11-24-2020

DCT4—A New Member of the Dicarboxylate Transporter Family in C4 Grasse

Sarit Weissmann
Donald Danforth Plant Science Center

Pu Huang
Donald Danforth Plant Science Center

Madeline A. Wiechert
Donald Danforth Plant Science Center

Koki Furuyama
Nagoya University

Thomas P. Brutnell
Biotechnology Research Institute

See next page for additional authors

DCT4—A New Member of the Dicarboxylate Transporter Family in C_4 Grasses

Sarit Weissmann[1], Pu Huang[1,†], Madeline A. Wiechert[1], Koki Furuyama[2], Thomas P. Brutnell[3], Mitsutaka Taniguchi[2], James C. Schnable[4,*], and Todd C. Mockler[1,*]

[1]Donald Danforth Plant Science Center, St. Louis, Missouri, USA

[2]Graduate School of Bioagricultural Sciences, Nagoya University, Aichi, Japan

[3]Chinese Academy of Agricultural Sciences, Biotechnology Research Institute, Beijing, China

[4]Computational Sciences Initiative, Center for Plant Science Innovation, Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Nebraska, USA

[†]Present address: BASF Corporation, Durham, NC, USA

[*]Corresponding authors: E-mails: schnable@unl.edu; tmockler@danforthcenter.org.
Accepted: 24 November 2020

Abstract

Malate transport shuttles atmospheric carbon into the Calvin–Benson cycle during NADP-ME C_4 photosynthesis. Previous characterizations of several plant dicarboxylate transporters (DCT) showed that they efficiently exchange malate across membranes. Here, we identify and characterize a previously unknown member of the *DCT* family, *DCT4*, in *Sorghum bicolor*. We show that SbDCT4 exchanges malate across membranes and its expression pattern is consistent with a role in malate transport during C_4 photosynthesis. *SbDCT4* is not syntenic to the characterized photosynthetic gene *ZmDCT2*, and an ortholog is not detectable in the maize reference genome. We found that the expression patterns of *DCT* family genes in the leaves of *Zea mays*, and *S. bicolor* varied by cell type. Our results suggest that subfunctionalization, of members of the *DCT* family, for the transport of malate into the bundle sheath plastids, occurred during the process of independent recurrent evolution of C_4 photosynthesis in grasses of the PACMAD clade. We also show that this subfunctionalization is lineage independent. Our results challenge the dogma that key C_4 genes must be orthologues of one another among C_4 species, and shed new light on the evolution of C_4 photosynthesis.

Key words: DCT4, new transporter gene, grass evolution, C_4 photosynthesis. .

Significance

Dicarboxylate-transporter-2 (*DCT2*) plays a key role during C_4 photosynthesis in *Zea mays*. Its orthologs are assumed to function the same in related species, as *Z. mays* is the main C_4 reference species. We introduce a new gene, *DCT4*, that assumed the role of *DCT2* in *Sorghum bicolor* and other C_4 grass species. By surveying related C_4 species, we propose that different members of the *DCT* family subfunctionalized for photosynthetic malate transport in the BS cells of C_4 grasses of the PACMAD clade. We suggest that rather than being static, biochemical adaptations continued after the divergence of the PACMAD lineages.

Introduction

Three subtypes of C_4 photosynthesis are generally recognized as defined by the primary decarboxylase in the bundle sheath (BS) cells: Chloroplastic NADP-dependent malic enzyme (NADP-ME); mitochondrial NAD-dependent malic enzyme (NAD-ME); and cytosolic phosphoenolpyruvate carboxykinase (PEPCK) (Hatch and Slack 1966; Hatch 1971; Rathnam and Edwards 1977). Different plant species may contain various

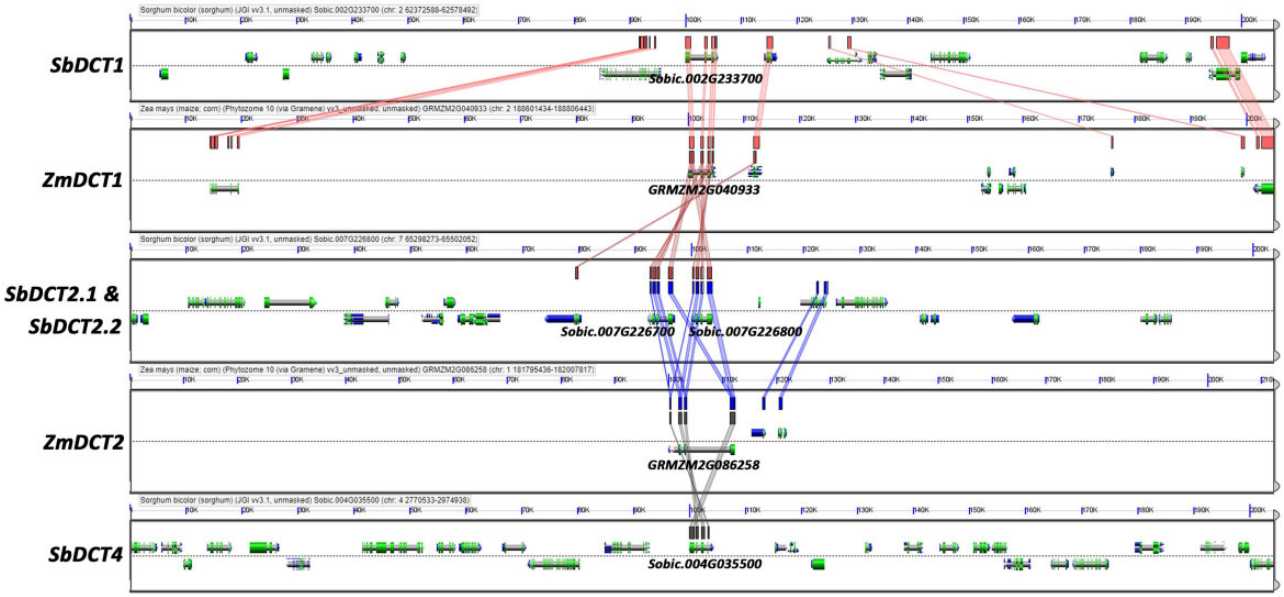


Fig. 1.—CoGe (https://genomevolution.org/coge/) genome viewer screenshots depicting the conservation and genomic contexts of *DCT* genes in *Zea mays* and *Sorghum bicolor*. Colored lines between panels show conserved genes. *SbDCT4* shows high sequence conservation with other *DCT* genes, but is not a syntenic ortholog of *ZmDCT2*, as shown by the lack of conservation in neighboring genes.

combinations of these three subtypes (Hatch 1971; Chapman and Hatch 1979; Furbank 2011; Pick et al. 2011; Wang, Brautigam, et al. 2014). The movement and exchange of malate across membranes, by dicarboxylate transporters (DCTs/DiTs), plays a significant role during photosynthesis in NADP-ME and NAD-ME C_4 species (Ding et al. 2015). In C_3 plants, DCTs are crucial to nitrate assimilation, such as the GS/GOGAT cycle and photorespiration (Linka and Weber 2010; Kinoshita et al. 2011). Taniguchi et al. characterized several plant DCTs that efficiently exchange malate across membranes (Taniguchi et al. 2002; 2004). The differential expression of C_4 photosynthesis genes in mesophyll (M) and BS cells (John et al. 2014; Tausta et al. 2014; Wang, Czedik-Eysenberg, et al. 2014) suggests that different malate transporters may be needed to move malate out of the chloroplasts of M cells and into the chloroplasts of BS cells. In *Zea mays*, an NADP-ME C_4 grass, dicarboxylate transporter-2 (*ZmDCT2*, GRMZM2G086258) moves malate into the chloroplast of BS cells during C_4 photosynthesis (Weissmann et al. 2016). *ZmDCT2* plays a critical role during C_4 photosynthesis in *Z. mays*, and its absence severely impairs plant growth and development (Weissmann et al. 2016). The role of DCTs in C_4 photosynthesis in other species, however, remains unknown.

Zea mays is the best characterized and functionally annotated C_4 grass species. As such, it is a useful reference for identification of photosynthesis-related genes in poorly characterized C_4 grasses and for resolving orthology (John et al. 2014; Ding et al. 2015; Huang et al. 2017). Microsynteny, the comparison of collinearity among related species, is a reliable approach to determine orthology and predict the function of a gene (Bennetzen and Freeling 1997; Chen et al. 1997; Tikhonov et al. 1999; Bennetzen 2000; Kumar et al. 2009; Jin et al. 2016). Davidson et al. (2012) showed that syntenic orthologs are likely to have conserved functions and expression patterns across lineages. Here, we identify a new member of the *DCT* family, *DCT4*, which is not syntenic to the photosynthetic gene *ZmDCT2* and is not detected in the maize reference genome. We demonstrate that *S. bicolor DCT4* (*SbDCT4*) efficiently exchanges malate across membranes, consistent with a malate transport role in C_4 photosynthesis. We characterize the diverse expression patterns of *DCT* genes in leaves of multiple grass species. We also propose that subfunctionalization of *DCT*s in grasses of the PACMAD clade (Sanchez-Ken and Clark 2010) occurred during independent recurrent evolution of C_4 photosynthesis.

Results

Identification of *DCT4* in *S. bicolor*

To learn more about C_4-related DCT in species evolutionarily related to maize, we identified the syntenic ortholog of *ZmDCT2* in *S. bicolor*. Two genes, Sobic.007G226700 and Sobic.007G226800, are present at the predicted syntenic orthologous position on chromosome 7. We refer to them as *SbDCT2.1* and *SbDCT2.2*, respectively (fig. 1). *ZmDCT2* is abundantly expressed (table 1), and its expression is enriched in BS cells of maize leaves (fig. 2) (Li et al. 2010; Tausta et al. 2014; Ding et al. 2015). In contrast, the expression profiles of *SbDCT2.1* and *SbDCT2.2* in *S. bicolor* leaves are low (table 1). *SbDCT2.1* expression is slightly enriched in the M cells

Table 1
Genomic Presence or Absence and Whole Leaf TPM Values for Dicarboxylate Transporter Genes in Grass Leaves[a]

Species	Presence in Genomic DNA			RNA-seq TPM Values			Photosynthesis	
	DCT1	*DCT2*	*DCT4*	*DCT1*	*DCT2*	*DCT4*	*OMT1*	
Brachypodium distachyon	+	+	–	33	119	N/A	18	C_3
Oryza sativa	+	+	–	16	11	N/A	101	C_3
Aristida congesta	+	+	+	393	18	0	871	C_4 NADP-ME
Eriachne aristidea	+	+	+	1348	89	0	874	C_4 NADP-ME
Chasmanthium laxum	+	+	–	8	11	N/A	18	C_3
Danthoniopsis dinteri	+	+	+	0	123	1555	305	C_4 NADP-ME
Anthephora pubensis	+	+	+	1611	0	0	514	C_4 NADP-ME
Echinochloa esculenta	+	+	+	1588	0	0	683	C_4 NADP-ME
Urochloa fusca	+	+	–	50	86	N/A	344	C_4 PEPCK
Setaria italica	+	+	+	29	41	658	328	C_4 NADP-ME
Dichanthelium oligosanthes	+	+	–	11	13	N/A	13	C_3
Paspalum vaginatum	+	+	+	591	0	666	354	C_4 NADP-ME
Arundinella hirta	+	+	+	0	23	481	398	C_4 NADP-ME
Sorghum bicolor	+	+	+	126	0[b]	229	483	C_4 NADP-ME
		+			9[c]			C_4 NADP-ME[d]
Zea mays	+	+	–	5	166	N/A	153	C_4 NADP-ME & PEPCK

[a]Note that interspecies comparison is not possible, because expression levels were normalized within each species.
[b]*SbDCT2.1*.
[c]*SbDCT2.2*.
[d]Despite no PEPCK activity detected in *Sorghum*, carbon moves into the BS through aspartate (Chapman and Hatch 1979).

whereas *SbDCT2.2* is enriched in BS cells (fig. 2). We also analyzed the transcript levels of two other *S. bicolor DCT*, *SbDCT1* (Sobic.002G233700) and *SbOMT1* (Sobic.008G112300). These genes are the orthologs of the *Z. mays* genes *ZmDCT1* (GRMZM2G040933) and *Zm*-oxoglutarate/malate transporter 1 (*ZmOMT1*; GRMZM2G383088), respectively. We found that *SbDCT1* expression, similar to that of *ZmDCT1*, is relatively low (table 1), and only slightly differentially expressed in M cells relative to BS cells (fig. 2). The expression of both *ZmOMT1* and *SbOMT1* is relatively high (table 1), and both are slightly enriched in M cells (fig. 2).

In C_4 species, the expression of many photosynthetic genes is enriched in either BS and M cells (Li et al. 2010; John et al. 2014; Tausta et al. 2014; Weissmann et al. 2016; Rao et al. 2016). In NADP-ME species, two transporters, one within the BS cells and another in M cells, move malate in and out of the chloroplast during C_4 photosynthesis (Brautigam et al. 2008; Weissmann and Brutnell 2012; John et al. 2014; Tausta et al. 2014; Wang, Czedik-Eysenberg, et al. 2014). However, in sorghum leaves, we found only one highly expressed *DCT*, *SbOMT1*, that showed slightly enriched expression in M cells. Therefore, we screened the sorghum genome for additional homologs of known maize *DCT*s. We identified the gene Sobic.004G035500 that showed homology to *ZmDCT1* and *ZmDCT2* but was not syntenic to either gene (fig. 1). *DCT3* is the name of the second transcript of *ZmDCT2* (Taniguchi et al. 2004), so we named this new gene *SbDCT4*. No syntenic ortholog of *SbDCT4* is present in the reference genomes of *Z. mays* or *Oryza sativa*. The absence of syntenic conservation between *S. bicolor* and *Z. mays* and the lack of direct orthologs in *Z. mays* or C_3 species prevented identification of *DCT4* in a previous bioinformatic screen for C_4 photosynthesis genes (Huang et al. 2017). The expression of *SbDCT4* is moderately abundant (table 1) and strongly enriched in the BS cells of *S. bicolor* leaves (fig. 2).

SbDCT4 is an Efficient Malate Transporter

To verify the ability of SbDCT4 to transport malate, we cloned coding sequences from the three sorghum *DCT* genes, *SbDCT1*, *SbDCT2*, and *SbDCT4*. We measured the malate transport activities of the recombinant proteins expressed in yeast. SbDCT4 was an efficient malate transporter (table 2). The K_m of SbDCT4 was similar to that of SbDCT1, and the affinity for malate was highest in SbDCT2 among the three SbDCTs (table 2), consistent with the relative malate transport activities reported for maize DCT1 and DCT2 (Taniguchi et al. 2004).

Phylogenetic Distribution of *DCT* Genes in Grasses

To understand the relationship of *SbDCT4* to other grass *DCT* genes, we searched the genomes of the grass species *Setaria italica*, *Urochloa fusca, Brachypodium distachyon*, and *Dichanthelium oligosanthes*. In *S. italica*, an NADP-ME C_4 species, we identified a *DCT*, Seita.9G375100, that showed no syntenic orthologous relationship with *DCT* genes in other available grass genomes. Phylogenetic analysis showed that this gene clustered with *SbDCT4* but not with *SbDCT1* and

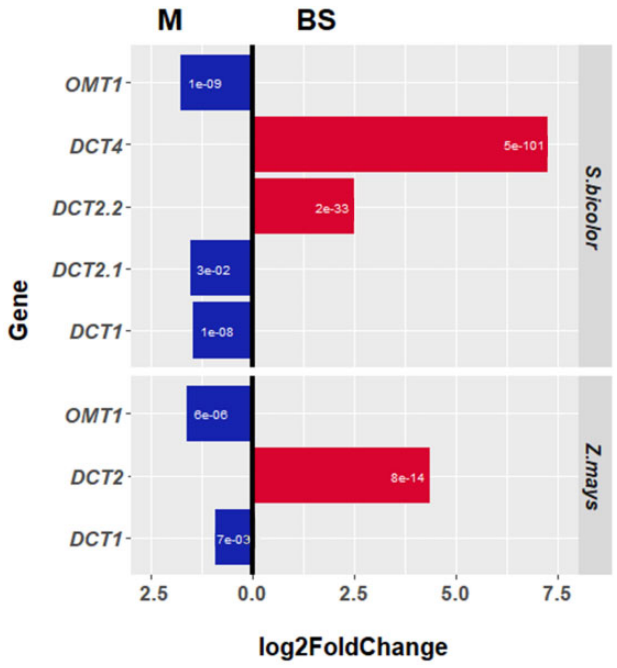


Fig. 2.—Differential expression of malate transporters in *Zea mays*, and *Sorghum bicolor* leaves between the BS and M cells. The genome of *Z. mays* has one copy of *DCT2* and does not contain *DCT4*, and *ZmDCT2* is highly enriched in BS cells. *Sorghum bicolor* has two copies of *DCT2* (*DCT2.1*, and *DCT2.2*) in the syntenic genomic location that are the result of gene duplication. *Sorghum bicolor* also expresses *DCT4*, which is highly enriched in BS cells. Both species express *OMT1* and *DCT1*, which are only slightly enriched in the M cells. Red bars represent enrichment in the BS cells. Blue bars represent enrichment in the M cells. The white numbers inside the bars represent the significance (*P*-value) of the log2(FoldChange).

Table 2

K_m of Malate for Recombinant DCT Proteins[a] Demonstrates the Ability of SbDCT4 to Transport Malate Efficiently

K_m (mM)			
	DCT1	**DCT2**	**DCT4**
Sorghum bicolor	1.24 ± 0.14	0.71 ± 0.10	1.13 ± 0.10
Zea mays[b]	1.1 ± 0.1	0.85 ± 0.44	N/A

[a]The values are the means of three independent experiments ± SE.
[b]Kinetic values from a previous report (Taniguchi et al. 2004).

SbDCT2 (fig. 3). We designated this gene *SiDCT4*. We did not detect orthologs, syntenic or otherwise, in *U. fusca*, a PEPCK C_4 species, or in the two C_3 species. To expand the search for *DCT4* in other grasses currently lacking genome assemblies, we examined leaf-derived transcript assemblies for *Aristida congesta, Eriachne aristidea, Chasmanthium laxum, Danthoniopsis dinteri, Anthephora pubensis, Echinochloa esculenta, Paspalum vaginatum*, and *Arundinella hirta* (Huang P, Mayfield-Jones D, Schnable J, Brutnell T, manuscript in preparation). We then used the predicted coding sequences of the *DCT* genes from available genomes and from the de novo leaf transcriptome assemblies to generate a phylogenetic tree of the *DCT* family. The resulting phylogeny shows that *DCT4* transcripts form a distinct subclade from the *DCT1* clade (fig. 3). The absence of *DCT4* transcript expression does not rule out the existence of the gene in the genome. We also used polymerase chain reaction (PCR) to survey for *DCT* genes in the genomes of grass species for which whole-genome assemblies were not available. We designed conserved primers (nondegenerate or minimally degenerate) to small regions unique to each of the three *DCT* genes using PrimaClade (Gadberry et al. 2005). We detected *DCT1* and *DCT2* in the genomes of all species tested (table 1, supplemental fig. 1, Supplementary Material online). *DCT4*, however, was detected only in the genomes of NADP-ME C_4 species of the PACMAD clade, excluding *Z. mays* (table 1, supplemental fig. 2, Supplementary Material online).

Expression of Malate Transporter Genes in NADP-ME C_4 Grasses

C_3 species and *U. fusca*, a PEPCK C_4 species, express both *DCT1* and *DCT2* at low levels in leaves (table 1). C_4 NADP-ME species of the PACMAD clade generally express one *DCT* gene in leaves at a high level and also express one or two other *DCT* genes at low levels (table 1). We did not find an apparent lineage-specific pattern for the expression of the predominant *DCT* gene in the NADP-ME species we analyzed. This finding is consistent with random evolutionary processes underlying the subfunctionalization of members of the *DCT* family. Interestingly, *Z. mays* is the only species we examined in which *DCT2* is the predominantly expressed *DCT* gene (table 1). We also examined the expression of the non-*DCT* malate transporter *OMT1* gene in the leaves of grasses (table 1). Interestingly, we found that although *OMT1* expression was generally abundant, there was no consistent pattern of relative expression between the *DCT* and *OMT* genes within the NADP-ME C_4 species (table 1).

Discussion

Evolution of the *DCT* Gene Family in Grasses

We identified *DCT4* as a new member of the *DCT* gene family in the grasses (fig. 1). Our analysis suggests that *DCT4* is present in some C_4 NADP-ME PACMAD grasses. *DCT1* and *DCT2* appear to have originated from a duplication of a single *DCT* gene after the monocot–eudicot split (Taniguchi et al. 2004) and *DCT4* arose from a duplication of *DCT1* at the root of the PACMAD grasses (fig. 3). The expression of *DCT* genes in the grasses that we analyzed exhibited no clear lineage-specific patterns (table 1). Therefore, we propose that different members of the *DCT* family subfunctionalized for photosynthetic malate transport in the BS cells of C_4 grasses of the PACMAD clade.

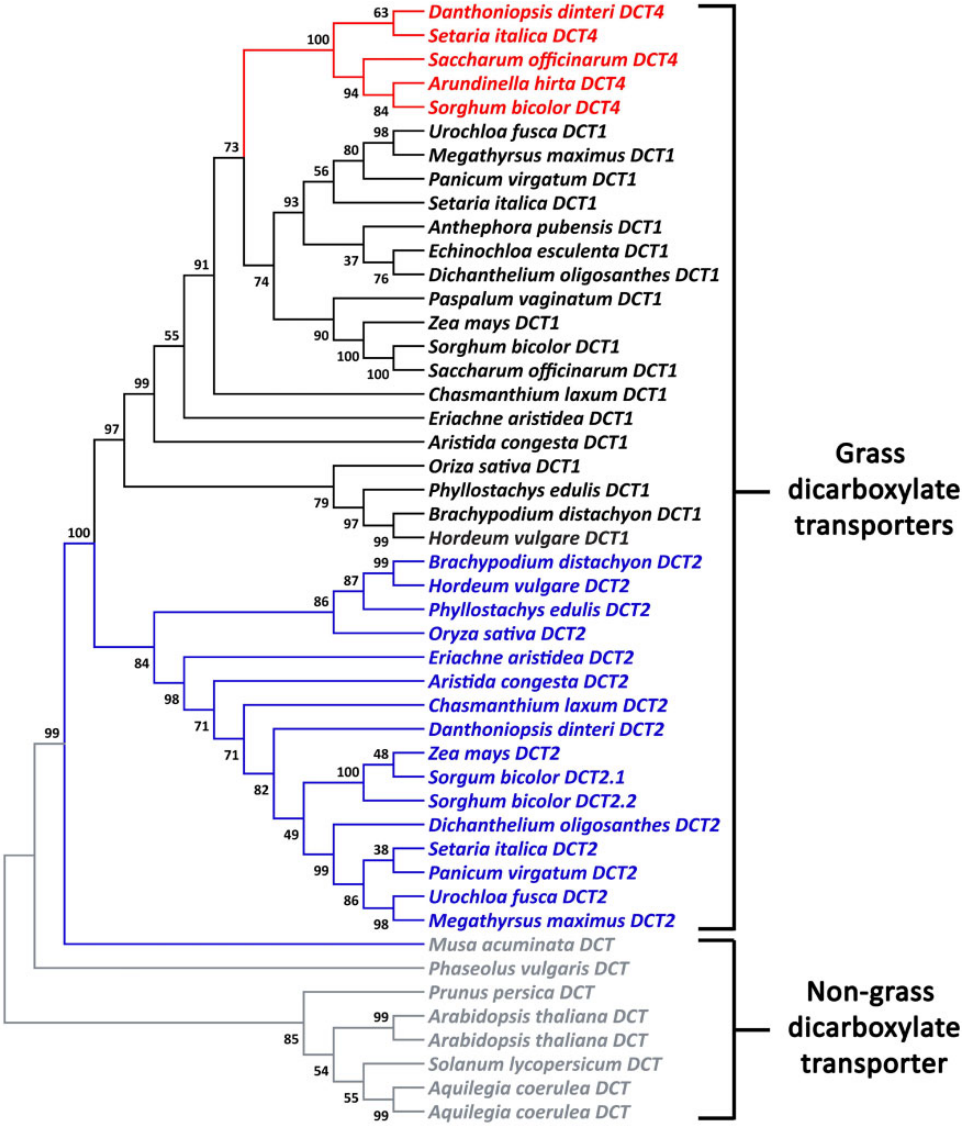


Fig. 3.—A phylogenetic tree of the *DCT* family in the grasses showing that *DCT4* is a subclade of *DCT1*. The *DCT1*, *DCT2*, and *DCT4* gene lineages are black, blue, and red, respectively. The length of the branches represents the evolutionary distance between ancestor to descendent nodes. The numbers represent the confidence level of the specific branch.

This work also challenges the dogma that key C_4 genes must be orthologues of one another, among species, and show that they can be paralogs. This confirms the importance of including syntenic and expression data in assigning orthology across species, and of developing multiple models for C_4 photosynthesis in the grasses. For example, *SiDCT4* was previously misannotated as the ortholog of *ZmDCT2* (John et al. 2014), likely because of the lower expression level of *SiDCT2* (Seita.9G375100) in leaf tissue. The use of different malate transporters, for example, DCT4 in *S. bicolor* and *S. italica*, or DCT2 in *Z. mays*, suggests that multiple evolutionary paths resulted in the development of an active C_4 NADP-ME photosynthetic cycle. It is interesting to note that common origins of C_4 photosynthesis are often defined based on the predominant decarboxylase utilized, thus maize and sorghum are considered to have evolved from a common C_4 ancestor. This analysis suggests that rather than being static, biochemical adaptations continued after the divergence of maize and sorghum lineages. Thus, optimizations of C_4 activities may be continuous as breeding pressures or climate change alters ecological niches of individual species.

Various C_4 Subtype Combinations Have Different Transport Requirements

The variation of expression levels among the different malate transporters within each NADP-ME species (table 1) suggests different transport requirements during C_4 photosynthesis.

This supposition is in agreement with the view that the three subtypes of C_4 photosynthesis are mixed rather than exclusive (Hatch 1971; Chapman and Hatch 1979; Furbank 2011; Pick et al. 2011; Wang, Brautigam, et al. 2014). For example, *Z. mays* utilizes both the NADP-ME (75%) and PEPCK (25%) pathways to fix carbon (Chapman and Hatch 1979; Wingler et al. 1999; Weissmann et al. 2016), and has similar expression levels of *DCT2* and *OMT1* and low expression of *DCT1*. *Sorghum bicolor* moves carbon through both malate and aspartate, although no PEPCK activity was detected in its leaves (Chapman and Hatch 1979). *Sorghum bicolor* has similar expression levels for *DCT4* and *DCT1* and high expression of *OMT1* (table 1). Other grass species may have *DCT* expression ratios that correspond to their unique combination of C_4 subsystems. For example, *OMT1* is highly expressed in *U. fusca*, ~3- to 7-fold higher than *DCT2* or *DCT1*, respectively. OMT1 transports dicarboxylates, excluding those containing an amino group (Taniguchi et al. 2002, 2004). Thus, in PEPCK C_4 plants, OMT1 may move oxaloacetate into the mesophyll chloroplast, and 2-oxoglutarate out, to support the high production of aspartate needed to maintain the photosynthetic cycle (Rathnam and Edwards 1977). Interestingly, both *ZmOMT1* and *SbOMT1* are only slightly differentially expressed in the M cells (fig. 2). As the loss of *DCT2* in *Z. mays* prevents movement of malate into the BS chloroplast (Weissmann et al. 2016), *OMT1* cannot be moving malate into the BS chloroplast alongside *DCT2*. But *OMT1* may also have a role in organic acid metabolism in both cell types, such as shuttling reducing equivalents in organelles other than the chloroplast (Pleite et al. 2005).

Conclusions

Our results show that the newly identified member of the *DCT* gene family, SbDCT4, is an efficient malate transporter. Based on the expression patterns of malate transporters among the grasses, we suggest that different members of the *DCT* family may have evolved multiple roles in C_4 photosynthesis. Further studies will be needed to verify the subcellular localization of these proteins and to define their specific metabolic functions. Characterizing the various combinations of C_4 photosynthetic subsystems in grasses will facilitate the exploitation of *DCT* genes, through breeding or engineering, to improve the performance of crop plants and increase yield.

Materials and Methods

Identification of *DCT4* Genes in *Sorghum* and Other Grasses

We used QUOTA-ALIGN (Tang et al. 2011) to identify syntenic orthologous regions in grass species with sequenced genomes, following the protocol described in Zhang et al. (2017). To find homologous genes at nonsyntenic locations, we used two complementary approaches. For species with sequenced genomes, we used LASTZ (Harris 2007) to align the coding sequence of the primary transcript annotated in Phytozome (https://phytozome-next.jgi.doe.gov, last accessed January 6, 2021) to the genome assembly. For species without assembled genomes, we used LASTZ to align the coding sequence of the primary transcript from Phytozome to transcript assemblies generated by Trinity (Grabherr et al. 2011).

Measurements of Malate Transport

We cloned each of the three *SbDCT* cDNAs between the promoter and terminator of yeast *GAL2* in the pTV3e vector (Nishizawa et al. 1995). We transformed the plasmids into yeast LBY416 cells and selected transformants on tryptophan-deficient agar plates. We prepared a crude membrane fraction from the selected yeast transformants. We used a freeze-thaw technique to reconstitute liposomes for the measurement of the uptake of [^{14}C]malate (Taniguchi et al. 2002).

Phylogenetic Analysis of *DCT* Homologs

DCT coding sequences for *Z. mays*, *S. bicolor*, *S. italica*, *B. distachyon*, *O. sativa*, *D. oligosanthes*, and *U. fusca* were from Phytozome (https://phytozome-next.jgi.doe.gov, last accessed January 6, 2021). We used BlastN (Altschul et al. 1990) to search *de novo* assembled leaf transcriptomes (Huang P, Mayfield-Jones D, Schnable J, Brutnell T, manuscript in preparation) from the C_4 grass species *A. congesta*, *E. aristidea*, *C. laxum*, *D. dinteri, A. pubensis*, *E. esculenta*, *P. vaginatum*, and *A. hirta* with the *DCT* sequences from maize, *Setaria*, and *Sorghum* as queries. We used ProGraphMSA to generate a codon-based sequence alignment (Szalkowski 2012). We used MEGA6 (Tamura et al. 2013), with default parameters and the branch support values based on 1,000 bootstraps, to generate the phylogenetic reconstruction with the maximum likelihood method and based on the nucleotides in the third position of codons (Simmons et al. 2006).

Analysis of Gene Expression for Decarboxylase Transporters in Grasses

For species with published leaf transcriptome profiles (Ouyang et al. 2007; Li et al. 2010; Zhang et al. 2012; Schnable 2014; Wang, Czedik-Eysenberg, et al. 2014; Studer et al. 2016), gene expression levels were calculated and normalized, for each species, as Transcripts Per Million (TPM). For the other species, the normalized TPM values were based on de novo transcriptome assemblies (Huang P, Mayfield-Jones D, Schnable J, Brutnell T, manuscript in preparation). The values in table 1 only allow for intraspecies comparisons among the decarboxylase transporters.

Table 3
Primers and PCR Conditions for the Amplification of Grass *DCT* Genes

Primer Pair	Forward 5′-3′[a]	Reverse 5′-3′[a]	Cycling Conditions
DCT1	CACCAACGAGGTCATCTGG	AGTAGGTGGCGATDCGGTC	94 °C 3 min, [94 °C 45 s, 58 °C 30 s, 72 °C 1 min] × 30, 72 °C 10 min, 4 °C ∞
DCT2	CVTGGATGTCRAATTGTGTTG	TGGCTTGCAAABADATAGTGAA	94 °C 3 min, [94 °C 45 s, 58–52 °C (–0.5 °C/cycle) 30 s, 72 °C 1 min] × 14, [94 °C 45 s, 52 °C 30 s, 72 °C 1 min] × 16, 72 °C 10 min, 4 °C ∞.
DCT4	CTTYGTCAAGTGGCTCGG	GACTTGATGATSGGCAGGA	94 °C 3 min, [94 °C 45 s, 60 °C 30 s, 72 °C 1 min] × 30, 72 °C 10.0 min, 4 °C ∞

[a]B = C + G + T, D = A + G + T, R = A + G, S = C + G, V = A + C + G, Y = C + T.

Identification of *DCT4* in Species without Sequenced Genomes

We aligned the coding sequences from each of the *DCT* genes from *Z. mays*, *S. bicolor*, *S. italica*, *U. fusca*, *B. distachyon*, *O. sativa*, *D. oligosanthes*, *A. congesta*, *E. aristidea*, *C. laxum*, *D. dinteri, A. pubensis*, *E. esculenta*, *P. vaginatum*, and *A. hirta* using PAL2NAL (Suyama et al. 2006). The resulting multiple sequence alignment enabled the design of non-degenerate or minimally degenerate PCR primers (table 3) using PrimaClade (Gadberry et al. 2005).

Jacob D. Washburn and J. Chris Pires (University of Missouri, Columbia) kindly provided genomic DNA from *A. congesta*, *E. aristidea*, *D. dinteri*, *A. pubensis*, *E. esculenta*, and *A. hirta* (Washburn et al. 2015). We used a CTAB-based method to extract genomic DNA from *C. laxum*, *P. vaginatum*, *Z. mays*, *S. bicolor*, *S. italica*, and *B. distachyon* (Weissmann et al. 2016). *Zea mays* and *B. distachyon* were the negative controls for *DCT4* and the positive controls for *DCT1* and *DCT2*. *Sorghum bicolor* and *S. italica* were the positive controls for *DCT1*, *DCT2*, and *DCT4*.

We conducted amplification of *DCT* genes by PCR using a 25-μl reaction mix and an ABI 2720 Thermal cycler. The reaction mixture included 2.5 μl of 10× Buffer, 2.5 μl of 10 μM solutions of forward and reverse primers, 2 μl of 2.5 mM dNTP stock, 14 μl of nuclease-free water, 0.5 μl of Choice Taq enzyme, and 1 μl of 100 ng/μl DNA. We performed PCR reactions as described in table 3 with 5 μl of loading dye added to each reaction. Aliquots of 13 μl were loaded on 3% agarose gels (Invitrogen UltraPure Agarose 1000, 1× TAE buffer, Invitrogen SYBR Safe Gel Stain) and electrophoresed for 30 min at 100 volts. We based size estimates on 100 bp and 50 bp DNA markers (GoldBio).

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We acknowledge funding from National Science Foundation (Award No. IOS-1546882) and Department of Energy BER (Award No. DE-SC0018277) to T.C.M. We would like to thank Dr Elizabeth Kellogg and Dr Doug Allen for their advice and consultations during the writing of this manuscript.

Data Availability

This work includes no new sequence data.

Literature Cited

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215(3):403–410.
Bennetzen JL. 2000. Comparative sequence analysis of plant nuclear genomes: microcolinearity and its many exceptions. Plant Cell. 12(7):1021–1029.
Bennetzen JL, Freeling M. 1997. The unified grass genome: synergy in synteny. Genome Res. 7(4):301–306.
Brautigam A, Hoffmann-Benning S, Weber AP. 2008. Comparative proteomics of chloroplast envelopes from C_3 and C_4 plants reveals specific adaptations of the plastid envelope to C_4 photosynthesis and candidate proteins required for maintaining C_4 metabolite fluxes. Plant Physiol. 148(1):568–579.
Chapman KS, Hatch MD. 1979. Aspartate stimulation of malate decarboxylation in *Zea mays* bundle sheath cells: possible role in regulation of C_4 photosynthesis. Biochem Biophys Res Commun. 86(4):1274–1280.
Chen M, et al. 1997. Microcolinearity in *sh2*-homologous regions of the maize, rice, and sorghum genomes. Proc Natl Acad Sci USA. 94(7):3431–3435.
Davidson RM, et al. 2012. Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. Plant J. 71(3):492–502.
Ding Z, et al. 2015. Identification of photosynthesis-associated C_4 candidate genes through comparative leaf gradient transcriptome in multiple lineages of C_3 and C_4 species. PLoS One 10(10):e0140629.
Furbank RT. 2011. Evolution of the C_4 photosynthetic mechanism: are there really three C_4 acid decarboxylation types? J Exp Bot. 62(9):3103–3108.
Gadberry MD, Malcomber ST, Doust AN, Kellogg EA. 2005. Primaclade—a flexible tool to find conserved PCR primers across multiple species. Bioinformatics 21(7):1263–1264.
Grabherr MG, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 29(7):644–652.
Harris RS. 2007. Improved pairwise alignment of genomic DNA. Philadelphia (PA): Pennsylvania State University.
Hatch MD. 1971. The C_4 pathway of photosynthesis. Evidence for an intermediate pool of carbon dioxide and the identity of the donor C_4 dicarboxylic acid. Biochem J. 125(2):425–432.

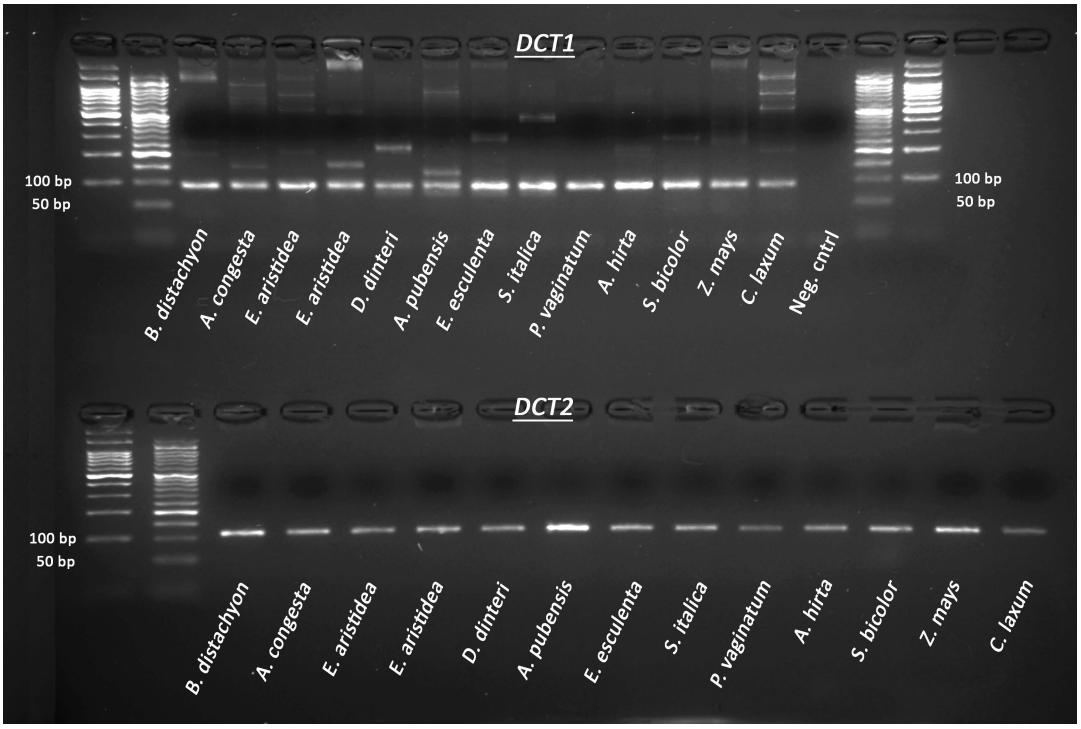
Hatch MD, Slack CR. 1966. Photosynthesis by sugar-cane leaves. A new carboxylation reaction and the pathway of sugar formation. Biochem J. 101(1):103–111.
Huang P, Studer AJ, Schnable JC, Kellogg EA, Brutnell TP. 2017. Cross species selection scans identify components of C_4 photosynthesis in the grasses. J Exp Bot. 68(2):127–135.
Jin J, et al. 2016. High level of microsynteny and purifying selection affect the evolution of WRKY family in Gramineae. Dev Genes Evol. 226(1):15–25.
John CR, Smith-Unna RD, Woodfield H, Covshoff S, Hibberd JM. 2014. Evolutionary convergence of cell-specific gene expression in independent lineages of C_4 grasses. Plant Physiol. 165(1):62–75.
Kinoshita H, et al. 2011. The chloroplastic 2-oxoglutarate/malate transporter has dual function as the malate valve and in carbon/nitrogen metabolism. Plant J. 65(1):15–26.
Kumar S, Mohan A, Balyan HS, Gupta PK. 2009. Orthology between genomes of *Brachypodium*, wheat and rice. BMC Res Notes. 2(1):93.
Li PH, et al. 2010. The developmental dynamics of the maize leaf transcriptome. Nat Genet. 42(12):1060–U1051.
Linka N, Weber APM. 2010. Intracellular metabolite transporters in plants. Mol Plant. 3(1):21–53.
Nishizawa K, Shimoda E, Kasahara M. 1995. Substrate recognition domain of the *Gal2* galactose transporter in yeast *Saccharomyces cerevisiae* as revealed by chimeric galactose-glucose transporters. J Biol Chem. 270(6):2423–2426.
Ouyang S, et al. 2007. The TIGR rice genome annotation resource: improvements and new features. Nucleic Acids Res. 35(Database):D883–887.
Pick TR, et al. 2011. Systems analysis of a maize leaf developmental gradient redefines the current C_4 model and provides candidates for regulation. Plant Cell. 23(12):4208–4220.
Pleite R, Pike MJ, Garcés R, Martínez-Force E, Rawsthorne S. 2005. The sources of carbon and reducing power for fatty acid synthesis in the heterotrophic plastids of developing sunflower (*Helianthus annuus* L.) embryos. J Exp Bot. 56(415):1297–1303.
Rao X, et al. 2016. Comparative cell-specific transcriptomics reveals differentiation of C_4 photosynthesis pathways in switchgrass and other C_4 lineages. J Exp Bot. 67(6):1649–1662.
Rathnam CKM, Edwards GE. 1977. C_4 acid decarboxylation and CO_2 donation to photosynthesis in bundle sheath strands and chloroplasts from species representing 3 groups of C_4 plants. Arch Biochem Biophys. 182(1):1–13.
Sanchez-Ken JG, Clark LG. 2010. Phylogeny and a new tribal classification of the Panicoideae s.l. (Poaceae) based on plastid and nuclear sequence data and structural data. Am J Bot. 97(10):1732–1748.
Simmons MP, Zhang LB, Webb CT, Reeves A. 2006. How can third codon positions outperform first and second codon positions in phylogenetic inference? An empirical example from the seed plants. Syst Biol. 55(2):245–258.
Studer AJ, et al. 2016. The draft genome of the C_3 panicoid grass species *Dichanthelium oligosanthes*. Genome Biol. 17(1):223.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34(Web Server):W609–W612.
Szalkowski AM. 2012. Fast and robust multiple sequence alignment with phylogeny-aware gap placement. BMC Bioinformatics 13(1):129.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 30(12):2725–2729.
Tang H, et al. 2011. Screening synteny blocks in pairwise genome comparisons through integer programming. BMC Bioinformatics 12(1):102.
Taniguchi M, et al. 2002. Identifying and characterizing plastidic 2-oxoglutarate/malate and dicarboxylate transporters in *Arabidopsis thaliana*. Plant Cell Physiol. 43(7):706–717.
Taniguchi Y, et al. 2004. Differentiation of dicarboxylate transporters in mesophyll and bundle sheath chloroplasts of maize. Plant Cell Physiol. 45(2):187–200.
Tausta SL, et al. 2014. Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C_4 related processes. J Exp Bot. 65(13):3543–3555.
Tikhonov AP, et al. 1999. Colinearity and its exceptions in orthologous *adh* regions of maize and sorghum. Proc Natl Acad Sci USA. 96(13):7409–7414.
Urochloa fusca LBJWC-52 annotated standard draft [Internet]. 2014. Joint Genome Institute. Available from: https://genome.jgi.doe.gov/portal/UrofusStandDraft_FD/UrofusStandDraft_FD.info.html.
Wang L, Czedik-Eysenberg A, et al. 2014. Comparative analyses of C_4 and C_3 photosynthesis in developing leaves of maize and rice. Nat Biotechnol. 32(11):1158–1165.
Wang Y, Brautigam A, Weber APM, Zhu XG. 2014. Three distinct biochemical subtypes of C_4 photosynthesis? A modelling analysis. J Exp Bot. 65(13):3567–3578.
Washburn JD, Schnable JC, Davidse G, Pires JC. 2015. Phylogeny and photosynthesis of the grass tribe Paniceae. Am J Bot. 102(9):1493–1505.
Weissmann S, Brutnell TP. 2012. Engineering C_4 photosynthetic regulatory networks. Curr Opin Biotechnol. 23(3):298–304.
Weissmann S, et al. 2016. Interactions of C_4 subtype metabolic activities and transport in maize are revealed through the characterization of *DCT2* Mutants. Plant Cell. 28(2):466–484.
Wingler A, Walker RP, Chen ZH, Leegood RC. 1999. Phosphoenolpyruvate carboxykinase is involved in the decarboxylation of aspartate in the bundle sheath of maize. Plant Physiol. 120(2):539–546.
Zhang G, et al. 2012. Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol. 30(6):549–554.
Zhang Y, et al. 2017. Differentially regulated orthologs in sorghum and the subgenomes of maize. Plant Cell. 29(8):1938–1951.

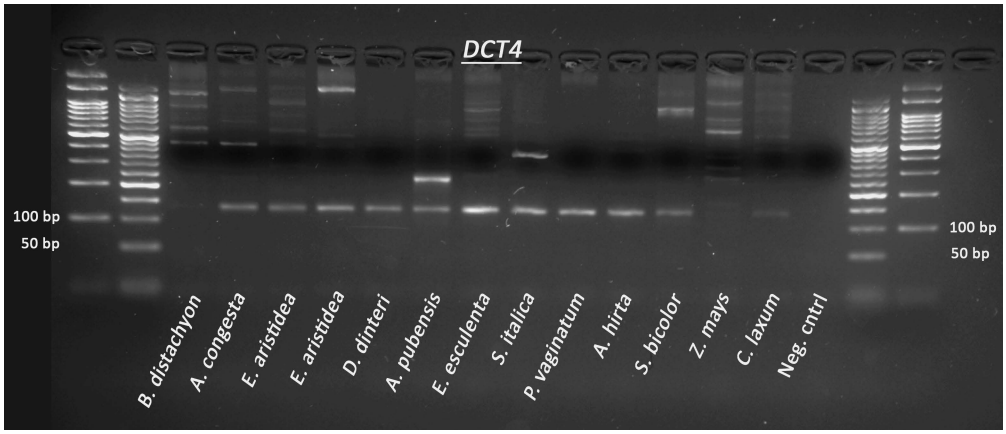
Associate editor: Adam Eyre-Walker

SUPPLEMENTAL FIGURES LEGEND

Supplemental Figure 1. Gel image showing that *DCT1* (92 bp PCR product) and *DCT2* (115 bp PCR product) genes are present in all grass species tested.

Supplemental Figure 2. Gel image showing the presence or absence of *DCT4* (132 bp PCR product) genes from species lacking genome assemblies. Negative controls were *Z. mays* and *B. distachyon*, and positive controls were *S. bicolor* and *S. italica*.


DCT1
100 bp
50 bp
100 bp
50 bp
B. distachyon
A. congesta
E. aristidea
E. aristidea
D. dinteri
A. pubensis
E. esculenta
S. italica
P. vaginatum
A. hirta
S. bicolor
Z. mays
C. laxum
Neg. cntrl
DCT2
100 bp
50 bp
B. distachyon
A. congesta
E. aristidea
E. aristidea
D. dinteri
A. pubensis
E. esculenta
S. italica
P. vaginatum
A. hirta
S. bicolor
Z. mays
C. laxum


DCT4
100 bp
50 bp
100 bp
50 bp
B. distachyon
A. congesta
E. aristidea
E. aristidea
D. dinteri
A. pubensis
E. esculenta
S. italica
P. vaginatum
A. hirta
S. bicolor
Z. mays
C. laxum
Neg. cntrl