHWILGIGDRHLSNFMVAMETGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPM-KETGLMY	3981
Query_10005	3916	HFASSHALMCISHWILGIGDRHLSNFMVNLTEGVMGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPM-KESGLLR	3994
Query_10006	3860	ELVVTHATISICQYFLGIGDRHLSNFMINLTKGMMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-GIKGLE	3938
Query_10007	3892	HFATSHALLCICHYLLGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-REKGLIE	3970
Query_10008	3756	KFTVSLAMCICIANYLEIGDRHLSNFMIDLESGVMGIDFGHAFGSATQHLVPELMPFRTRQMLNLLPL-KESGLLY	3834
Query_10009	3938	SFARSLATFSCISYIIGIGDRHLSNFMIDLESGVMGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPM-DSVGLLS	4016
Query_10010	3844	EFIKCHAVINICGYLIGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPM-GSKGLE	3922
Query_10011	3858	QFARSLSVTSCIGYVAGVDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-TAVGLLR	3936
Query_10012	3774	SFACSYAVLCIAHWLLGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-EETGLFR	3852
Query_10013	3974	LFGRTLATMNICQYILGIGDRHLSNFMIDLESGVMGIDFGHAFGSATQELPIPEMIPFRTRQLEAVFLPH-KKEGLVQ	4052
Query_10014	3834	EFIKSYAVASTCHYLLGIGDRHLSNFMIDLESGVMGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-EORGLFE	3912
Query_10015	3845	AFVSYAALCISQWLLGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-GQVGLMK	3923
Query_10016	3889	NFITSHAVLCISQWLLGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-KEKGVIE	3967
Query_10017	3937	TFIQSLSVNICQYILGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-KVSGQLR	4015
Query_10018	3909	HFARTHATVSIICHYLLGIGDRHLSNFMIDLESGVMGIDFGHAFGSATQFLVPELMPFRTRQMLNLLPL-RENGQLK	3987
Query_10019	3948	HFASSHAELCICQYILGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-KKSGTLQ	4026
Query_10020	3748	NFLKSLAVFSACSYVLGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-DGSPLLA	3826
Query_10021	4177	TLVRSHAAMCGAHWLLSVGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-PLSGQAFV	4256
Query_10022	4004	QFTKSLATFSIASYLLGIGDRHLSNFMVLDTEGVMIGIDFGHAFGSATQFLVPELMPFRTRQFINLMLPL-PLSGVYK	4081
Query_10001	3962	GTMVHSLRALRQHEQLISTMDVFIKEPSLDWQNA---KK EIQ A- --MSNRKedtSKLTLDSQ WFP 4022	
Query_10002	4049	SSMVHTLRALRQNHLELLNTMDVFIKEPSLDWKTNA---SK QAK AQ[3]GrVLEQGsdpDGNLSDES WYP 4115	
Query_10003	4060	STMVHLRALRSPNVLLNTMDVFIKEPSLDWKNFA---AK QM- --[2]GmVGE----DDDLDLIS WYP 4117	
Query_10004	3982	SIMVHALRAFSPDGLLNTMDVFKEPSFDWKNFE---QK MLK KG G--SWIQ---EINVAEKN WYP 4040	
Query_10005	3995	SIMVHGLRAFRLDPLLSTMDVFKEPSLDWKNLE---LK QMK KG G--TWNK---DVNTREIH WYP 4053	
Query_10006	3939	NTMHTTRALRQDYLLVNTMDVFKEPSLDWLNFA---EK QRN E- --MLMTA---DD-EEST WYP 3994	
Query_10007	3971	STMVHALRALRANYDLLNTMDVFKEPSLDWLAFA---EK QLQ T- --TMSQ---EDKEDAK WYP 4027	
Query_10008	3835	STMKYCLKAFRDSYILLNTMDVFKEPSLDWIVEA---RR QIG[1]-- GrTLWNDe--ANLRDSFK WYP 3895	
Query_10009	4017	HNMHTLTAIQNNKDFLLSTMDVFKEPSLDWIKLAsrlGK[1]QKG[2]KS[3]----- WFP 4071	
Query_10010	3923	YPMKIKLQIMQANKDILLNSMDFVKEPSLDWLNRAAndqAK QQR[2]-- --NSNeigEGDSQTV E WYP 3985	
Query_10011	3937	SDMVRIIMTALRSGRLGISAVMEVFKEPSLDWVQRQAvklQR[1]RAR[4]-- StALEADaganNHMAAFS[3] ]-HV 4008	
Query_10012	3853	ETMHCRLRTRLNNDMLLATMEVFKEPSLDWLEFA---SR HES[1]-- ----NK---KSHEESQT WYP 3907	
Query_10013	4053	EVMVHLRALRANQRLLSTMDVFIKEPSLDWQQA---RG DVA -- ---- WYP 4096	
Query_10014	3913	SSMVHCLKALRENNDLLVCILDVFIKEPSLDWIGSA---VK IAK[4]ED[3]TlGLQ----- -YA 3969	
Query_10015	3924	EIMVSALGALQESRHVLTAVLEAFVKEPTKDWLDFV---QK QEG[1]-- ----TQ---DSKVEL-- -FS 3975	
Query_10016	3968	ATMISALRALRDNADILLNTMDVFIKEPSLDWQENA---LK QKR[10]SD[3]YpGNTESddpNHSECFN WYP 4044	
Query_10017	4016	SCMHTLRALRHSPTLLNTMDVFIKEPSLDWESNA---KK QAL AQ -----KINTDNT WYP 4069	
Query_10018	3988	CSMVHTLRALRKNPDLVNTMDVFIKEPSLDWQNAFA---NR QNA KQ G--LKTE---S---DDFS WYP 4043	
Query_10019	4027	SVMHTLRALRSHDILLNTMDVFIQEPSLDWQVYA---RK QAK TQ G--INEE---E---QKAD WYP 4082	
Query_10020	3827	REMGTVDALRSKRVVSEVMNVFLHDPDLDWQSSkikqK[9]DLH[9]KD[6]TrESQRkkaStLPATGN[4]WLP 3921	
Query_10021	4257	RCLGHCLAAALRRERELVLSVLQVFINDVADWQQLAqmlTA[3]ESS[9]AG[3]AsASSAaaggSGSIASFS[4]WFS 4342	
Query_10022	4082	QTMVHALNALKHKRKSIMDTCEIFVKEPSLDWVKDAkqrQK[2]ENN[1]--[1]GmVDSLIsEeyEELQELA WYP 4150	
Query_10001	4023	KEKVNIVRQKLEGNPSHVMRAELALG --QLGSYLKYfviVLMGDK QADIRARLPADG---LSSEQVAALLN 4090	
Query_10002	4116	REKVNFAQKKLEGANPSDITQRELALG HSRNPAFSAI-KEVVGDR SQNIRASSPKSG---LSVETQVECLID 4184	
Query_10003	4118	REKIKFAAKKLGQVNPHTITKAELQLG HKNLPWFKSF-CLVAMGDG GKDVRAQKPAEG---LSVEDQVACLID 4186	
Query_10004	4041	RQKICYAKRKLKAGANPAVITCDELLG HEKAPAFRDY-VAVARGSK DHNIRAQEPESG---LSEETQVKCLMD 4109	
Query_10005	4054	MQKVRYARRKLEGANFAAITRDELQLG HETEAYSY-VAVAMGDK EHNIRARVPEDG---LSVEIQVDCLID 4122	
Query_10006	3995	KQKQYALRRFKGDNPFCHITKDELYLG HSRSPAIKHFekCV-LGDA KDNVRATLPSTG---LTVQVEALID 4063	
Query_10007	4028	KQKVAFARRKLGANPWCITRDELELG HRRG--KTYqAFIAITQG QDNTRTKFAEVD---LTPQVQVACLID 4095	
Query_10008	3896	EKISAAARKLRGDNPFVIMQDELMNS[3]LSFPLHQI--NCIYGKK[8]DQSPARGYPLQA---LTVBQVTSCLIE 3974	
Query_10009	4072	KEKVDIARRKLERWNPAYITSEELANS[1]HGQIYEKSI-QNVVKGDP[1]-HNIRAKACQIC---STAKEQVDCID 4141	
Query_10010	3986	RQKLDIARRKLEENFPAYIVCSELEAG HSNKPWFNSI-IKIAKGD[2]-NIRARARQKC---FSVKEQVECLID 4054	
Query_10011	4009	QKLVETAQCKLDLWNPAYITISELGF[1]HANKPYTRAL-BEIVRGDP[1]-RNIRARVAGNV--cVSVQEQIDCLLD 4079	
Query_10012	3908	QKQLDLAKLLEGANPVTITTEBELKAG[6]LAK-----F-LECVKGHN --SVRSQMPYSQ---LPLPEQVDSLLD 3974	
Query_10013	4097	EQKTAFAAKRKLKAGNPAHLLISELESN[3]TSDRSWHRKA-KDIIAGDA QSNIRATMPEHG---LSVEDQVRLLD 4168	
Query_10014	3970	KDRVNSVKQKLFGMHPSVIMRNDLCSG[1]HKDSKYLYQF-VDVLMGNK[12]SKILAEENGEKRY---LTVIEQVKCID 4051	
Query_10015	3976	SKRIELNDKLSINFPAYITLWAVS--[4]VAKQAYENL-KQAVLGDG GESVRSVARGYR---LSTHQVVDILLE 4046	
Query_10016	4045	KQRILFAKRLKLEGNPCYITKEELRLG HSNKPIFQRL-TDVCLGDR -KSLRQNLPKNG---LSVKQVKSLLID 4112	
Query_10017	4070	KEKVNICRKKLGFNSHMLLDLEHG FKNKPEFKYL-KGILLGDS STNLRARVGRKC---DSVEDQVDCID 4138	
Query_10018	4044	REKVTFAKRLKLGANPAFVTRDELRLC HHLST----L-ESFALGDK QYNERARMPADN---LSVEDQVTCID 4108	
Query_10019	4083	KQKQVSAWRKLEGNFPAYVMRDDLQLG HSKQSWYQOM-ESVCLGDR NANVRAREPESG---LSVESQVSCID 4151	
Query_10020	3922	NVKTIAARRKLEGVSPALLLLQELSQN[3]HKHLNKYSLvNAASTGED[1]-----ADEiaaLSSLERQAQLLA 3989	
Query_10021	4343	TEKLVVVRQKLSGAHPQCTLSELLLD[1]HRGKPYTSGM-LAVIDRRR[1]--SVGGDAER---LVPMQAELLVD 4411	
Query_10022	4151	RKKIEVMKDKLNGLNPVRIKAKELQDS[1]HKKPYFKVI-LNAIKQE[1]--SIRYKLLAQNrryLEVDFIDCLIE 4222	

*Supplementary Figure 5: COBAL analysis DNA-PKcs from representative examples from all major phyla (short list)*

Query_10001	4091	LATDPNILGRTFAGWEPWV	4109
Query_10002	4185	QATDPNILGRVYAGWEAWM	4203
Query_10003	4187	QATDPNILGRTWQGWQSPM	4205
Query_10004	4110	QATDPNILGRTWEGWEPWM	4128
Query_10005	4123	QATDPNILGRVWEGWEPWM	4141
Query_10006	4064	QATDPNILGRVWQGWAEAV	4082
Query_10007	4096	QATDPNILGRAYEGWEPWV	4114
Query_10008	3975	MATDSVILSRWFEGEAWI	3993
Query_10009	4142	QSTDPNILARAWAGWSSWI	4160
Query_10010	4055	LSTDPYVLGSMWVGWQSWI	4073
Query_10011	4080	QATDPNLLGRIFAGWQPWV	4098
Query_10012	3975	QATDYNILGKTWQGWEPWM	3993
Query_10013	4169	QATDPCILGVTWAGWEPWK	4187
Query_10014	4052	QSTDTNILGRTWIGWQPPC	4070
Query_10015	4047	QATDPNILGRSWIGWDPFL	4065
Query_10016	4113	QATDSGILSRIYFGWDPWF	4131
Query_10017	4139	HATDSNVLGRTYGGWEAVV	4157
Query_10018	4109	QATDPNILGRTWQGWEPWM	4127
Query_10019	4152	QATDANILGRTWQGWEPWV	4170
Query_10020	3990	IATAPDLLGRTFQGWMPWL	4008
Query_10021	4412	MATDPALLARTYGGWQPDL	4430
Query_10022	4223	QSRDPNILGRTWIGWSPYV	4241