

Additional File 13. Alignment of Dact protein sequences.

Alignment of protein sequences for all Dact genes used in the phylogenetic analysis. Gaps are indicated by '-', missing sequences by 'X'. Also shown is the position of the conserved protein motifs.

Genomic coordinates and sequence alignments for various species. Includes labels like 'Cm_3', 'M1_2', 'M2_2', etc., and corresponding nucleotide sequences.

Genomic coordinates and sequence alignments for various species. Includes labels like 'M1_1', 'M2_1', 'M3_1', etc., and corresponding nucleotide sequences.

Genomic coordinates and sequence alignments for various species. Includes labels like 'M1_1', 'M2_1', 'M3_1', etc., and corresponding nucleotide sequences.

Genomic coordinates and sequence alignments for various species. Includes labels like 'M1_1', 'M2_1', 'M3_1', etc., and corresponding nucleotide sequences.

```
Dr_1 ---YTSPTA YV-----GSD-SEY--SABCA SLPHSTILT S---EDE--RSNYTNC-----FDGSESAS EADYV--AE SS--TSSSD EGSAG---VNRQI SQAGSGSH-----GMSA QAKAVFKIK---ASBHLK
On_1 ---YTSPTA YTA-----GSD-SEY--SABCA SLPHSTIVDT S---EDE--RSNYTNC-----FDGSESAS EADYV--AE ST--WTSDD EESGGG---MRRGR QLRGAVRAT GQ-----EMTSA QAKAVFKIK---ASBHLK
On_1 ---QASQVC GIR-----GSD-SEY--SABCA SLPHSTIVDT S---EDE--RSNYTNC-----FDGSESAS EADYV--AE ST--WTSDD EESGGG---MRRGR QLRGAVRAT GQ-----EMTSA QAKAVFKIK---ASBHLK
Hs_3 ---GFSPPA P---QRRLL YGCA--GSD-----S-ECS AG---RLL GPLGRG--P AGGYG--GG--YGESSESAS EGESP--AF SS--ASSDS EGSOGG---LWFPQ L-----VA ASAAGG--AG GAG--APAG PAKVFKIK---ASBHALK
Md_3 ---GFSPPA P---PRLLL YGCA--GSD-----S-ECS AGGGGAGRR APPLGVSGG YGGYG--PA--LWFPQ L-----VA ASAAGG--AG GAGG--Q G-KVFKIK---ASBHALK
Ac_1 ---TAI-----TG YTTNT--GSD-----S-ECP AE--PPRQ APFQMEG---IW---YGESSESAS DGESE--AVY SS--GASDD DSGG---LWFPQ LFPOL--VS AS--GP-QK TSGG---Q P-KVFKIK---ASBHALK
Cpb_3 ---SP--PCHVL YGTA--GSD-----S-ECS GGGG--GRR APVCELEG---LA---YGESSESAS EGESP--TF ST--G-SDD DSGG---LWFPQ LFPOL--VA SA--GP-QG AGG--R P-KVFKIK---ASBHALK
Xt_3 -----PRSSM YGTA--ESD-SEY--SADC-----S-VRC APVCEPDD---LSSYTMNC-----FDGSESAS EVESP--GF ST--CSDT DSGG---LWFPQ VEPOL--VA SS-----GSL--QSSM P-KVFKIK---ASBHALK
Lc_3 -----ETDF QER--ALFCR P-----EGB---YTPFC-----FDGSESAS EADSP--GS SS--LSSDS DSGG---LWFPQ LFPOL--A A-----RP-PS FSS-PPCPQ P-KAFVKIK---ASBHALK
Dr_3a ---ORL-----QY-----R-----ME---ASFCP-----TWGDESLA EFPSP--GT SS--VSDS DSGG---LWFPQ LFPRL--FP SSSS-----GRR FPAVFKIK---ASBHALK
Tr_3a ---ORR-----LDV QER--ALCQ P-----EED---SYATI-----PASESMS EVYSP--AS SS--LSSDS DSGG---LWFPQ LFPRL--AS TSSSSP-SP QTA-NATQ P-KAFVKIK---ASBHALK
On_3a ---ORR-----QDF QES--ALCQ P-----GEE---SYGAA-----LWFPQ LFPRL--AS TSSSSP-SP QATA-NTFS P-KAFVKIK---ASBHALK
Dr_3b ---ORR-----QDF QES--ALCQ P-----LRO---YT-AY-----PGESESLA EADSP--GS SS--LSSDS DSGG---LWFPQ LFPQL--AP T-----SP-SH FPGH--GDLQ P-KAFVKIK---ASBHALK
Tr_3b ---ORR-----EQY PRRYFMQ P-----SED---YMRP-----QGESSETS EADSP--DS SS--LSSDS DSGG---LWFPQ LFPQL--SL A-----SP-B FFF--GAQQ P-KAFVKIK---ASBHALK
On_3b ---ORR-----QIM-----EPQ LERDAQLC A-----TDD---YSHAG-----QVESSEMS EADSP--DS SS--LSSDS DSGG---LWFPQ LFPRL--S-----LP-ST LAFS-GDLQ P-KAFVKIK---ASBHALK
Cm_3 XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
Pm_A FVMVVPAA VYGGGALBS QRYV--ESD-SEY--SABCE SLFRP-SFGA T--GSDDD-----DDDDR DEEDDDB DEDA-RGG- GDVTVPAV ABGGRAEPA-----LASSLA SSSSSSSP FXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
Lj_A FVMVVPAA VYGGGALBS QRYV--ESD-SEY--SABCE SLFRP-SFGA T--GSDDD-----DDDDR DEEDDDB DEDA-RGG- GDVTVPAV ABGGRAEPA-----LASSLA SSSSSSSP FXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
Pm_B XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
Lj_B ---ARAY VVAVARADS-----DSER SEY--SABCE SLFRGVTMT S---EDE--MSDTNRF EKERRAPAS FRQKVDVVD RD--DDEEE VADTDKRRR RRRLRREGF SAEDEEEDD VDTGGRQSK GLP-LARRG GQKAPVVA P-VAARGEARG G--GGARG QQAACRIT---ASRSLK
Lj_C PGPVIVPAA GCPCELQLR-----SAIAR M-----SEB-----MGLNRMFL ASGAGLAD F----LDBY R-----DAD DADA-----DAGRLP ADG-SAF-----RSS--GSE GLRG-----CWSQ QKAVFPAVK
Lj_D GGGVTRGMS FQKRTKQD PFP-----DGR AQC-FRDFP PRSFGYSES KQVAKRCS WELNRMFL ASGAGLAD F----LDBY R-----DAD DADA-----DAGRLP ADG-SAF-----RSS--GSE GLRG-----CWSQ QKAVFPAVK
Pm_D XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
Hs_2 ---GR PSLV---RQ DAYTRDSEP SKU--SABCD PRFSPVIFET S---EGE--SSDTNRF-----FDGSESAS DEEGGASD CD---LALY VAA---GH AE--LAW---TOE-----AP-V S---SPLLS FVKLRIK---ASKALK
Md_2 ---SR SKIP---SH HYTKASESDH SEY--SABCA SLPHSTIVET S---EED--VSDYTNRC-----FDGSESAS DSDA-CRSE SS--LTLY EA-----VG SE--LWPKA IVQPL---VR--QQ-A A--SKPSBH FVKLRIK---ASKALK
Ac_2 ---AR SKSR--FPFH QYAKESDI SEY--SABCA SLPHSTIAET S---EGE--VSDYTNRC-----FDGSESAS DSDG---SSS SS--LALY DGG---DE SE--LWPKA SVRQ-----GTQA G--SK-LP FVKLRIK---ASKALK
Cpb_2 ---AR SKSR--H QYAKESDI SEY--SABCA SLPHSTIAET S---EGE--VSDYTNRC-----FDGSESAS DSDG---SSS SS--LTLY DDA---DE SE--LWPKA SVRQ-----AVQA G--SK-LP FVKLRIK---ASKALK
Lc_2 ---TH QARYRISAI RSRARSDSD SEY--SABCA SLPHSTIVET S---EED--RSDYTNRC-----FDGSESAS DSDAT--SD SS--LTLY DDT---DA SE--PWFEA SS-QP--AA SS--Q-PAVA SVARSK-RI TFKLRIK---ASKALK
Tr_2 ---RPSYQY VWRH--PSS YGCLLSQV SEY--SABCA SLPHSTIAE S---EGE--LSDYTNRC-----FDGSESAS DSDG--SDS LS--LWKE D---M--ED GE--LWPKA TV-----GTAA G--LPLSR FPKCRIK---ASKALK
Dr_2 SCFSYLSRY PAAP---VSS RHPRCESE SEY--SABCA SLPHSTIAS S---DCE--MSDTNRF-----FDGSESAS SQT-ADF SS--LSDI ED-LLEEE DEGOLWQA AM-----GTAA GFL-LOQLR SESACRIT---ASRSLK
Tr_2 ---FLMLNARY PPSPF-MR HYPRCE--SEY--SABCA SLPHSTIAS S---EGE--MSDTNRF-----FDGSESAS FLSP--SDS SS--LSDI DYVDNGEE RN--LWADD SL-----GTAA G--LQQLR FPKCRIK---ASKALK
On_2 ---FLMLNARY PPSPF-VTS HYPRCE--SEY--SABCA SLPHSTIAS S---EGE--MSDTNRF-----FDGSESAS FLSP--SDS SS--LSDI GEQADSEEE GS--WVAAA AL-----GTAA GRLKQLR FPKCRIK---ASKALK
Pm_A ---VSRF RYFEN--FSV HYSARSDQ SEY--SABCA SLPHSTIET S---EED--RSDYTNRC-----FDGSELGQ SVSD--TASN SS--LSDI DDA-----LWPKA LV-----AR N-----EPTSS S--G-F-Ri SEQCRIT---ASKALK
Lj_A -----H LAF--SDSP NSFCFL--EG F-----ER--SL-----ATTSSS--LAKVATA N--LND-----WLOQ LQDN-----PH- A--DSDVIV RYKLRSA---SKPEK
Ac_4 -----H LAF--SDS-----CFP--DG S---QLE---VLRQ LQPRAR-----LWAS ATLDSNLSAM ASFKLRTR---SKPEK
Cpb_4 -----H LAF--SDSP NSFCFL--EG F-----ER--SL-----ATTSSS--LAKVATA N--LND-----WLOQ LQDN-----PH- A--DSDVIV RYKLRSA---SKPEK
Lc_4 -----H RLF--SDCP SFPSPLEP S---QTD--LNTNRY-----LYGES--LSVDTATC TS--LWNE-----LIED RSIARLGE-----KFS NFDLNMMN GSILYKTC---SKPEK
Dr_4 XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
Tr_4 -----H WVS--SIDVA GGLVYSDE A-----VR--LSTK-----EKLSSTA SLFL-SRQ S---LEGC H-----SIS-----PL SPSPSPFP SMPLRSR---SLRDLG
On_4 -----H WVS--SIDVA GGLVYSDE A-----VR--LSTK-----EKLSSTA SLFL-SRQ S---LEGC H-----SIS-----PL SPSPSPFP SMPLRSR---SLRDLG
Md_4t XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
Lc_4t XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
Dr_4t XXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
Bfl F-----SR GVRLLK-----KAGQR STR-LVDY T--SVLSS S---EE--SSS-----DSETR DSES--SE S-----S DAGG-----LVLA---KAKWP AVVGTAT---R
Motifs *****motifid T(ex4/ex5)***** **x11f ***** **x11f ***** **x11f ***** **x11f ***** **x11f ***** **x11f ***** **x11f ***** **x11f ***** **x11f *****
```

1810 1820

```
Hs_1 KKILRFSSG LKIMTV---
Md_1 KKILRFSSG LKIMTV---
Gg_1 KKILRFSSG LKIMTV---
Ac_1 KKILRFSSG LKIMTV---
Cpb_1 KKILRFSSG LKIMTV---
Xt_1 KKILRFSSG LKIMTV---
Lc_1 KKILRFSSG LKIMTV---
Dr_1 KKILRFSSG LKIMTV---
On_1 KKILRFSSG LKIMTV---
Cm_1 KKSCAEXXX XXXXXXXXXXXX XXX
Hs_3 KKILRFSSG LKIMTV---
Md_3 KKILRFSSG LKIMTV---
Ac_3 KKIMRFSSG LKIMTV---
Cpb_3 KKIMRFSSG LKIMTV---
Xt_3 KKIMRFSSG LKIMTV---
Lc_3 KKIMRFSSG LKIMTV---
Dr_3a KKILRFSSG LKIMTV---
Tr_3a KKILRFSSG LKIMTV---
On_3a KKILRFSSG LKIMTV---
Dr_3b KKILRFSSG LKIMTV---
Tr_3b KKILRFSSG LKIMTV---
On_3b KKILRFSSG LKIMTV---
Cm_3 XXXXXXXXXXXX XXXXXXXXXXXX XXX
Pm_A XXXXXXXXXXXX XXXXXXXXXXXX XXX
Lj_A KKILRSSG LKIMTV---
Pm_B KKMMHFRSSG LKIVTTV---
Lj_B KKMMHFRSSG LKIVTTV---
Lj_C KKIRIRLDA ARTYDTV---
Lj_D RALLRCHVZ PALTYTV---
Pm_D XXXXXXXXXXXX XXXXXXXXXXXX XXX
Hs_2 KKIRRFQPA LKIMTMV---
Md_2 KKIRRFQPA LKIMTMV---
Gg_2 KKIRRFQPA LKIMTMV---
Ac_2 KKIRRFQPA LKIMTMV---
Cpb_2 KKIRRFQPA LKIMTMV---
Lc_2 KKIRRFQPA LKIMTMV---
Lc_2 KKIRRFQPA LKIMTMV---
Dr_2 KKIRRFQPA LKIMTMV---
Tr_2 KKIRRFQPA LKIMTMV---
On_2 KKIRRFQPA LKIMTMV---
Cm_2 KKIRRFQPA LKIMTMV---
Ac_4 KAMSRLR--SRATK---
Cpb_4 RYVSRSLRPL APKASASK---
Lc_4 RYVSRSMRF SRFSVSTR---
Lc_4 XXXXXXXXXXX XXXXXXXXXXX XXX
Dr_4 RYVSRMSRL SFRKHSKX KX-
Tr_4 RYVSRMSRL SLKRKSKK---
On_4 RYVSRMSRL SLKRKSKK---
Cm_4 XXXXXXXXXXX XXXXXXXXXXX XXX
Lc_4t XXXXXXXXXXX XXXXXXXXXXX XXX
Dr_4t RYL-SGDRL SRMFPCHLV TEL
Motifs *****pD6-binding*****
```