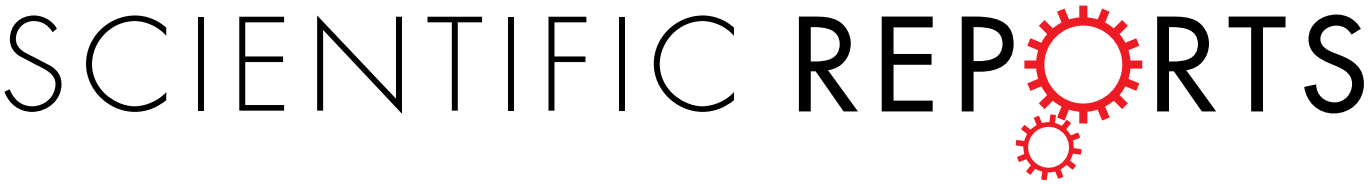

OPEN

New evidence for grain specific C_4 photosynthesis in wheat

Parimalan Rangan[1,2], Agnelo Furtado[1] & Robert J Henry[1]

Received: 22 February 2016
Accepted: 22 July 2016
Published: 17 August 2016

The C_4 photosynthetic pathway evolved to allow efficient CO_2 capture by plants where effective carbon supply may be limiting as in hot or dry environments, explaining the high growth rates of C_4 plants such as maize. Important crops such as wheat and rice are C_3 plants resulting in efforts to engineer them to use the C_4 pathway. Here we show the presence of a C_4 photosynthetic pathway in the developing wheat grain that is absent in the leaves. Genes specific for C_4 photosynthesis were identified in the wheat genome and found to be preferentially expressed in the photosynthetic pericarp tissue (cross- and tube-cell layers) of the wheat caryopsis. The chloroplasts exhibit dimorphism that corresponds to chloroplasts of mesophyll- and bundle sheath-cells in leaves of classical C_4 plants. Breeding to optimize the relative contributions of C_3 and C_4 photosynthesis may adapt wheat to climate change, contributing to wheat food security.

One of the key biological innovations was development of the ability of an organism to use light as the source of energy to generate chemical energy (ATP and NAD(P)H) for metabolic activities[1] in the process commonly known as photosynthesis[2]. Evolutionarily, six phyla of prokaryotic bacteria have the ability to photosynthesize[3], five of them using anoxygenic photosynthesis with bacteriochlorophyll and only one, the cyanobacteria, having oxygenic photosynthesis with chlorophyll[4]. Endosymbiotic associations of cyanobacteria in eukaryotes resulted in their ability to photosynthesize through chloroplasts in the process designated as "photosyntax" or "photosynthesis" in 1893 by Charles Reid Barnes[5]. Chemical energy generated from light energy is captured and used to synthesize organic compounds in higher plants in 'dark reactions'[6]. There are many different photosynthetic pathways reported in higher plants[7]; four types *viz*., C_3, C_4, CAM (Crassulacean acid metabolism), and C_3-C_4 intermediates are widely known, while, C_4-like (less advanced C_4), C_3-CAM, and C_4-CAM intermediates have also been reported. These photosynthetic pathways, able to use CO_2 as a carbon source, evolved in cyanobacteria around 3.5 billion years ago[8]. The key enzyme in C_3 photosynthesis, ribulose diphosphate carboxylase (RuBisCO), was reported to have evolved around the same time as cyanobacteria[9]. The C_4 pathway originated approximately 30 Mya (million years ago)[10] and was first described 50 years ago[11]. The pathway provides enhanced radiation- water- and nitrogen- use efficiency[12] especially in sub-optimal environments[10,13].

Three classical C_4 photosynthesis subtypes, NADP-ME (NADP- dependent malic enzyme), NAD-ME (NAD-dependent malic enzyme) and PEPCK (phosphoenolpyruvate carboxykinase) have been defined based upon the decarboxylation reactions involved[14]. These photosynthetic pathways explain the high growth rates of C_4 plants such as maize. Anatomical, biochemical, and molecular evidence has been commonly used to distinguish C_4-(sub)types from C_3-types[15]. Kranz anatomy with reactions compartmentalized in different cell types has been considered essential for C_4 photosynthesis[16] but spatial compartmentalization in a single-cell has been demonstrated more recently[17]. The stem and petiole of C_3 plants (tobacco and celery) was reported to accomplish NAD-ME type C_4 photosynthesis in cells surrounding vascular bundles[18]. Photosynthesis in cereal grains is less well defined. Ear photosynthesis in wheat contributes from 10% to 44% of grain yield[19]. Grain photosynthesis accounts for 33-42% of this photosynthesis depending on the genotype and environment[20].

Wheat is a major food crop critical to global food security. The current increase in wheat production of around 1% per year is not keeping pace with the rate of yield growth required to achieve the target of doubling crop production by 2050[21]. The likely impact of climate change makes progress in advancing wheat productivity more urgent. Increasing total plant biomass through efficient carbon capture by photosynthesis is now more crucial in improving wheat productivity since advances in grain yield by improving harvest index have plateaued[22]. Plants with the C_4 pathway are known to contribute 25% of total photosynthesis although they represent just 3% of species[10]. Converting C_3 crops to C_4 provides the possibility of improving yield by 30% through improved water- and

[1]Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane QLD 4072, Australia. [2]Division of Genomic Resources, ICAR-National Bureau of Plant Genetic Resources, New Delhi-110012, India. Correspondence and requests for materials should be addressed to R.J.H. (email: robert.henry@uq.edu.au)

Gene name	Copy number	C_4-type^	C_3-type^
ppc	2	3ABDL	5ABDL
aat	6	3ABDL (cyt) – *aat1*	1ABDL
		7ABDL (mt) – *aat2*	6ABDL
			6ABDS
			7ADS4AL
mdh	2	5ABDS (mt) – *mdh2*	1ABDL (cyt) – *mdh1*
me2	2	2ABDS (mt)	1ABDS (plastid)
gpt	2	2ABDS & 5ABDS (both cyt)	
ppdk	4	1ABDL (cp)	1ABDL (cyt)

Table 1. Wheat genes identified as being involved in NAD-ME type C_4 photosynthesis, chromosomal localization, and C_3-, C_4-type specificity. ^A, B and D represent the there sub-genomes in hexaploid wheat; 'L' and 'S' represent the long and short chromosomal arms; ***aat***: aspartate aminotransferase (also known as *got*); **cp**: chloropolastic; **cyt**: cytoplasmic; ***gpt***: alanine aminotransferase; ***mdh***: malate dehydrogenase; ***me2***: NAD-dependent malic enzyme; **mt**: mitochondrial; ***ppc***: phospho*enol*pyruvate carboxylase; ***ppdk***: pyruvate, orthophosphate dikinase..

nitrogen- use efficiency[23]. Engineering C_3 food crops like wheat and rice to use the C_4 pathway has long been explored to enhance global food security[24]. We now report an analysis of the transcriptome of genes associated with C_4 photosynthesis in the developing wheat grain. Genes identified as transcripts were located in the genome and their sequences analysed to determine likely specificity. This allowed an evaluation of substantial new evidence for C_4 photosynthesis in wheat grains.

Results

Remarkably, transcriptome analysis and functional annotation of genes expressed in developing wheat grains revealed the presence and expression of all genes specific to NAD-ME type C_4-photosynthesis. When added to earlier evidence dispersed in the literature, the present discoveries suggest the functioning of a form of C_4-photosynthesis specifically in the developing wheat grain. The transcriptome of the developing caryopsis from 35 diverse wheat genotypes (31 and 32 genotypes respectively from 14 and 30 days-post-anthesis stage with 28 genotypes in common) was analyzed by RNA-Seq. Annotation of the differentially expressed genes in the wheat grain transcriptome between 14 and 30 dpa (days-post-anthesis) indicated the presence of NAD-ME type C_4 photosynthesis during wheat grain development. This was an unexpected finding with wheat being a well-known C_3 crop. Wheat genes involved in C_4 photosynthesis, the number of copies expressed in developing wheat grains and their C_4 specificity (based on cytological and evolutionary evidence) are listed in Table 1.

Molecular evidence. Phospho*enol*pyruvate carboxylase (*ppc*) genes were localized in wheat on the long arms of chromosomes 3 and 5. The mean expression value (in RPKM) for *ppc* across 31 genotypes at 14 dpa (chromosome 3) was 36.2 (Fig. 1A, sum of three sub-genomes A, B, and D) while only 0.29 (mean of three growth stages – Z10, Z23, and Z71 with the expression values on the Y-axis representing the sum of the three sub-genomes) for leaves[25] (Fig. 2A), indicating a 125 fold up-regulation in the developing wheat caryopsis. Conversely, *ppc* from chromosome 5 was upregulated in leaves (Fig. 2A). It is well-known that C_4 plants have less RubisCO protein (reflecting transcript abundance) than C_3 plants[26]. The mean *rbcS* gene expression value was 512.3 and 39166 for the wheat caryopsis at 14 dpa and leaves respectively indicating a 76 fold down-regulation in the developing wheat caryopsis. This shows an enormous, 9500 fold, difference between developing wheat caryopsis and leaves for the relative expression of *ppc* and *rbcS* genes.

Aspartate aminotransferase (*aat*; also known as *got*) is the most up-regulated among six C_4 pathway genes in the developing wheat caryopsis. This is also the most up-regulated gene in the leaf tissues between C_3 and C_4 plants[26]. Of six copies (in each sub-genome) of the *aat* gene in wheat, only two copies were the C_4 type (cytoplasmic 3L – *aat1* and mitochondrial 7L – *aat2*). RNA-Seq analysis indicated that these genes were differentially up-regulated at 14 dpa in the developing caryopsis (Fig. 1B) when compared with leaves (Fig. 2B)[25].

Two copies of malate dehydrogenase (*mdh*) gene were localized on the long and short arm of chromosome 1 (cytoplasmic – *mdh1*) and chromosome 5 (mitochondrial – *mdh2*) respectively across the three sub-genomes. The gene copy from chromosome 1 was differently expressed (Figs 1C and 2C) compared to the one from chromosome 5 in both grain and leaf tissues[25]. The mitochondrial targeted *mdh2* gene from chromosome 5 is likely to be involved in C_4 photosynthesis.

Two copies of the NAD-dependent malic enzyme coding gene (*me2*) with one each targeted to chloroplast and mitochondria were localized on chromosomes 1 and 2 respectively. The mitochondrial targeted gene (chromosome 2) copy supports C_4 photosynthesis, converting malate into pyruvate with release of CO_2 for further fixation through the C_3 cycle[15]. The mitochondrial isoform was up-regulated in the developing wheat caryopsis (Fig. 1D) while, the plastidic isoform was up-regulated in leaves (Fig. 2D)[25].

Two copies of alanine transaminase (*gpt*) genes were localized to the short arm of chromosomes 2 and 5 of hexaploid wheat. This cytoplasmic enzyme converts pyruvate to alanine and vice-versa in bundle sheath and mesophyll cells respectively in a classical NAD-ME type C_4 pathway[14]. Both genes were expressed in similar

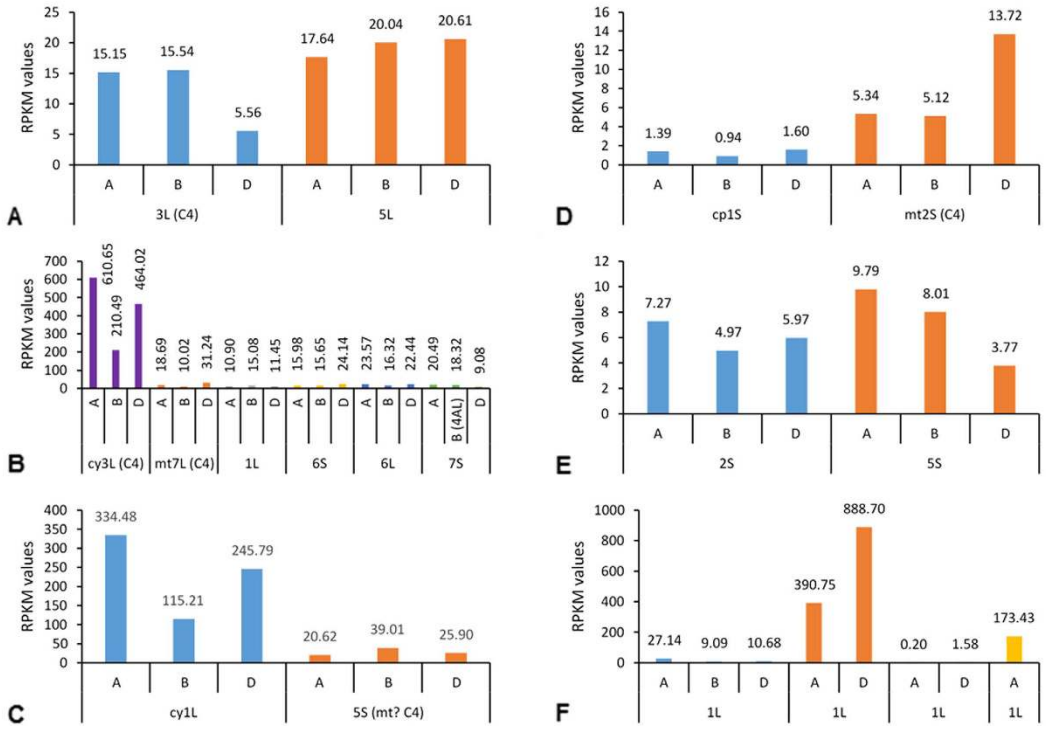

Figure 1. Grain specific expression of genes involved in NAD-ME type C_4 photosynthesis as RPKM values (mean of 31 genotypes) partitioned to the sub-genome level at 14 days-post-anthesis period. (**A**) *ppc* (phospho*enol*pyruvate carboxylase); (**B**) *aat* (aspartate aminotransferase); (**C**) *mdh* (malate dehydrogenase); (**D**) *me2* (NAD-dependent malic enzyme); (**E**) *gpt* (alanine aminotransferase); (**F**) *ppdk* (orthophosphate, pyruvate dikinase). **Abbreviations used in X-axis panel–ABD:** three sub-genomes of hexaploid wheat; **cp:** chloroplast targeted; **cy:** cytosolic; **L:** chromosomal long arm; **mt:** mitochondria targeted; **S:** chromosomal short arm.

proportions in the developing wheat caryopsis at 14 dpa (Figs 1E and 2E); while the gene on chromosome 2 was more highly expressed in leaves[25].

Pyruvate, orthophosphate dikinase (*ppdk*) gene was localized to the long arm of chromosome 1 in hexaploid wheat. All four gene copies (although a full length sequence was not available) were used to assess the RPKM expression levels in the developing wheat caryopsis at 14 dpa (Fig. 1F) and in leaf (Fig. 2F) tissues[25]. Earlier reports indicate the role of a dual promoter in regulating a single gene copy during light and dark in the chloroplast and cytoplasm respectively with the second promoter region in the first intron for cytoplasmic expression[27]. Aoyagi and co-workers showed the presence of PPDK and RubisCO in the green pericarp, but failed to envision the possibility of C_4 photosynthesis due to the lack of Kranz anatomy in developing wheat grains[28].

Six genes (excluding carbonic anhydrase) were involved in the NAD-ME type C_4 pathway, phospho*enol*pyruvate (PEP) carboxylase (*ppc*), aspartate aminotransferase (*aat*; also known as *got*), malate dehydrogenase (*mdh*), NAD- dependent malic enzyme (*me2*), alanine aminotransferase (*gpt*), and pyruvate, orthophosphate dikinase (*ppdk*)[15]. Grain specific expression of genes involving NAD-ME type C_4 photosynthesis *viz*., *ppc*, *aat*, *mdh*, *me2*, *gpt*, and *ppdk*; in all three (A, B, and D) sub-genomes (Fig. 1) indicates a possible evolutionary diversification point well before the speciation of the diploid progenitors in the Triticeae tribe. Endosperm and aleurone transcripts[29] do not express all of these genes demonstrating that the C_4 pathway is restricted to the wheat pericarp.

Varied expression pattern between wheat genotypes. The presence of all C_4 specific genes in the genome confirms that natural selection may have already explored the options being considered by plant breeders[30]. The levels of expression for all six genes at 14 dpa in NAD-ME type C_4 pathway varied across 31 genotypes (Fig. 3) suggesting potential for genetic selection for this trait in wheat breeding.

C_4 specificity of gene sequences. Four of the six genes involved in NAD-ME type C_4 photosynthesis, (*aat*, *mdh*, *me2*, and *ppdk*) had sub-cellular targeting that suggests C_4-type specificity[15]. The other two genes (*ppc* and *gpt*) require sequence information to distinguish between the copies specific for C_3- or C_4- pathways. Analysis of *gpt* genes in wheat suggested both C_3 and C_4 forms were expressed at similar levels (Figs 1E and 2E) across photosynthetic and non-photosynthetic tissues. While the *ppc* gene copies clearly show different expression patterns between developing grains and leaves (Figs 1A and 2A); sequence differences are the only way to distinguish the C_3- and C_4- isoforms. Specific amino acid substitutions have been associated with C_4 functionality[13]. Increased tolerance to feedback inhibition by malate involves $\mathbf{G_{884}}$ (Glycine) in C_4-isoforms rather than $\mathbf{R_{884}}$ (Arginine) as found in C_3-isoforms. The translated sequence of the *ppc* gene from chromosome 3 ($\mathbf{S_{885}}$) and 5 ($\mathbf{R_{891}}$) of wheat

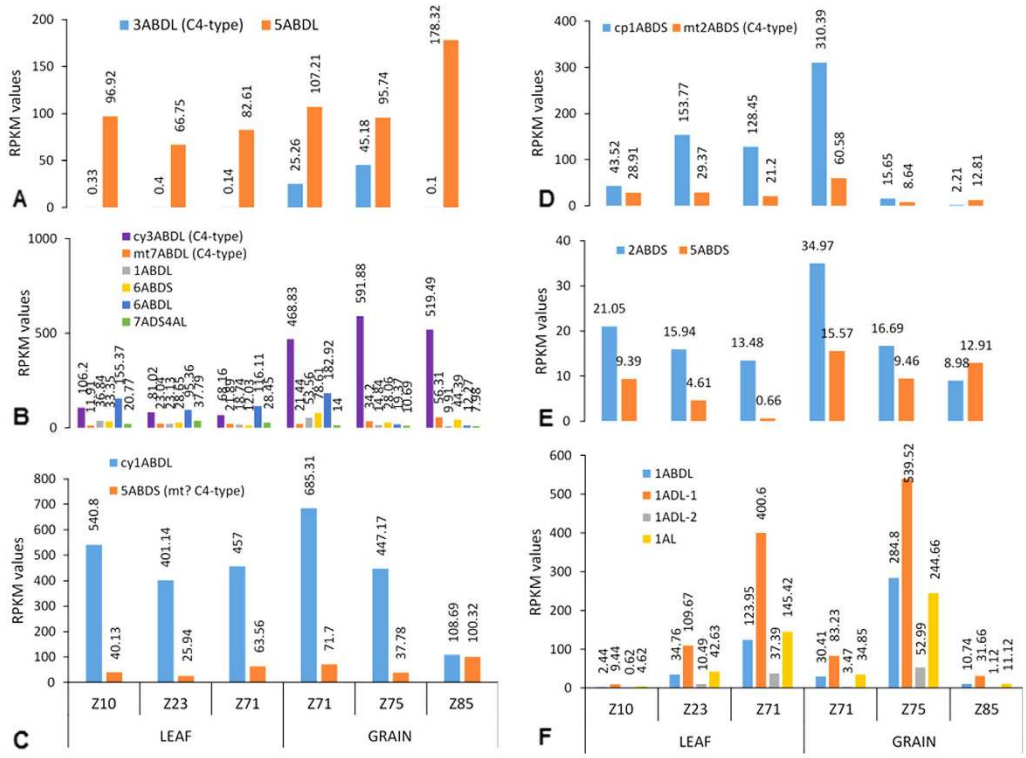

Figure 2. Comparison of expression levels (measured as RPKM) for genes involved in NAD-ME type C_4 photosynthesis in leaf and grain tissues of hexaploid wheat. for X-axis, refer bottom most panel. (**A**) *ppc* (phospho*enol*pyruvate carboxylase); (**B**) *aat* (aspartate aminotransferase); (**C**) *mdh* (malate dehydrogenase); (**D**) *me2* (NAD-dependent malic enzyme); (**E**) *gpt* (alanine aminotransferase); (**F**) *ppdk* (orthophosphate, pyruvate dikinase). **Abbreviations used in the X-axis panel and series' legends – ABD:** three sub-genomes of hexaploid wheat; **cp:** chloroplast targeted; **cy:** cytosolic; **L:** chromosomal long arm; **mt:** mitochondria targeted; **S:** chromosomal short arm; **Zadok's scales Z10:** seedling stage; **Z23:** tillering stage; **Z71:** watery kernel stage; **Z75:** early-grain filling stage (app.14 days-post-anthesis); **Z85:** late-grain filling stage (app. 30 days-post-anthesis).

cDNA (IWGSC – international wheat genome sequencing consortium, release-23 version) indicates the gene copy from chromosome 5 is C_3-type; while the chromosome 3 copy is non-C_3 type. The gene sequences from wheat and related species[31] were analyzed using the translated amino acid sequence of the *ppc* gene (IWGSC cDNA database release-23) from chromosome 3. Results indicated that most of the Triticeae tribe members have five copies of *ppc* gene (Table 2) although in the hexaploid wheat cDNA database we found only two copies (3L and 5L). IWGSC cDNA database (release-23)[31] was used to perform tblastn analysis with the translated *ppc* gene sequence confirming that gene sequence copies from chromosomes 3S, 6 and 7 are not in frame suggesting the presence of insertions or deletions in these genes. However, one *ppc* gene copy from all Triticeae members had $\mathbf{S_{885}}$ indicating a non C_3-type; while the other four copies revealed a C_3-type – $\mathbf{R_{891}}$ (the corresponding amino acid position) across all the Triticeae members studied (Table 2). Since the amino acid position is neither **R** nor **G**, we studied different species acting as diversification points in the evolution of these species in order to compare them with respect to known C_4 types (Table 3). This gave an indication that from Bryophytes to Angiosperms, the C_3 type amino acid position was invariably conserved with **'R'** (Table 3). Whereas the C_4 type amino acid position was either **S** (*Panicum* and Triticeae tribe) or **Q** (*Alloteropsis*, *Setaria*) or **G** (*Alloteropsis*, *Panicum*, *Zea*, and *Sorghum*) or **I** (*Amaranthus*) depending on the species or taxonomic group (Table 3).

Discussion

Wheat is widely known as a classical C_3 plant. Close examination of the literature shows many reports of components of the case for C_4 photosynthesis in the grain especially in early studies. However, this evidence has been overlooked because of the knowledge of C_3 photosynthesis in the leaves and a lack of understanding of the possibility of different pathways in different parts of the plant. Indeed many studies have attempted to explain away the evidence that did not fit with the knowledge that wheat was a C_3 plant. This study has identified a complete set of C_4 specific genes in wheat genome for the first time. This finding addresses the apparent anomaly of this subfamily (Pooideae) of the Poaceae being uniquely seen to lack C_4 photosynthesis. We have also shown for the first time that all the required genes are expressed in the required compartmentalization, specifically in the pericarp, a tissue with an anatomy that is suitable for supporting a C_4 pathway. The possibility of photosynthesis in the pericarp of wheat grains was predicted in the early 1960s[32]. Phospho*enol*pyruvate carboxylase (PPC) from the wheat or barley pericarp tissues of developing grain was reported to be 50-100 times as active in carbon fixation as ribulose

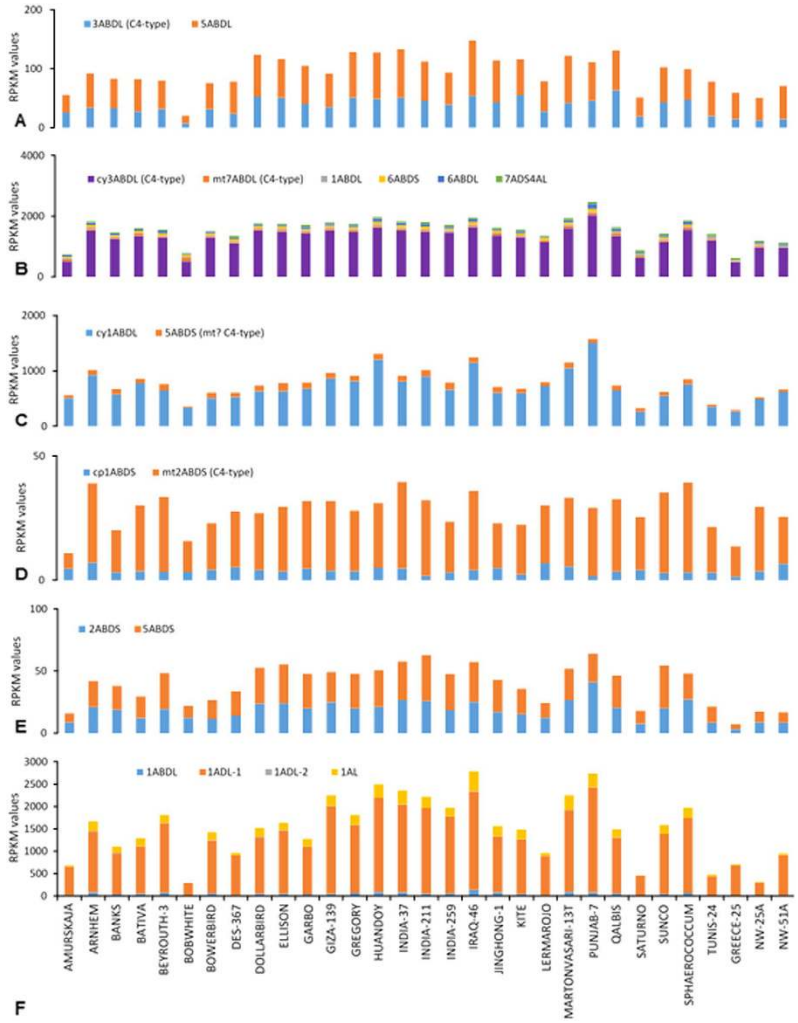

Figure 3. Variation in expression levels (measured as RPKM) for genes involved in NAD-ME type C_4 photosynthesis in 31 wheat genotypes at 14 days-post-anthesis. For X-axis, refer 'panel F'. (**A**) *ppc* (phospho*enol*pyruvate carboxylase); (**B**) *aat* (aspartate aminotransferase); (**C**) *mdh* (malate dehydrogenase); (**D**) *me2* (NAD-dependent malic enzyme); (**E**) *gpt* (alanine aminotransferase); (**F**) *ppdk* (orthophosphate, pyruvate dikinase). **Abbreviations used in series' legends – ABD:** three sub-genomes of hexaploid wheat; **cp:** chloroplast targeted; **cy:** cytosolic; **L:** chromosomal long arm; **mt:** mitochondria targeted; **S:** chromosomal short arm.

diphosphate carboxylase (RuBisCO)[33]. Based on the enzyme activity for malate dehydrogenase, malic enzyme, and pyruvate-orthophosphate dikinase in pericarp tissues of developing grain, Duffus and Rosie[33] indicated the possibility of C_4 photosynthesis. A little later, Wirth, *et al.*[34] studied different reproductive parts from wheat and oat – glume, lemma, palea, and pericarp – along with leaves and reported that the pericarp tissues of developing grains seemed to "possess carbon metabolism different to that of the other tissues". They also analyzed and reported the possibility of refixation of the CO_2 released through respiration or photorespiration. Assimilation of $^{14}CO_2$ to malate and 3-phosphoglyceric acid in wheat ears and flag leaf respectively; along with higher enzyme activities for enzymes of C_4 and C_3 metabolic pathways in ears and flag leaf respectively suggested the possibility of C_4 photosynthesis in ears[35]. Carbon isotope discrimination (Δ) values were used to distinguish plants between C_3- and C_4-type[36]. Although wheat was considered a C_3 plant, Δ values were used to study the plants' water-use- or transpiration efficiency[37,38]. Their results indicate a clear difference between flag leaf and grain Δ values in different wheat genotypes. Although the difference is not as distinct as it is with classical C_4 photosynthesis. This might be due to either inefficient less advanced C_4 type photosynthesis or the fact that grain photosynthesis accounts for only 33-42% of ear photosynthesis[20] with the remainder translocated from leaf or stem tissues with C_3-type photosynthesis thereby diminishing the difference in Δ values between flag leaf and grain to a marginal level. Similarly but in reverse, in a maize plant with C_4-type, maize husk leaves were reported to be C_3-type and their Δ values were marginally higher than leaves[39].

In spite of this evidence (enzyme activity, $^{14}CO_2$ in malate, Δ values), earlier researchers failed to explore C_4 photosynthesis in wheat grains due to the view that Kranz anatomy was required for C_4 photosynthesis[16,28]. In

Species[a]	Amino acid with flanking region	Chr no	Type	Reference
Brachypodium distachyon	LEGDPYLKQRLRLRDPY	1	C_3	Kersey, *et al.*[31]
	LESDPYLRQRLMLRDSY	2	C_3	
	LEGDPYLRQRLRLRE??	2	C_3	
	LEGDPYLKQRLRLRESY	3	C_3	
	LEGDPYLKQRLRLRDAY	4	C_3	
Hordeum vulgare	LESDPYLRQRLMLRDSY	3	C_3	Kersey, *et al.*[31]
	LEGDPYLRQ**S**LRLRDSY	3	**C_4?**	
	LEGDPYLKQRLRLRDAY	5	C_3	
	LEGDPYLKQRLRLRESY	6	C_3	
	LEDDPYLKQRLRLRDPY	7	C_3	
A. sha, spe, tau; T. mon, ura; T. durum (cappelli)	LESDPYLRQRLLLRDSY	3?	C_3	IWGSC[63]
	LEGDPYLRQ**S**LRLRDSY	3?	**C_4?**	
	LEGDPYLKQRLRLRDAY	5?	C_3	
	LEGDPYLKQRLRLRESY	6?	C_3	
	LEDDPYLKQRLRLRDPY	7?	C_3	
T. durum (strongfield)	LESDPYLRQRLLLRDSY	3?	C_3	IWGSC[63]
	LEGDPYLRQ**S**LRLRDSY	3?	**C_4?**	
	LEGDPYLKQRLRLRESY	6?	C_3	
	LEDDPYLKQRLRLRDPY	7?	C_3	
T. aestivum	LESDPYLRQRLLLRDSY	3S	C_3	Mayer, *et al.*[65]
	LEGDPYLRQ**S**LRLRDSY	3L	**C_4?**	
	LEGDPYLKQRLRLRDAY	5L	C_3	
	LEGDPYLKQRLRLRESY	6ASBDL	C_3	
	LEDDPYLK?????????	7DL		
	LEGDPYLRQRLQLRDPY	4DS	C_3	

Table 2. Amino acid (10th position, bold) for C_4 specificity across species that are taxonomically related to *Triticum* spp. [a]**A. sha:** Aegilops sharonensis; **spe:** A. speltoides; **tau:** A. tauschii; **T. mon:** Triticum monococcum; **ura:** T. uratu;

2001 and 2002, the occurrence of C_4 photosynthesis without Kranz anatomy was reported in single cells and in the petioles of C_3 plants respectively[17,18].

In the late 1990s, there were reports of the C_4 pathway being found selectively at different developmental stages of some plants (*Salsola* spp. and *Haloxylon* spp.) of cotyledons and leaves exhibiting C_3 and C_4 type photosynthesis respectively[40,41]. Similarly there have been reports on the selective use of the C_4 pathway in different environments like terrestrial or submerged situations[42,43] or high or low CO_2[44]. Selective expression of the C_3 pathway was reported in the husk leaves (hypsophylls) of the maize plant[45] which is otherwise a C_4 plant. Evidence of 4-carbon compounds specifically in wheat leaf bases[46] agreed with a much later report of C_4 pathways in C_3 plants[18]. Altered C_4 and C_3 enzymatic activity has recently been reported in wheat ears under water stress[47] but the significance of this was not clear given the C_3 status of wheat. This evidence suggests the presence of a diversified range of regulatory patterns of C_4 pathway in plants with ontogeny and varied environmental cues. Operation of different pathways (C_3 or C_4) at different growth stages allows wheat to have a lifecycle that extends across seasons with varying environments (cool, wet during vegetative growth; hot, dry during grain filling).

Molecular and cytological evidence. Functional annotation and differential expression of C_4-specific gene copies (Figs 1 and 2) for genes of NAD-ME type photosynthesis specifically in developing wheat grains adds evidence for the C_4 pathway in wheat grains as suggested in early reports[34,35]. With multiple copies in a genome, species preferentially co-opt the same neo-functionalized gene lineage for C_4 photosynthesis[48] although these genes were present well before the evolution of the C_4 pathway but with different anaplerotic functions[49].

Reports indicate that cross- and tube-cells in pericarp of developing wheat grain are photosynthetic in nature and contribute to the grain weight[50]. Thorough re-examination of this report indicates the presence of numerous mitochondria, and dimorphic chloroplasts – stacked grana in cross-cells and reduced stacking in tube-cells being structurally similar to classical C_4 types[51]. The presence of numerous mitochondria specifically in bundle sheath cells of NAD-ME type C_4 pathway has also been reported[52]. These pieces of cytological evidence in addition to our molecular evidence suggests NAD-ME type C_4 photosynthesis operates in developing wheat grains (Fig. 4) with cross- and tube-cells paralleling mesophyll and bundle sheath cells in a classical C_4 pathway. In contrast to the classical C_4 photosynthesis that is associated with little or no starch granules in mesophyll cells[53], the presence of starch granules in cross-cells (mesophyll like) was reported by Morrison[50]. This led us to question the possibility of NAD-ME type C_4 photosynthesis in wheat grains. However, there is evidence for the presence of RuBisCO in both mesophyll and bundle sheath cells of young amaranth leaves[54] suggesting a C_3 cycle in both mesophyll and bundle sheath cells. In some *Flaveria* spp., reports of the presence of RuBisCO in both mesophyll and bundle sheath cells supporting both the C_3 and C_4 cycle simultaneously led to their classification as having C_4-like type (less advanced)

Species[a]	Amino acid with flanking region	Type	Taxonomy	Mya	Reference
Physcomitrella patens	LQGNPTLKQRLRLREPY	C_3	Bryophytes	450	Rensing, *et al.*[66]
	LQGNPSLKQRLRLREPY	C_3			
	LQGNPTLKQRLRLREPV	C_3			
	LEGNPTLKQRLRLREQY	C_3			
Selaginella moellendorffii	LAGNPILKQRLSLREPF	C_3	Lycophytes	410	Banks, *et al.*[67]
	LEENPTLKQRLRLREPF	C_3			
PA, AS, GG, JC, PS, TB	LEGDPYLKQRLRLRDSY	C_3	Gymnosperm	300	Nystedt, *et al.*[64]
Amborella trichopoda	LEGDPYLKQRLRLRDSY	C_3	Basal Angiosperm	130	Soltis, *et al.*[68]
Oryza sativa	LEGDLYLKQRLRLRNAY	C_3	Oryzoideae (BOP clade)[p]	60[q]	Kersey, *et al.*[31]
	LEGDPYLRQRLRIRDSY	C_3			
	LEGDPYLKQRLRLRDAY	C_3			
	LEGDPYLKQRLRLRESY	C_3			
Guadua sp.	LEGDPYLKQRLRLRESY	C_3	Bambusoideae (BOP clade)[p]		Christin, *et al.*[69]
	LESDPYLRQRLMLRDSY	C_3			
Brachypodium	Refer Table 2	C_3	Pooideae (BOP clade)[p]	35[q]	
Triticeae	Refer Table 2	?		11.6[q]	
Alloteropsis spp.	LEGDPYLKE**G**LRLRNPY	C_4[b]	Panicoideae (PACMAD clade)[p]	20-27[q]	Christin, *et al.*[69]
	LEGDPYLKQ**Q**LRLRDPY	C_4[c]			
	LEGSPGLKQRLRLRDPY	C_3[d]			
	LEGDPYLKQRLRLRESY	C_3[e]			
	LEGDPYLKQRLRIRDSY	C_3[f]			
	LEGDPYLKQRLRLRDAY	C_3[g]			
	LEGGPYLKQRLRLRDPY	C_3[h]			
	LEGDLYLKQRLRLRDAY	C_3[i]			
	LEGDPYLKQRLRLRDPY	C_3[j]			
Panicum spp.	LEGDPFLKQ**S**LRLRNPY	C_4[k]	Panicoideae (PACMAD clade)[p]	18-27[q]	Christin, *et al.*[70] Christin, *et al.*[69]
	LEGDPYLKQ**G**LRLRNPY	C_4[l]			
	LEADPFLKQ**S**LRLRNPY	C_4[m]			
	LEGDLYLKQRLRLRDAY	C_3[n]			
	LEGDPYLKQRLRLRDPY	C_3[o]			
Setaria italica	LESDPGLKQ**Q**LRLRDPY	C_4	Panicoideae (PACMAD)[p]	16-27[q]	Bennetzen, *et al.*[71] Christin, *et al.*[69]
	LEGDPYLKQRLRLRESY	C_3			
	LESDPGLQQ**Q**LMLRDSY	C_3			
	LEGDLYLKQRLRLRDAY	C_3			
	LEGDPYLKQRLRIRDSY	C_3			
Zea mays	LEGDPFLKQ**G**LVLRNPY	C_4	Panicoideae (PACMAD)[p]	16[q]	Schnable, *et al.*[72]
	LEGDLYLKQRLRLRDAY	C_3			
	LEGDPYLKQRLRIRDSY	C_3			
	LEGDPYLKQRLRLRESY	C_3			
Sorghum bicolor	LEGDPYLKQ**G**LRLRNPY	C_4	Panicoideae (PACMAD)[p]	13[q]	Paterson, *et al.*[73]
	LEGDLYLKQRLRLRDAY	C_3			
	LEGDPYLKQRLRIRDSY	C_3			
	LEGDPYLKQRLRLRESY	C_3			
	LEGDPYLKQRLRLRDAY	C_3			
Amaranthus hypochondriacus	LDADPYLKQ**I**LRLRDPY	C_4	Dicot	-	ADO15315 (Accession number)

Table 3. Amino acid (10th position, bold) for C_4 specificity across species with different diversification point in evolutionary scale (in Mya). [a]Species arranged in the order of evolution except Amaranthus**; PA:** Picea abies; **AS:** Abies sibiria; **GG:** Gnetum gnemon; **JC:** Juniperus communis; **PS:** Pinus sylvestris; **TB:** Taxus baccata; **Oryza sativa**: includes indica and japonica; **Alloteropsis spp.**: A. cimicina[d-g], A. angusta[c,f,g,i], A. semialata subsp semialata[b,f-i], A. s. subsp eckloniana[g,i,j]; **Panicum spp.**: P. bisulcatum[n], P. capillare[l], P. coloratum[l], P. fluviicola[l], P. laetum[k,m], P. miliaceum[k,l], P. millegrana[o], P. phragmitoides[l], P. turgidum[l]. [p]Soreng, *et al.*[56]. [q]Chalupska, *et al.*[57].

photosynthesis[7]. Occurrence of the C_3 cycle (presence of RuBisCO) in both mesophyll and bundle sheath cells along with the C_4 pathway in some species might be due to the fact that compartmentalization of RuBisCO is the final step in the evolution of C_4 from C_3 photosynthesis[55]. These considerations lead us to propose the occurrence of C_4-like

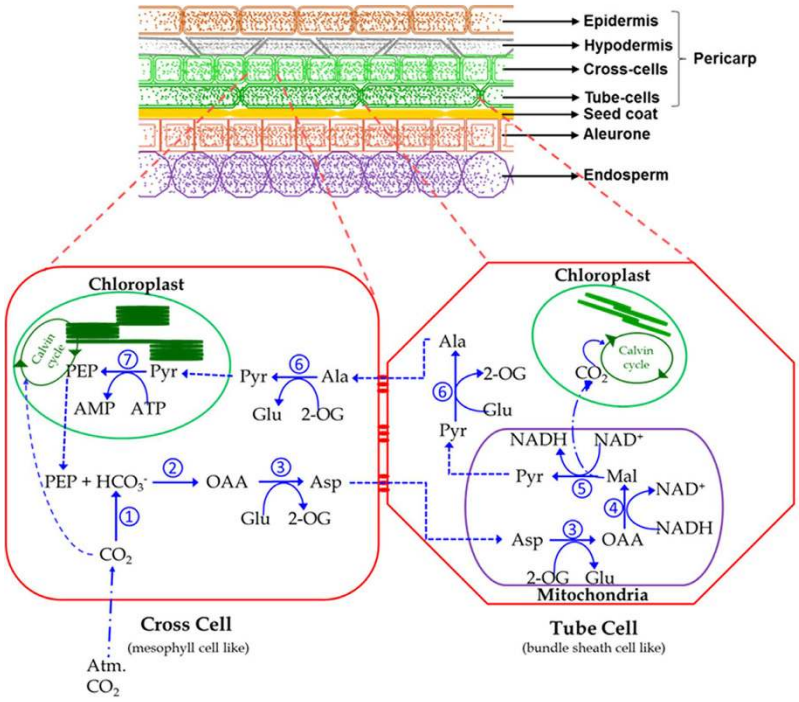

Figure 4. Schema depicting a longitudinal cross section of developing wheat grains with one cross- and tube-cell in enlarged view illustrating the C_4-like type photosynthetic pathway through NAD-dependent malic enzyme (NAD-ME) subtype [adapted and modified from Sage and Monson[15]]. 2-OG: 2-Oxyglutarate; **Ala:** Alanine; **Asp:** aspartate; **Glu:** glutamate; **Mal:** malate; **OAA:** oxaloacetate; **PEP:** phospho*enol*pyruvate; **Pyr:** pyruvate; **1,** Carbonic anhydrase (CA, EC 4.2.1.1); **2,** PEP carboxylase (PPC, EC 4.1.1.31); **3,** Aspartate aminotransferase (AAT, EC 2.6.1.1); **4,** Malate dehydrogenase (MDH, EC 1.1.1.31); **5,** NAD-dependent-malic enzyme (ME2, 1.1.1.39); **6,** Alanine aminotransferase (GPT, EC 2.6.1.2); **7,** Pyruvate orthophosphate dikinase (PPDK, EC 2.7.9.1).

type (less advanced) photosynthesis in developing wheat grains through the cross- and tube-cell layers of pericarp paralleling the mesophyll and bundle sheath cells of classical C_4 photosynthesis (Fig. 4).

Taxonomical and evolutionary evidence. Around 41% of grasses are known to fix carbon through the C_4 pathway[56]. Hence, an overview at evolutionary scale linking speciation events with C_4 photosynthesis might shed light on the evolution of the C_4-like type photosynthetic pathway in developing wheat grains. The Poaceae family is monophyletic and consists of 12 subfamilies with three at the basal level, followed by the BOP and PACMAD clades consisting of three and six subfamilies each with the Triticeae tribe included in the subfamily Pooideae (cool season grasses) of the BOP clade[56]. To date, no species from the Pooideae have been reported to be C_4. The Aristidoideae (PACMAD) subfamily has been reported to have at least two independent evolutions of the C_4 pathway[56]. In this study, knowledge of specific amino acids (**G** in C_4 and **R** in C_3) in the *ppc* gene product required for efficient carbon fixation by the C_4 pathway[13] was used to show that wheat and all related species (including *Hordeum* and *Brachypodium*) had five *ppc* copies in their genome with four of them having the amino acid **R** indicating their C_3 nature while one copy (3L, as in hexaploid wheat) has an **S** – a non-C_3 type in place of **R** except *Brachypodium* (Table 2).

The C_3 specific amino acid position (**R**) was apparently conserved (Table 3) from Bryophytes (around 450Mya) to Angiosperms. The amino acid position with C_4 specificity appears to have evolved at least four times in the last 30Mya (origin of C_4) with either **S** (*Panicum* and Triticeae tribe) or **Q** (*Alloteropsis*, *Setaria*) or **G** (*Alloteropsis*, *Panicum*, *Zea*, and *Sorghum*) or **I** (*Amaranthus*). This suggest that various amino acid substitutions at that site might result in differing efficiency of carbon fixation through the C_4 pathway by altering tolerance to feedback inhibition by malate[13]. Analysis of enzyme kinetics with each of the four C_4-specific amino acids individually might help to rank their photosynthetic efficiency. However, the weakest form among the four will probably be much more efficient in carbon fixation than the C_3 type (**R**). The presence in wheat of an amino acid specific to a known C_4-type (**S** in *Panicum laetum* and *P. miliaceum*) is strong evidence when taken together with the grain specific pattern of expression of the C_4 specific *ppc* gene (Fig. 2A). The tribe Brachypodieae (*Brachypodium distachyon*) has amino acids corresponding to the C_3-type for all the five copies; while members from tribe Triticeae (*Aegilops*, *Hordeum*, *Triticum*) have one copy of the C_4-type and four copies of the C_3-type (Table 2). This fits with the evolutionary time line for C_4 photosynthesis around 30Mya[10]; with Brachypodieae evolving around 35Mya[57] having only C_3-type genes (*Brachypodium*). Unfortunately, there are no diversification points between *Brachypodium* (35 Mya) and *Hordeum* (11.6 Mya)[57] to establish the exact timing of C_4 evolution in the Pooideae tribe. Derived traits like those associated with C_4 photosynthesis appear

at later evolutionary stages and are expressed later in plant development[58]. This is consistent with the observations of C_4 photosynthesis in wheat and its relatives specifically in the grain.

Based on this molecular, cytological, taxonomical and evolutionary evidence, we propose the occurrence of C_4-like type photosynthesis specifically in developing wheat grains. Recognition of both C_3 and C_4-like type photosynthetic pathways in wheat provides a basis for interpretation of wheat performance as a crop adapted to maturation in hot dry environments, suggesting that the plant may rely more on C_4 photosynthesis under conditions of water stress during the grain filling stage. Photosynthates from pericarp, glumes and awns are critical[59] when other parts of the plant lose photosynthetic capacity due to terminal drought often experienced in the environments in which wheat evolved. This may be especially important in the development of wheat varieties to adapt to climate change[60] and associated temperature extremes. The operation of C_4 photosynthesis specifically in these tissues provides an adaptive advantage to the wheat plant while C_3 photosynthesis is adequate during early vegetative growth under more temperate conditions. The potential for genetic manipulation to extend C_4 photosynthesis throughout the wheat plant seems much more realistic given the existing expression of the entire pathway in the grain. This supports the view that plant species have evolved specific photosynthetic pathways in different organs, at specific developmental stages and in different environments suggesting that the classification of plants as C_3 or C_4 or CAM in a broad fashion cannot simply be based upon leaf anatomy. Research to establish the variation in flux through this pathway in wheat and its progenitors will shed much light on the share of carbon fixation through the C_3 and C_4 pathway under varying environmental conditions. This has the potential to suggest new options for the development of higher yielding wheat genotypes.

Experimental procedures. *Experimental material.* Thirty-five wheat genotypes *viz*., Amurskaja, Arnhem, Banks, Bativa, Beyrouth-3, Bobwihte-26, Bowerbird, Des-367, Dollarbird, Ellison, Garbo, Giza-139, Gregory, Huandoy, India-37, India-211, India-259, Iraq-46, JingHong-1, Kite, LermaRojo, Martonvasari-13T, Punjab-7, Qalbis, Saturno, Sunco, Sphaerococcum, Tunis-24, Greece-25, NW-25A, NW-51A, NW-93A, NW-108A, Pelada, and Vega were used for transcriptome analysis. Seeds for these genotypes were procured from the Australian Winter Cereals Collection. Seeds were germinated and plants grown under controlled conditions as described in Furtado, *et al.*[58]. Developing grains were collected from wheat spikes at 14 days- and at 30 days-postanthesis (dpa) as described elsewhere[58].

RNA isolation, library preparation and NGS sequencing. RNA isolation, cDNA synthesis, library preparation and next generation sequencing was carried out and described by Furtado, *et al.*[58]. Libraries for 31 samples from 14 dpa and 32 samples from 30 dpa with 28 genotypes in common were prepared and sequenced as described in Furtado, *et al.*[58]. Libraries were not prepared for four cultivars *viz*., NW-93A, NW-108A, Pelada, and Vega at 14 dpa, and three cultivars *viz*., Greece-25, NW-25A, and NW-51A at the 30 dpa stage due to lack of sufficient starting material.

Sequencing data processing and analysis. Sequencing data obtained was imported into CLC genomics workbench ver. 7.0.4 (CLC Bio-Qiagen, Denmark) and further processing and analysis were done within this environment unless otherwise stated. Quality checking, trimming, and RNA-Seq analysis were performed as described in Furtado, *et al.*[58] using the TaGI (*Triticum aestivum* Gene Indices, The Computational Biology and Functional Genomics Laboratory, Dana Farber Cancer Institute and Harvard School of Public Health) cDNA database as reference, containing 221,925 sequences (release 12.0)[61].

Differential transcript and statistical analyses. Transcripts that were differentially expressed between 14 and 30 dpa were analyzed using the RNA-Seq experimentation tool with default parameters. Statistically significantly differentially expressed transcripts were identified using both Gaussian (mean based) and Empirical analysis of Differential Gene Expression (EDGE, count based) statistics facilitated through CLC workbench (CLC Bio-Qiagen, Denmark) with *p*-value using false discovery rate (FDR) corrected least the significant difference set at 0.01 level.

Functional annotation and data mining. In total, 26,477 transcripts that are common for both Gaussian and EDGE statistics were significant at FDR corrected value 0.01. Among them, 319 and 181 transcripts were unique to 14 dpa and 30 dpa respectively; while 16237 and 9740 transcripts were differentially up-regulated at 14 dpa and 30 dpa respectively. Transcript sequences for these four groups (unique 14 dpa, unique 30 dpa, differential 14 dpa and differential 30 dpa) were extracted from the reference database (TaGI) and subjected to blastx analysis against the non-redundant protein database. Blast results obtained were converted to a BLAST2GO project file and exported in ".dat" format files using the plug-in version within CLC workbench (CLC Bio-Qiagen, Denmark). Functional annotation for these four groups was performed independently using BLAST2GO Pro ver 3.0.10 with default parameters[62]. Annotations were augmented using InterProScan and followed by Run-annex options. Annotations pertaining to the plant database were retained using the GO (gene ontology)-slim option. Finally, KEGG (Kyoto encyclopedia of genes and genomes) pathway maps for these four annotated sequence groups were retrieved using GO-enzyme code mapping option. The differential 14 dpa group highlighted the presence of a complete C_4 photosynthetic pathway existing in developing wheat caryopsis.

Chromosomal localization and IWGSC transcript retrieval. Based on enzyme code mapping, TaGI transcript IDs pertaining to those enzyme code (EC numbers) for the genes involved in the C_4 photosynthetic pathway from the differential 14 dpa group were retrieved using CLC workbench (CLC Bio-Qiagen, Denmark). A total of 62 transcripts for the six genes (phosphoenolpyruvate carboxylase – *ppc*; aspartate aminotransferase – *aat*; malate dehydrogenase – *mdh*; decarboxylating dehydrogenase – *me2*; alanine aminotransferase – *gpt*; and pyruvate,

orthophosphate dikinase - *ppdk*) were retrieved. Blast searches for the 62 transcripts from the TaGI database were performed against the IWGSC cDNA database containing 100,717 sequences (release-23)[31] for retrieval of IWGSC transcripts (since the TaGI transcripts are lesser in length and mostly incomplete).

Modified reference and RNA-Seq analysis. Based on blast analyses using the 62 transcripts of TaGI as reference[55] transcripts from the IWGSC cDNA database (release-23)[31] were obtained. Using sub-genome sequence information and sequence alignment, 10 of 55 transcripts were found to be actually five genes each being two parts of the same transcript with or without overlap. Based on homology and sequence alignment between the sub-genome copies, those 10 transcripts were stitched into five transcripts resulting into a total of 50 transcripts for six genes that accomplish NAD-ME type C_4 photosynthesis. In order to construct a modified reference, 10 transcripts (that are used to stitch) were replaced with the five stitched transcripts in the 100,717 sequences of the IWGSC cDNA (release-23)[31]. The resulting database containing 100,712 sequences was named "modified IWGSC cDNA (release-23)"and used for performing RNA-Seq analysis as described above[58] to obtain RPKM values for the 31 genotypes at the 14 dpa stage. Although researchers use FPKM instead of RPKM for paired-end reads, we used RPKM with an option of counting mapped paired-end reads as "two" and singleton reads that are mapped as "one" to avoid confusion between FPKM and RPKM terminologies.

RNA-Seq analysis for tissue specific transcriptome data. Raw reads (100 bp paired-end sequencing on Illumina HiSeq2000) of different tissues (leaf, and grain) at three different growth stages for hexaploid wheat ('Chinese Spring') were available online[25]. These raw sequence reads were downloaded, and processed through the CLC workbench (CLC Bio-Qiagen, Denmark). Quality checking, trimming and RNA-Seq analysis using the modified IWGSC cDNA (release-23) containing 100,712 sequences as reference were performed to obtain RPKM values and represented in pictorial form.

Taxonomical and evolutionary relation for C_4-specificity. Specific amino-acid positions for PPC (PEPCase) that are functionally related to C_3 and C_4-specificity were reported recently[13]. In order to identify these in wheat and related species (Table 2), whole genome sequence details[63] were downloaded and translational blast analysis was performed using CLC workbench ver. 8.5.1 (CLC Bio-Qiagen, Denmark).

Similar analyses were performed for species (for which genome sequence was available) including taxa from bryophytes to angiosperms[31,64,65] corresponding to various diversification points in an evolutionary timeline (Table 3). Although whole genome data for some well-known C_4 species was not available, *ppc* gene sequences in public databases was used to study the evolutionary pattern at specific amino acid positions (Table 3) that are functionally related to C_3 or C_4 specificity.

References

1. Shih, P. M. Photosynthesis and early Earth. *Current Biology* **25,** R855–R859 (2015).
2. Gest, H. History of the word photosynthesis and evolution of its definition. *Photosynthesis research* **73,** 7–10 (2002).
3. Blankenship, R. E. Early evolution of photosynthesis. *Plant physiology* **154,** 434–438 (2010).
4. Lockhart, P. J., Larkum, A., Steel, M., Waddell, P. J. & Penny, D. Evolution of chlorophyll and bacteriochlorophyll: the problem of invariant sites in sequence analysis. *Proceedings of the National Academy of Sciences* **93,** 1930–1934 (1996).
5. Barnes, C. R. On the food of green plants. *Botanical Gazette* **18,** 403–411 (1893).
6. Ashida, H., Danchin, A. & Yokota, A. Was photosynthetic RuBisCO recruited by acquisitive evolution from RuBisCO-like proteins involved in sulfur metabolism? *Research in microbiology* **156,** 611–618 (2005).
7. Moore, B., Ku, M. & Edwards, G. Expression of C4-like photosynthesis in several species of Flaveria. *Plant, Cell & Environment* **12,** 541–549 (1989).
8. Bowman, J. L., Floyd, S. K. & Sakakibara, K. Green genes—comparative genomics of the green branch of life. *Cell* **129,** 229–234 (2007).
9. Badger, M. & Andrews, T. In *Progress in photosynthesis research* 601-609 (Springer, 1987).
10. Sage, R. F., Sage, T. L. & Kocacinar, F. Photorespiration and the evolution of C4 photosynthesis. *Annual Review of Plant Biology* **63,** 19–47 (2012).
11. Hatch, M. & Slack, C. Photosynthesis by sugar-cane leaves, a new carboxylation reaction, and the pathway of sugar formation. *Biochemical Journal* **101,** 103–111 (1966).
12. Wang, Y., Bräutigam, A., Weber, A. P. & Zhu, X.-G. Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *Journal of Experimental Botany* **65,** 3567–3578, doi: 10.1093/jxb/eru058 (2014).
13. Paulus, J. K., Schlieper, D. & Groth, G. Greater efficiency of photosynthetic carbon fixation due to single amino-acid substitution. *Nature Communications* **4,** 1518, doi: 10.1038/ncomms2504 (2013).
14. Hatch, M. D. C 4 photosynthesis: a unique blend of modified biochemistry, anatomy and ultrastructure. *Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics* **895,** 81–106 (1987).
15. Sage, R. F. & Monson, R. K. *C4 plant biology*. (Academic Press, CA, 1999).
16. Brown, W. V. Variations in anatomy, associations, and origins of Kranz tissue. *American Journal of Botany* **62,** 395–402 (1975).
17. Voznesenskaya, E. V., Franceschi, V. R., Kiirats, O., Freitag, H. & Edwards, G. E. Kranz anatomy is not essential for terrestrial C4 plant photosynthesis. *Nature* **414,** 543–546 (2001).
18. Hibberd, J. M. & Quick, W. P. Characteristics of C4 photosynthesis in stems and petioles of C3 flowering plants. *Nature* **415,** 451–454 (2002).
19. Kriedemann, P. The photosynthetic activity of the wheat ear. *Annals of Botany* **30,** 349–363 (1966).
20. Evans, L. & Rawson, H. M. Photosynthesis and respiration by the flag leaf and components of the ear during grain development in wheat. *Australian journal of biological sciences* **23,** 245–254 (1970).
21. Ray, D. K., Mueller, N. D., West, P. C. & Foley, J. A. Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* **8,** e66428, doi: 10.1371/journal.pone.0066428 (2013).
22. Parry, M. A. *et al.* Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *Journal of Experimental Botany* **62,** 453–467, doi: 10.1093/jxb/erq304 (2011).
23. Long, S. P., Marshall-Colon, A. & Zhu, X.-G. Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential. *Cell* **161,** 56–66 (2015).
24. Leegood, R. C. Strategies for engineering C 4 photosynthesis. *Journal of Plant Physiology* **170,** 378–388 (2013).

25. Pingault, L. *et al.* Deep transcriptome sequencing provides new insights into the structural and functional organization of the wheat genome.. *Genome biology* **16,** 29 (2015).
26. Bräutigam, A. *et al.* An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology* **155,** 142–156 (2011).
27. Aoyagi, K. & Chua, N.-H. Cell-Specific Expression of Pyruvate, Pi Dikinase *In Situ* mRNA Hybridization and Immunolocalization Labeling of Protein in Wheat Seed. *Plant physiology* **86,** 364–368 (1988).
28. Aoyagi, K., Bassham, J. A. & Greene, F. C. Pyruvate orthophosphate dikinase gene expression in developing wheat seeds. *Plant physiology* **75,** 393–396 (1984).
29. Gillies, S. A., Futardo, A. & Henry, R. J. Gene expression in the developing aleurone and starchy endosperm of wheat. *Plant biotechnology journal* **10,** 668–679 (2012).
30. Henry, R. J. & Nevo, E. Exploring natural selection to guide breeding for agriculture. *Plant biotechnology journal* **12,** 655–662 (2014).
31. Kersey, P. J. *et al.* Ensembl Genomes 2013: scaling up access to genome-wide data. *Nucleic acids research* **42,** D546–D552 (2014).
32. Jennings, A. & Morton, R. Changes in carbohydrate, protein, and non-protein nitrogenous compounds of developing wheat grain. *Australian Journal of Biological Sciences* **16,** 318–331 (1963).
33. Duffus, C. & Rosie, R. Some enzyme activities associated with the chlorophyll containing layers of the immature barley pericarp. *Planta* **114,** 219–226 (1973).
34. Wirth, E., Kelly, G., Fischbeck, G. & Latzko, E. Enzyme activities and products of CO2 fixation in various photosynthetic organs of wheat and oat. *Zeitschrift für Pflanzenphysiologie* **82,** 78–87 (1977).
35. Singal, H., Sheoran, I. & Singh, R. *In vitro* enzyme activities and products of 14CO2 assimilation in flag leaf and ear parts of wheat (Triticum aestivum L.). *Photosynthesis Research* **8,** 113–122 (1986).
36. von Caemmerer, S., Ghannoum, O., Pengelly, J. J. & Cousins, A. B. Carbon isotope discrimination as a tool to explore C4 photosynthesis. *Journal of experimental botany* **65,** 3459–3470 (2014).
37. Monneveux, P. *et al.* Relationships between grain yield, flag leaf morphology, carbon isotope discrimination and ash content in irrigated wheat. *Journal of Agronomy and Crop Science* **190,** 395–401 (2004).
38. Merah, O., Deleens, E., Nachit, M. & Monneveux, P. Carbon isotope discrimination, leaf characteristics and grain yield of interspecific wheat lines and their durum parents under Mediterranean conditions. *CEREAL RESEARCH COMMUNICATIONS*. 143–149 (2001).
39. Yakir, D., Osmond, B. & Giles, L. Autotrophy in Maize Husk Leaves Evaluation Using Natural Abundance of Stable Isotopes. *Plant physiology* **97,** 1196–1198 (1991).
40. Pyankov, V. I. *et al.* Occurrence of C3 and C4 photosynthesis in cotyledons and leaves of Salsola species (Chenopodiaceae). *Photosynthesis Research* **63,** 69–84 (2000).
41. Pyankov, V. I. *et al.* Features of photosynthesis in Haloxylon species of Chenopodiaceae that are dominant plants in Central Asian deserts. *Plant and Cell physiology* **40,** 125–134 (1999).
42. Ueno, O. Induction of Kranz anatomy and C4-like biochemical characteristics in a submerged amphibious plant by abscisic acid. *The Plant Cell* **10,** 571–583 (1998).
43. Ueno, O., Samejima, M., Muto, S. & Miyachi, S. Photosynthetic characteristics of an amphibious plant, Eleocharis vivipara: expression of C4 and C3 modes in contrasting environments. *Proceedings of the National Academy of Sciences* **85,** 6733–6737 (1988).
44. Gu, S., Yin, L.-y. & Wang, Q.-f. Phosphoenolpyruvate carboxylase in the stem of the submersed species Egeria densa may be involved in an inducible C 4-like mechanism. *Aquatic Botany* **125,** 1–8 (2015).
45. Langdale, J. A., Zelitch, I., Miller, E. & Nelson, T. Cell position and light influence C4 versus C3 patterns of photosynthetic gene expression in maize. *The EMBO journal* **7,** 3643–3651 (1988).
46. Aoyagi, K. & Bassham, J. A. Appearance and accumulation of C4 carbon pathway enzymes in developing wheat leaves. *Plant physiology* **80,** 334–340 (1986).
47. Jia, S. *et al.* Response of wheat ear photosynthesis and photosynthate carbon distribution to water deficit. *Photosynthetica* **53,** 95–109 (2015).
48. Christin, P.-A., Arakaki, M., Osborne, C. P. & Edwards, E. J. Genetic enablers underlying the clustered evolutionary origins of C4 photosynthesis in angiosperms. *Molecular biology and evolution* **32,** 846–858 (2015).
49. Aubry, S., Brown, N. J. & Hibberd, J. M. The role of proteins in C3 plants prior to their recruitment into the C4 pathway. *Journal of Experimental Botany* **62,** 3049–3059 (2011).
50. Morrison, I. The structure of the chlorophyll-containing cross cells and tube cells of the inner pericarp of wheat during grain development. *Botanical Gazette*, 85–93 (1976).
51. Hibberd, J. M. & Covshoff, S. The regulation of gene expression required for C4 photosynthesis. *Annual review of plant biology* **61,** 181–207 (2010).
52. Long, J. J., Wang, J.-L. & Berry, J. O. Cloning and analysis of the C4 photosynthetic NAD-dependent malic enzyme of amaranth mitochondria. *Journal of Biological Chemistry* **269,** 2827–2833 (1994).
53. Majeran, W., Cai, Y., Sun, Q. & van Wijk, K. J. Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. *The Plant Cell* **17,** 3111–3140 (2005).
54. Wang, J.-L., Turgeon, R., Carr, J. P. & Berry, J. O. Carbon sink-to-source transition is coordinated with establishment of cell-specific gene expression in a C4 plant. *The Plant Cell* **5,** 289–296 (1993).
55. Peisker, M. Models of carbon metabolism in C3-C4 intermediate plants as applied to the evolution of C4 photosynthesis. *Plant, Cell & Environment* **9,** 627–635 (1986).
56. Soreng, R. J. *et al.* A worldwide phylogenetic classification of the Poaceae (Gramineae). *Journal of Systematics and Evolution* **53,** 117–137 (2015).
57. Chalupska, D. *et al.* Acc homoeoloci and the evolution of wheat genomes. *Proceedings of the National Academy of Sciences* **105,** 9691–9696 (2008).
58. Furtado, A. *et al.* A novel highly differentially expressed gene in wheat endosperm associated with bread quality. *Scientific reports* **5,** 10446, doi: 10.1038/srep10446 (2015).
59. Maydup, M. *et al.* The contribution of ear photosynthesis to grain filling in bread wheat (Triticum aestivum L.). *Field Crops Research* **119,** 48–58 (2010).
60. Henry, R. J., Rangan, P. & Furtado, A. Functional cereals for production in new and variable climates. *Current Opinion in Plant Biology* **30,** 11–18 (2016).
61. TGI. ftp://occams.dfci.harvard.edu/pub/bio/tgi/data/Triticum_aestivum/, 2010) date accesed 15/07/2016.
62. Götz, S. *et al. et al.* High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research* **36,** 3420–3435 (2008).
63. IWGSC. *TGAC WGS assemblies of other wheat species*,<https://urgi.versailles.inra.fr/download/iwgsc/TGAC_WGS_assemblies_of_other_wheat_species/>(2014) date accessed 15/07/2016.
64. Nystedt, B. *et al.* The Norway spruce genome sequence and conifer genome evolution. *Nature* **497,** 579–584 (2013).
65. Mayer, K. F. *et al.* A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* **345,** 1251788 (2014).
66. Rensing, S. A. *et al.* The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science* **319,** 64–69 (2008).

67. Banks, J. A. *et al.* The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science* **332,** 960–963 (2011).
68. Soltis, D. E. *et al.* The Amborella genome: an evolutionary reference for plant biology. *Genome Biol* **9,** 10.1186 (2008).
69. Christin, P.-A. *et al.* Adaptive evolution of C 4 photosynthesis through recurrent lateral gene transfer. *Current Biology* **22,** 445–449 (2012).
70. Christin, P.-A., Salamin, N., Savolainen, V., Duvall, M. R. & Besnard, G. C 4 photosynthesis evolved in grasses via parallel adaptive genetic changes. *Current Biology* **17,** 1241–1247 (2007).
71. Bennetzen, J. L. *et al.* Reference genome sequence of the model plant Setaria. *Nature biotechnology* **30,** 555–561 (2012).
72. Schnable, P. S. *et al.* The B73 maize genome: complexity, diversity, and dynamics. *science* **326,** 1112–1115 (2009).
73. Paterson, A. H. *et al.* The Sorghum bicolor genome and the diversification of grasses. *Nature* **457,** 551–556 (2009).

Acknowledgements

We thank Grain Foods CRC for supporting the generation of the transcriptome data. PR is supported by an Indo-Australian Career Boosting Gold Fellowship from the Department of Biotechnology, Government of India.

Author Contributions

P.R. and A.F. conducted the analysis. All authors conceived and designed experiments and wrote and edited the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Rangan, P. *et al.* New evidence for grain specific C_4 photosynthesis in wheat. *Sci. Rep.* **6**, 31721; doi: 10.1038/srep31721 (2016).