

VVCCD4a	MDAFSSS-FLSSTFTFPSLTTRP---PIAPSSLPQIPSLNISAVRIEEKQPSLTAETSS	56
VVCCD4b	MNPLFCP-FLSSTLPHPKPLVSPSLTTTTPSSSPYPPFLHISAIRNVEDKLHSTFYATPT	59
AtNCED4	MDSVSSSSFLSSTFSLHHSLLRR---RSSS---PTLLRINSVVEERSPITNPSDND	52
VVCCD4c	---MITP-MVSFAFFQIQPLQRS---NQPNSPISFFNGNIAQYCSLPKKNMIIASELST	52
VvCCD1.1	-----	
VVCCD1.2	-----	
AtCCD1	-----	

VVCCD4a	QSSKTQVHKPPPPPPRRAALPTRNIPKKGAAPSLPVTIFNALDDVINNFIDPPLRSSVD	116
VVCCD4b	TS-----QFPEIPTTVITAKKRPVPSLLVTIFNGLDDFINNFIDPPLPPSID	106
AtNCED4	RRNK-----PKTLHNRTNHTLVSSPPKLRPEMTLATALFTTVEDVINTFIDPPSRPSVD	106
VVCCD4c	SEAP-----SPKKLPNIGSDRDTNSEGLGKGALEFLESFLSFFTSLLYFINPPLHPSVD	108
VvCCD1.1	-----MAEKEEQGGTGVVVDPKPSKGFTSKAVDWLEKLIVKLMYDSSQ	44
VVCCD1.2	-----MAEKEEQGGAGVVVDPKPSKGFTSKAVDWLEKLIVKLMYDSSQ	44
AtCCD1	-----MAEKLSDGSS-IISVHPRPSKGFSSKLLDLLERLVVKLMHDASL	43

: . : :: :

VVCCD4a	PRYVLSQNFAP-VEELPPTCEVTDGSLPPWLDGAYIRNGPNPQFLPRGPHYHFDGDGML	175
VVCCD4b	PKHVLSGNFAP-VDELPPTECEVIEGSLPPCLDGAYIRNGPNPQFYPRGPHHFDGDGML	165
AtNCED4	PKHVLSDNFAPVLDDELPPDCEIIGHGTLPLSLNGAYIRNGPNPQFLPRGPHYHFDGDGML	166
VVCCD4c	PKHVLITGNFAQ-VDELPPIDCLVVEGELPQSLNGTYIRNGPNPLHQPRGPHHFEFGDGM	167
VvCCD1.1	PLHYLSGNFAPVRDETPPKNLPVIGYLPCLNGEFVVRVGNPKFSPVAGYHWFDDGMI	104
VVCCD1.2	PLHYLSGNFAPVRDETPPKNLPVIGYLPCLNGEFVVRVGNPKFSPVAGYHWFDDGMI	104
AtCCD1	PLHYLSGNFAPVRDETPPKNLPVIGYLPCLNGEFVVRVGNPKFSDAVAGYHWFDDGMI	103

* : * : *** : * * * . * * * * : * : * * * * . . . : * * : * * * :

VVCCD4a	HSIRISQGRAILCSRYVKTYKYTIERRAGSPILPN-----	210
VVCCD4b	HSIRISHGRPIFCSRYVKTYKYIIEKRAGSPVIPNLFSSYR---SFARSAVAIARLLTG	221
AtNCED4	HAIKIHNGKATLCSRYVKTYKYINVEKQTGAPVMPNVFSGFNGVTASVARGALTAARVLTG	226
VVCCD4c	HSIRLSDGRATFCSRYVKTYKYALEDNVGFPIFPNLSGFHS-VVDLGRCAIAIGRVMKG	226
VvCCD1.1	HGLRIKDGKATYVSRVYVRTSLRKQEEYFGGAKFTRFGDLKG--LFGLLMVNMQMLRAKLK	162
VVCCD1.2	HGLHIKDGKATYVSRVYVRTSLRKQEEYFGGAKFMRIKDLKG--LFGLLMVNMQMLRAKLK	162
AtCCD1	HGVRIKDGKATYVSRVYKTSRLKQEEFFGAAGFMKIGDLKG--FFGLLMVNVQQLRTKLK	161

* . : : . * : . * * * : * : * * * . : .

VVCCD4a	-FNPVNGIGLANTSLALFGGRLYALGESDLPYSRLRLKPDGDIETLGRHDFDGKLVMSMTA	269
VVCCD4b	QFDPVNGVGLANTSVAFFCGHLYALAESDLPYAVRLTPDGIKTIGRYDFDGKLSMSMTA	281
AtNCED4	QYNPVNGIGLANTSLAFFNSRLFALGESDLPYAVRLTESGDIETIGRYDFDGKLAMSMTA	286
VVCCD4c	QIDLRLKGFGLANTSLSLFNNLRFALGESDLPYSIHLSEEGDIETIGRCDFDGKAFINMTA	286
VvCCD1.1	ILDVSYGTGTGNTALVYHHGKLLALSEADKPYVLKVLLEDGDLQTLGLLDYDKRLTHSFTA	222
VVCCD1.2	ILDVSYGTGTGNTALVYHHGKLLALSEADKPYVLKVLLEDGDLQTLGLMDYDKRLTHSFTA	222
AtCCD1	ILDNTYGNGTANTALVYHHGKLLALQEADKPYVLKVLLEDGDLQTLGIIDYDKRLTHSFTA	221

: * * . * : : . : * * * * * * : : . * : : * * : : . * * :

VVCCD4a	HPKVDPETGEAFAFRYGPVPPFLTYFRFDAQGRKQPDVPFISLTSFSLHDFGITHKYAI	329
VVCCD4b	HPKIDPSTGEAFAFRYSPVRPFLTYFRFDAQGRKQPDVPFISLSCPSFFHDFAITNRYAI	341
AtNCED4	HPKTDPIETGETFAFRYGPVPPFLTYFRFDSAGKKQPDVPFISMTSPSLHDFAITKRHAI	346
VVCCD4c	HPKIDPETGETFAFRCSPIPPYITFFSIDKEGSKQPDVPFISMTDPTFVHDFSITKQYIV	346
VvCCD1.1	HPKVDPFTGEMFTFGYSHTPPYITYRVISKDGFMEHPVPIT-ISDPIMMHDFAITENYAI	281
VVCCD1.2	HPKVDPFTGEMFTFGYSHTPPYITYRVISKDGFMEHPVPIT-ISDPIMMHDFAITENYAI	281
AtCCD1	HPKVDPVTGEMFTFGYSHTPPYLYRVISKDGMHDPVPIT-ISEPIMMHDFAITETYAI	280

* * * * * * : * . * : : : . : * : : * * : : * : : * * : : : :



VVCCD4a	FADIQIGM--NPVEMVT-GGSPVGTVPNKVPRLGIIIPRYAKDESEMRWFNVPGFNIVHSI	386
VVCCD4b	FPDIQMMW--NPVKMIIRGGSPVGTDPKVPVRVGIIPRYAKDESEMRWIDVPGFNIHAI	399
AtNCED4	FAEIQLGMRMMLDLVLEGGSPVGTDNKTPRLGVIPKYAGDESEMKWFEVPGFNIHAI	406
VVCCD4c	FSESQIEM--NPLRLMMCKGMPVSAELDKVPRIGVLPRYASTDSEIRWFEAPGFNAMI	404
VvCCD1.1	FMDLPLYFR--PKEMVKEKKLIFTFATKKARFVLPKYAKNELHIKWFELPNCFIENNA	339
VVCCD1.2	FMDLPLYFR--PKEMVKEKKLIFTFATKKARFVLPKYAKNELHIKWFELPNCFIENNA	339
AtCCD1	FMDLPMHFR--PKEMVKEKKMIYSFDPKTKARFVLPKYAKDELMIRWFEELPNCFIENNA	338

* : : : : : * . * : : * * : : : : * .



VVCCD4a	NAWDEED--AIIMVAPNLSVEHT---LERLDMIHASVEMVRIDLKTGMVTRHPLSTRN	440
VVCCD4b	NAWDEEDGDIVMVAAPNLIPIEHA---LERMDLVHGSLEKVRIDLKTGTVTRHRLSQWN	455
AtNCED4	NAWDEDDGNSVVLIAPNIMSIEHT---LERMDLVHALVEKVKIDLVTGIVRRHPISARN	462
VVCCD4c	NAWEEGD-EEIILVAPNAISIENL---FHSIEKVHFSLEKVRINLRSGSVTRTTLSQKN	459
VvCCD1.1	NAWEEDEDEVVLITCRLENPDLDLVGNVKEKLENFANELYEMRFNMKTGIASQRKLSASS	399
VVCCD1.2	NAWEEDEDEVVLITCRLENPDLDLVGGDVKEKLENFGNELYEMRFNMKTGIASQRKLSASS	399
AtCCD1	NAWEEDEDEVVLITCRLENPDLDVSGVKVKEKLENFGNELYEMRFNMKTGSASQKLSASA	398

* * : * * : : : : . : : : : : : * : : : *

VVCCD4a	LDFAVINPGYVGKKNKYVYAAVGNPMPKISGVVKLDVVSQTERK-ECIVGSR-----MY	492
VVCCD4b	LEFAVINPGYLGKKNKYVYSAVGDPLPKISGIVKLDVSRSDRRQECIVAKR-----MY	508
AtNCED4	LDFAVINPAFLGRCSRYYVAAIGDPMPKISGVVKLDVSKGDRD-DCTVARR-----MY	514
VVCCD4c	LELGINSYVGKKNKYGYMGIGKMPKMSGVVKIDL-----LECEVSR-----LY	507

```

VvCCD1.1      VDFPRVNESYTGRRQRYVYGTILDSIAILDSIAKFDLHAEFDTGKSKLEVGGNVQGIFDL 459
VVCCD1.2      VDFPRVNESYTGRRQRYVYGTILDSIAKVTGIIKFDLHAEFDTGKSKLEVGGNVQGIFDL 459
AtCCD1        VDFPRINECYTGKKQRYVYGTILDSIAKVTGIIKFDLHAEAEATGKRMLEVGGNIKGIYDL 458
               ::  : *  : *  : : *  *  : . . . : : * : * : . :

VVCCD4a       GPGCYGGEPPFFVAREPDNPEAEEDDGYIVSYVHDEKSGESKFLVMDAKTPNLDIVAAVRL 552
VVCCD4b       EPGCYGGEPPFFVAKPEPDNPEAEEDDGYVLSYVHDEQSGKSRFIVMDAQSPDLDIVAAVKL 568
AtNCED4       GSGCYGGEPPFFVARDPGNPEAEEDDGYVVTYVHDEVTGESKFLVMDAKSPLEIVAAVRL 574
VVCCD4c       GAGCFGGEPLFVAKDG---ASEEDDGYIVSYVHDEKSGASRFVVMDAKSQTLDVVAVKL 564
VvCCD1.1      GVGGRFGSEAVFVPREPGI-TSEEDDGYLIFVHDEKTKSYVNVINAKTMSDPVAIVEL 518
VVCCD1.2      GVGGRFGSEAVFVPREPGI-TSEEDDGYLIFVHDEKTKSYVNVIDAKTMSDPPIAIVEL 518
AtCCD1        GEGRYGSEAIYVPRE----TAEEDDGYLIFVHDENTGKSCVTVIDAKTMSAEPVAVVEL 514
               *  : * . * . : : : : : : : : : : * * . * : : : : : : : : : : : : * * * *
               ↑
VVCCD4a       PRRVPYGFHGLFVRERDIKGL--- 573
VVCCD4b       PTRVPYGFHGLFVKGCDLKMD--- 589
AtNCED4       PRRVPYGFHGLFVKESDLNKL--- 595
VVCCD4c       PRRVPYGFHGLFVKDGDIREIH-- 586
VvCCD1.1      PNRVPYGFHAFFVTEEQLEQAKL 542
VVCCD1.2      PNRVPYGFHAFFVTEEQLEQAKL 542
AtCCD1        PHRVPYGFHALFVTEEQLEQTLI 538
               *  * * * * * : : : : : : : : : : : :
               ↑

```

Additional file 5. Clustal multiple protein alignments of carotenoid cleavage dioxygenase encoding sequences of *A. thaliana* and (At-) and *V. vinifera* (Vv) orthologues. Arrows indicate the conserved histidine residues.