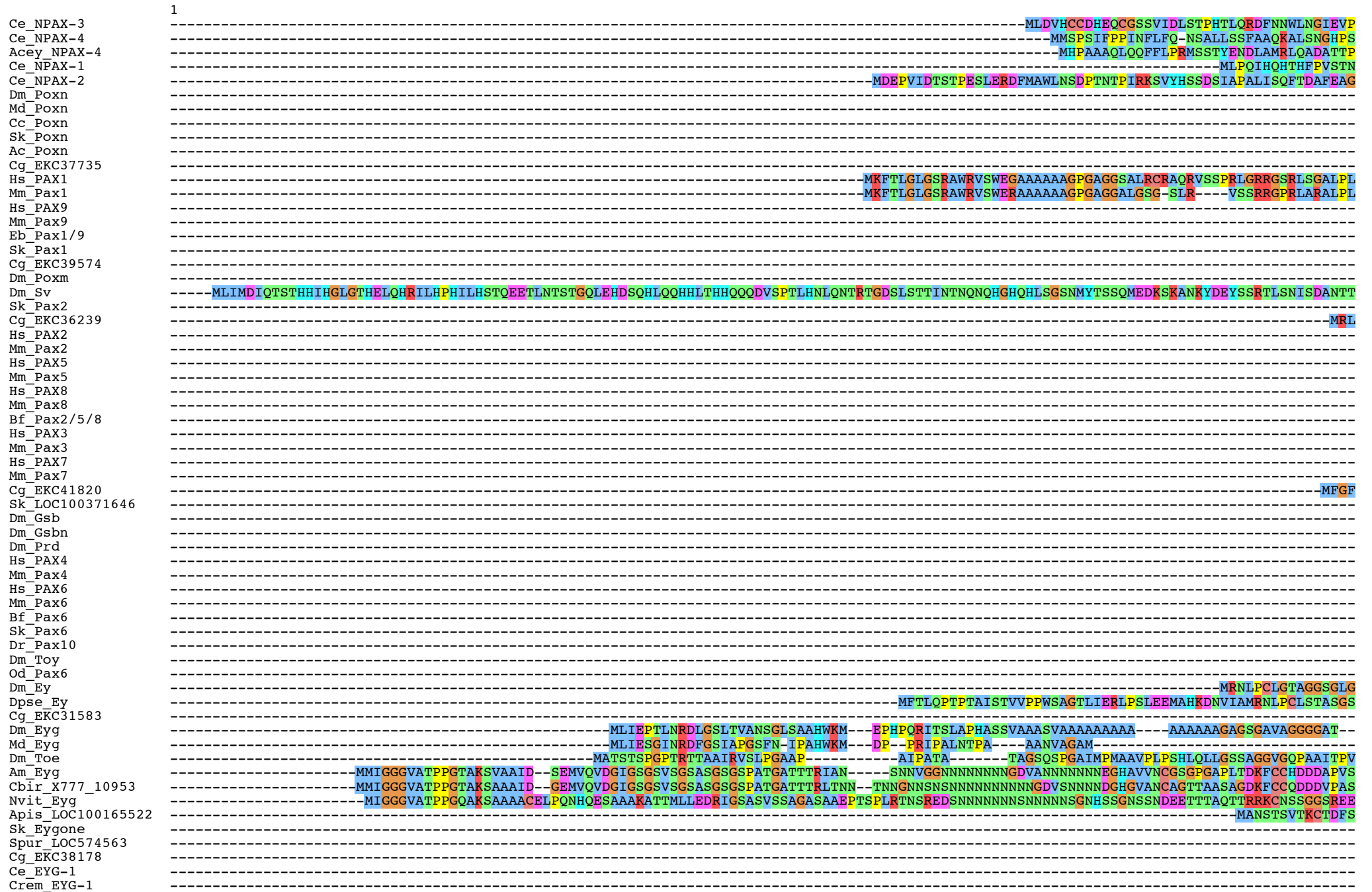


Sup. Fig. S5 Sequence alignment of selected PRD class proteins



	β 1	β 2	α 1	α 2	α 3
Ce_NPAX-3	NSYESQQSDSSSSVDESGSKTSSPTSMOQNPSODGRK-NRKKGTNLYGRPYCPGRPLS		MEERTRIIQIHLNN	GMKV-NAISKSLC	ISHGCVSKII-SRFRATGVLLP
Ce_NPAX-4	GSSLAALEKLSKAFRRPDLHIEKLASSCIDKQDLKMTNVVPSRNRYGRPYISGRPLL		TCDRQKIVECYKK	GMKK-IHIAKQLG	ITHSCVSKII-RRYAEATGEIVA
Acey_NPAX-4	HMENSSLNLLMLESFQAQSLSPRNPIMEAAVPSOPLLNVPYPSRNRYGRPYISGRPLL		TCDRQKIVQLFEN	GMKK-IHIAKQLG	ITHSCVSKII-RKYQDTGMVEA
Ce_NPAX-1	GSTGSPNNNTLPPPIIPYLLSFAFYSLSSANISTSPEINEGEKVKAVGRSYNPGRPLC		LEDKRIKIVRLYEE	GCRV-SHIARLIG	VTHSCVSKII-SRYRRTGSVQP
Ce_NPAX-2	MYNRYYPYPTSDSGSPSSLSGSSLDYDVPYPPYQYKEDVARSSGRNLGRYISGRLPLS		MCEREEIVKLFQE	GWKI-CDISKRLC	VTHSCVSKII-NRYRQTGSVKP
Dm_Poxn	-----MPHTGQAGVNQLGGVFNVRPLP		DCVRRRIVDLALC	GVRP-CDISRQLL	VSHGCVSKII-TRFYETGSIRP
Md_Poxn	-----MPHTGQAGVNQLGGVFNVRPLP		DCVRRRIVDELALC	GVRP-CDISRQLL	VSHGCVSKII-TRFYETGSIRP
Cc_Poxn	-----MPHTGQAGVNQLGGVFNVRPLP		DCVRRRIVQLALC	GVRP-CDISRQLL	VSHGCVSKII-TRFYETGSIRP
Sk_Poxn	-----MPHTGQAGVNQLGGVFNVRPLP		DFIRARIVELAHL	GVRP-CDISRQLL	VSHGCVSKII-TRFYETGSIRP
Ac_Poxn	-----MPHTGQAGVNQLGGVFNVRPLP		EHVRRRIVELAQM	GVRP-CDISRQLL	VSHGCVSKII-TRFYETGSIRP
Cg_EKC37735	-----MDPLOWGOTGVNQLGGVFNVRPLP		EHVRRRIVELAHM	GVRP-CDISRQLL	VSHGCVSKII-TRYETGSIKP
Hs_PAX1	CLSRGGGGAQALPDCAGPSPGHPGHGPGARQLAGPLAMEQTYGEVNQLGGVFNVRPLP		NAIRLRIVELAQL	GIRP-CDISRQLR	VSHGCVSKII-ARYNETGSILP
Mm_Pax1	CLSG-GGGARALPDCAGPSPRR--SGARQLAGPRAMEQTYGEVNQLGGVFNVRPLP		NAIRLRIVELAQL	GIRP-CDISRQLR	VSHGCVSKII-ARYNETGSILP
Hs_PAX9	-----MEPAFGEVNQLGGVFNVRPLP		NAIRLRIVELAQL	GIRP-CDISRQLR	VSHGCVSKII-ARYNETGSILP
Mm_Pax9	-----MEPAFGEVNQLGGVFNVRPLP		NAIRLRIVELAQL	GIRP-CDISRQLR	VSHGCVSKII-ARYNETGSILP
Eb_Pax1/9	-----MDQTFGEVNQLGGVFNVRPLP		NAIRLRIVELAQL	GIRP-CDISRQLR	VSHGCVSKII-ARYNETGSILP
Sk_Pax1	-----MHYGMEOTFGEVNQLGGVFNVRPLP		NGIRLRIVELAQL	GIRP-CDISRQLR	VSHGCVSKII-ARYNETGSILP
Cg_EKC39574	-----MRKLLDTEQSTFGEVNQLGGVFNVRPLP		NAIRLRIVELAQL	GVRP-CDISRQLR	VSHGCVSKII-ARYNETGSILP
Dm_Poxm	-----MDPESQCPQYGEVNQLGGVFNVRPLP		NATRRRIVELARL	GIRP-CDISRQLR	VSHGCVSKII-ARYHETGSILP
Dm_Sv	PSANNFITOSQGIIEWITAMNDIONGAEDSHSSQGISIGDGHGGVNQLGGVFNVRPLP		DVVRQRIVELAHN	GVRP-CDISRQLR	VSHGCVSKII-SRYETGSFKA
Sk_Pax2	-----MPQSGRSGHGGVNQLGGVFNVRPLP		DVVRQRIVDLAHS	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIKP
Cg_EKC36239	DFFSKAGIGLETMONTWAGFYGPPTFFPYQLQPRAYDVFAGHGGVNQLGGVFNVRPLP		DVVRQRIVELAHQ	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIKP
Hs_PAX2	-----MDMHCKADPFSAMHPGHGGVNQLGGVFNVRPLP		DVVRQRIVELAHQ	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIKP
Mm_Pax2	-----MDMHCKADPFSAMH-RHGGVNQLGGVFNVRPLP		DVVRQRIVELAHQ	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIKP
Hs_PAX5	-----MDLEKNYPTPRTSRTGHGGVNQLGGVFNVRPLP		DVVRQRIVELAHQ	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIKP
Mm_Pax5	-----MDLEKNYPTPRTIRTGHGGVNQLGGVFNVRPLP		DVVRQRIVELAHQ	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIKP
Hs_PAX8	-----MPHNSIRSGHGLNQLGGAFVNGRPLP		EVVRQRIVDLAHQ	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIRP
Mm_Pax8	-----MPHNSIRSGHGLNQLGGAFVNGRPLP		EVVRQRIVDLAHQ	GVRP-CDISRQLR	VSHGCVSKII-GRYYETGSIRP
Bf_Pax2/5/8	-----MDRMTTMAASGSMQOHHGDSGGHGGVNQLGGVFNVRPLP		DVVRQRIVELAHQ	GVRP-CDISRQLR	VSHGCVSKII-RRYYETGSIKP
Hs_PAX3	MTTLGAVPRMMRPGPGQNYPRSGFPLEVSTPLGQGRVNQLGGVFNVRPLP		NHIRHKIVEMAAH	GIRP-CVISRQLR	VSHGCVSKII-GRYYETGSIRP
Mm_Pax3	MTTLGAVPRMMRPGPGQNYPRSGFPLEVSTPLGQGRVNQLGGVFNVRPLP		NHIRHKIVEMAAH	GIRP-CVISRQLR	VSHGCVSKII-CRYQETGSIRP
Hs_PAX7	MAALPGTVPRMMRPGPGQNYPRSGFPLEVSTPLGQGRVNQLGGVFNVRPLP		NHIRHKIVEMAAH	GIRP-CVISRQLR	VSHGCVSKII-CRYQETGSIRP
Mm_Pax7	MAALPGTVPRMMRPGPGQNYPRSGFPLEVSTPLGQGRVNQLGGVFNVRPLP		NHIRHKIVEMAAH	GIRP-CVISRQLR	VSHGCVSKII-CRYQETGSIRP
Cg_EKC41820	VNGFGVYLINERLIVRYPTAMFSYHLAALGLMPSFOMEGQGRVNQLGGVFNVRPLP		NHIRLKIVELAAQ	GVRP-CVISRQLR	VSHGCVSKII-CRYQETGSIRP
Sk_LOC100371646	-----MTIQGRSFAAPGFPLEGQGRVNQLGGVFNVRPLP		NHIRYKIVEMAAH	GIRP-CQISRQLR	VSHGCVSKII-CRYQETGSIRP
Dm_Gsb	MAVSALNMTPYFGGYPFQOGRVNQLGGVFNVRPLP		NHIRRQIVEMAAA	GVRP-CVISRQLR	VSHGCVSKII-NRYQETGSIRP
Dm_Gsbn	-----MDMSSANSRLPLFAGYPFQOGRVNQLGGVFNVRPLP		NHIRLKIVEMAAAS	GVRP-CVISRQLR	VSHGCVSKII-NRYQETGSIRP
Dm_Prd	MTVTAFAAAMHRPFFNGYSTMDMNSGQGRVNQLGGVFNVRPLP		NNIRLKIVEMAAD	GIRP-CVISRQLR	VSHGCVSKII-NRYQETGSIRP
Hs_PAX4	-----MHODGISSMNLGGLFVNGRPLP		LDTROQIVRLAVS	GMRP-CDISRILK	VSNCGVSKII-GRYYRTGVLEP
Mm_Pax4	-----MOQDGLSSVNQLGGLFVNGRPLP		LDTROQIVQLAIR	GMRP-CDISRILK	VSNCGVSKII-GRYYRTGVLEP
Hs_PAX6	-----MQNSHSGVNQLGGVFNVRPLP		DSTRQKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIRP
Mm_Pax6	-----MQNSHSGVNQLGGVFNVRPLP		DSTRQKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIRP
Bf_Pax6	-----MPHKAWTLORPADEHAQYSPVOADPGHSGVNQLGGVFNVRPLP		DSTRRKIVELAHQ	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIRP
Sk_Pax6	-----MMEFSTDSANGHSGVNQLGGVFNVRPLP		DSTRQKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIRP
Dr_Pax10	-----		-----	-----	-----
Dm_Toy	-----MMLTTEHIMHGHPS--SVGQSTLFGCSTAGHSGINQLGGVFNVRPLP		DSTRQKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIKP
Od_Pax6	-----MEGLAAEDPMNRALTAQKNFYALDFSGSGLGHSVNQLGGVFNVRPLP		DSLROKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-ARYYETGSIKP
Dm_Ey	GIAGKPSPTMEAVEASTASHRHSTSSYFATYYHLTDDECHSGVNQLGGVFNVRPLP		DSTRQKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIRP
Dpse_Ey	GLGGKSATAMEAVDATTGAQPHSTSSYFTTYHYHLTDDECHSGVNQLGGVFNVRPLP		DSTRQKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIRP
Cg_EKC31583	-----MFVSLTDDTASNSVSGPDDVDSFNKRSRQRKHSVNQLGGVFNVRPLP		DSTRQKIVELAHS	GARP-CDISRILQ	VSNCGVSKII-GRYYETGSIRP
Dm_Eyg	-----STSPTSQIIPP-----FGGGGTTGGIIPHAGGPTGSLTNT		LMSQHRLLFEFSRFG	GLRGY-DIAQHML	TOQGAVSKLL-GSLRP
Md_Eyg	-----PPLHPNVGLVPS-----GQIP-GSAIFTGGANGQSPSTLSA		LVSQORLLELSRFS	GLRGY-DIAQHML	TOQGAVSKLL-GSLRP
Dm_Toe	SPTAATTVLAQPHSLSPSTPILTOGAPPAAALGGIFCGGSA-AAGPGAASAAAATNLNA		LASQHRLLLELSRF	GLRGY-DIAQHML	TOQGAVSKLL-GTLRP
Am_Eyg	TGVARPEPPSPPTFSQTRSHLQVHPHHHQH--PAHHHQMTVPPVSLIDGVSLONCTAG		PFASGNNGGSSRRLELSHGLGALRHYNDLANHVLNLNQQGAVVTKLL	-----	GTLRP
Cbir_X777_10953	TGVTRPWEPPSPPTFSQTRSHLQVHPHHHQH--PAHHHQMTVPPVSLIDGVSLONCTAG		PFASGNNGGSSRRLELSHGLGALRHYNDLANHVLNLNQQGAVVTKLL	-----	GTLRP
Nvit_Eyg	RTTESSSAGQOQPSVARPWEHRHTAQPTAAQIAAVAAAHQHQQHPLPALIEGISLPCNGAGLFTSQTSGAGGNSNNSSSGGTGGSRRLELSHGLGALRHYNDLANHVLNLNQQGAVVTKLL		-----	-----	GTLRP
Apis_LOC100165522	VSSLCRDMKTSPPDPAIVVGGTGTNMKLTPLAACARFSSNINGPTAITAVNGRPPPSSTSPSPSPSP		-----ATLRELYAMAASQHQHLSLLELSNR	-----YRGPPEIAHVF	-----NQOAVTKLL
Sk_Eygone	-----MDTTTVHPGLNGLPOLGGMFLNGRPLP		DPLRHRIVELAHC	GTRP-CDISRQLL	ISHGCVSKII-GRYYEMGAIRP
Spur_LOC574563	-----MAMQGTLPYGTFLPAEPGRLLA		EDLRGMVSMILPC	NSRA-CDLTHPLLS	IPOGYSKLF-GRLYEPTPTRS
Cg_EKC38178	-----MKLHENAPNLMSGFSNSASPSGVG		NDVAGKAALFRPY	ALPP-SPSVRYLDRTS	PLPPVAHRVMWAPYIDSVIGNR
Ce_EYG-1	-----MAAQOPLS		QEQLLHFVGAH	NTMG-KMTRRTSQ	SDGGKPSGNSNA
Crem_EYG-1	-----MASQOPLS		QEQLLHFVGAH	NTMG-KMTRRSQ	HEPGKPPGNSNS

P A I subdomain



Table of protein sequences for various species including Ce_NPAX-3, Dm_Poxn, Hs_PAX1, and others. The sequences are color-coded and aligned with domain markers alpha 4, alpha 5, and alpha 6. A linker region is also indicated.

linker R E D subdomain

Ce_NPAX-3	
Ce_NPAX-4	
Acey_NPAX-4	
Ce_NPAX-1	
Ce_NPAX-2	
Dm_Poxn	APPTPAATPSIARYAKPPALMMNSAG
Md_Poxn	LPPATMALASSYRYADINMPTASNTPAHAAAS
Cc_Poxn	LHRASLTPLSTAHHAHTSTHLTAANAHPHNAAPHALTLPAQFRYSSAAAAAAMSAANESNRGNNAATTTQDARLLASFGSGVEMPIKPAK-HPVAAHMTALTGAGGSTSALAQLSATVSSAVSTGHLSPQDLSY-AA
Sk_Poxn	
Ac_Poxn	
Cg_EKC37735	
Hs_PAX1	
Mm_Pax1	
Hs_PAX9	
Mm_Pax9	
Eb_Pax1/9	
Sk_Pax1	
Cg_EKC39574	
Dm_Poxm	LGHQHTPGTVMGSGSSSG
Dm_Sv	
Sk_Pax2	
Cg_EKC36239	
Hs_PAX2	
Mm_Pax2	
Hs_PAX5	
Mm_Pax5	
Hs_PAX8	
Mm_Pax8	
Bf_Pax2/5/8	
Hs_PAX3	
Mm_Pax3	
Hs_PAX7	
Mm_Pax7	
Cg_EKC41820	
Sk_LOC100371646	
Dm_Gsb	
Dm_Gsbn	
Dm_Prd	
Hs_PAX4	
Mm_Pax4	
Hs_PAX6	
Mm_Pax6	
Bf_Pax6	
Sk_Pax6	
Dr_Pax10	
Dm_Toy	
Od_Pax6	
Dm_Ey	
Dpse_Ey	LSSSTDLMQ
Cg_EKC31583	LSSGGDLMQ
Dm_Eyg	
Md_Eyg	
Dm_Toe	
Am_Eyg	
Cbir_X777_10953	
Nvit_Eyg	
Apis_LOC100165522	
Sk_Eygone	
Spur_LOC574563	TEQA
Cg_EKC38178	
Ce_EYG-1	
Crem_EYG-1	

