

ARTICLE

Received 9 Jun 2010 | Accepted 17 Aug 2011 | Published 20 Sep 2011

DOI: 10.1038/ncomms1482

# Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae

V. Prasad<sup>1</sup>, C.A.E. Strömberg<sup>2</sup>, A.D. Leaché<sup>2</sup>, B. Samant<sup>3</sup>, R. Patnaik<sup>4</sup>, L. Tang<sup>5</sup>, D.M. Mohabey<sup>6</sup>, S. Ge<sup>5</sup> & A. Sahni<sup>7</sup>

Rice and its relatives are a focal point in agricultural and evolutionary science, but a paucity of fossils has obscured their deep-time history. Previously described cuticles with silica bodies (phytoliths) from the Late Cretaceous period (67–65 Ma) of India indicate that, by the latest Cretaceous, the grass family (Poaceae) consisted of members of the modern subclades PACMAD (Panicoideae–Aristidoideae–Chloridoideae–Micrairoideae–Arundinoideae–Danthonioideae) and BEP (Bambusoideae–Ehrhartoideae–Pooideae), including a taxon with proposed affinities to Ehrhartoideae. Here we describe additional fossils and show that, based on phylogenetic analyses that combine molecular genetic data and epidermal and phytolith features across Poaceae, these can be assigned to the rice tribe, Oryzeae, of grass subfamily Ehrhartoideae. The new Oryzeae fossils suggest substantial diversification within Ehrhartoideae by the Late Cretaceous, pushing back the time of origin of Poaceae as a whole. These results, therefore, necessitate a re-evaluation of current models for grass evolution and palaeobiogeography.

<sup>1</sup> Birbal Sahni Institute of Palaeobotany, 53 University Road, Lucknow 226 007, India. <sup>2</sup> Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, Washington 98195, USA. <sup>3</sup> Postgraduate Department of Geology, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur 440001, India. <sup>4</sup> Centre for Petroleum and Applied Geology, Panjab University, Chandigarh 160014, India. <sup>5</sup> State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, PR China. <sup>6</sup> Geological Survey of India, Nagpur 440006, India. <sup>7</sup> Centre of Advanced Study in Geology, Lucknow University, Lucknow 226007, India. Correspondence and requests for materials should be addressed to C.A.E.S. (email: caestrom@uw.edu).

The critical dependence of humans on rice (*Oryza sativa*) as a crop species has made it a major research focus in both basic and applied plant science<sup>1–3</sup>. Numerous studies have sought to elucidate its domestication history<sup>4–7</sup> clarify the relationships of species within the genus *Oryza* and to other genera in the rice tribe<sup>2,8,9–11</sup>, and track the evolutionary history of the subfamily that rice belongs to, Ehrhartoideae, within the grass family as a whole<sup>12–14</sup>. In these studies, the Ehrhartoideae is reconstructed as consisting of the tribes Oryzeae, Ehrharteae and Phyllorachidae<sup>15</sup>. The subfamily is commonly found to be nested within a monophyletic group consisting of the Bambusoideae, Ehrhartoideae and the Pooideae (the BEP clade)<sup>13,16,17</sup>; alternatively, the Bambusoideae, Ehrhartoideae and the Pooideae clades form a paraphyletic grade with reference to the PACMAD (Panicoideae–Aristidoideae–Chloridoideae–Micrairoideae–Arundinoideae–Danthonioideae) clade, with Pooideae as the sister taxon to the PACMADs<sup>18,19</sup>. Recent molecular clock estimates calibrated by selected grass fossils indicate that the Ehrhartoideae arose after the K–Pg (Cretaceous–Paleogene) boundary, by the Early Eocene (~53 Ma)<sup>17</sup>, that the Oryzeae diverged from the Ehrharteae by the Late Eocene–Early Oligocene ( $34.5 \pm 6.8$  Ma)<sup>16</sup> and that *Oryza* emerged very recently, by the Middle Miocene period (15–14)<sup>10</sup>. However, beyond these analytical age-estimates, very little is known about the evolutionary history of ehrhartoid grasses, and the rice tribe in particular. The macrofossil record of these groups is virtually absent, with the oldest known macrofossil of the Oryzeae extending back only to the Late Miocene<sup>20,21</sup>. Prasad *et al.*<sup>22</sup> described microscopic silica bodies (phytoliths) of grasses in dinosaur coprolites from the Late Cretaceous period (67–65 Ma) of India, some of which showed affinity to the Ehrhartoideae (*Matleylites indicum*) but could not be clearly assigned to this subfamily based on available data.

The coprolites studied by Prasad *et al.* were recovered from the Maastrichtian (67–65 Ma) Lameta Formation, which consists of alluvial to limnic sediments associated with the Deccan Volcanics (infra- and inter-trappean) of Central India. These deposits are known for their wealth of fossils, including dinosaur bones, coprolites, egg-shell fragments, remains of turtles, frogs, fishes and mammals, as well as macro- and micro-floral remains<sup>23–25</sup>. Continued sampling<sup>22,26,27</sup> both from Lameta Formation sediments and associated titanosaurid sauropod coprolites from Pisdura (east and south sections) and sediments of the Lameta Formation at Dhamni, Piraya and Polgaon has yielded grass pollen and additional leaf cuticle with embedded grass silica short cell (GSSC) phytoliths similar to those found in oryzoid grasses (often referred to as ‘oryzoid’ silica bodies<sup>28</sup>).

Here we report on these newly recovered fossil cuticles and associated phytoliths (referred to as morphotypes *Changii indicum* and *Tateokai deccana*) and use them to suggest a different scenario and timing for the emergence and diversification of the rice tribe than what is widely assumed among evolutionary biologists. We first infer a phylogenetic tree for the Poaceae based on molecular and morphological data, which allows us to deduce the placement of the new fossils in the context of 56 extant grass species. Divergence times for Poaceae lineages are then estimated using seven previously published fossils and our new fossils, and this framework is employed to investigate the evolutionary history of seven leaf epidermis and phytolith morphology characters. These characters were also scored for 14 taxa formerly used in phylogenetic analysis of the Ehrhartoideae<sup>10</sup>. Our analysis reveals that *Changii indicum* and *Tateokai deccana* most likely were members of the Oryzeae, and that diversification of the rice tribe began before the end of the Cretaceous.

## Results

### Systematic palaeontology.

**Fossil phytolith morphotype.** *Changii indicum* (Fig. 1a).

**ICPN name.** Vertical bilobate/cross with pointed lobes epidermal short cell + rectangular sinuate multi-papillate epidermal long cells + papillate subsidiary cells.

**Etymology.** Genus named in honour of Te-Tzu Chang, who has done pioneering work on the biosystematics, evolution and biogeography of rice, including theories regarding a Gondwana origin of rice: species referring to the country India.

**Holotype.** One piece of cuticle, BSIP 13160 (Q-14-3).

**Locality and horizon.** Coprolites from Red clays, Lameta Formation; Pisdura East and Pisdura South (Fig. 2).

**Other referred specimens:** BSIP 13152 (H-28-1), BSIP 13152 (H-27), BSIP 13153 (T-14), BSIP 3161 (Q-29-1).

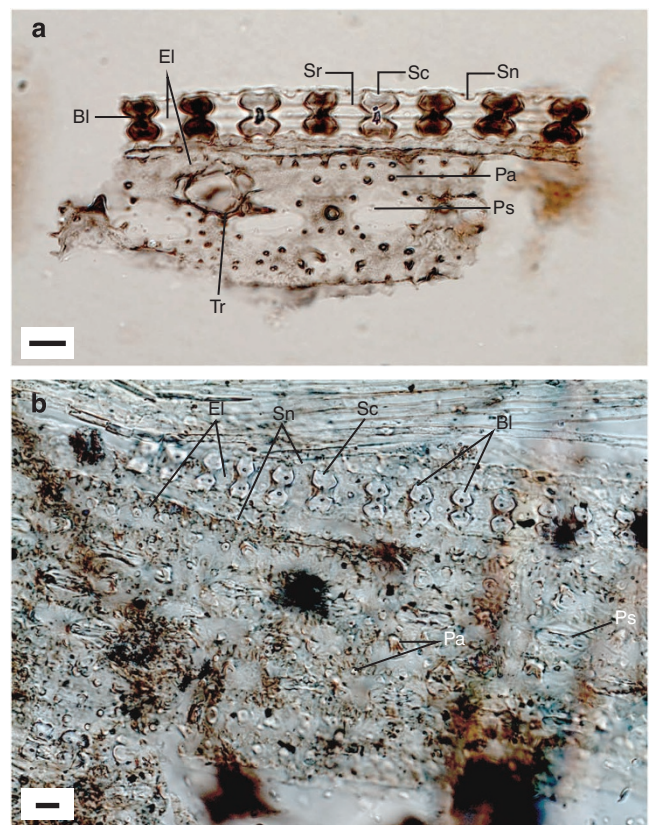
**Generic Diagnosis.** Costal area (overlying vein) of grass leaf consisting of eight closely spaced, bilobate GSSC arranged in a row on sinuate, distinctly silicified, ribbed, thickened, epidermal long cells. The remaining epidermal long cells have numerous thin-walled, more or less evenly distributed, simple conical papillae, some of which show bifurcation at the base. The subsidiary cells are distinctly quadri-papillate. The GSSC are highly regular in their morphology with well-developed shanks and lobes. The bilobates have markedly scooped ends and are oriented with their long axis at right angles to the papillate epidermal long cells (vertically), showing the characteristic ‘oryzoid’ pattern.

**Specific diagnosis.** As for genus.

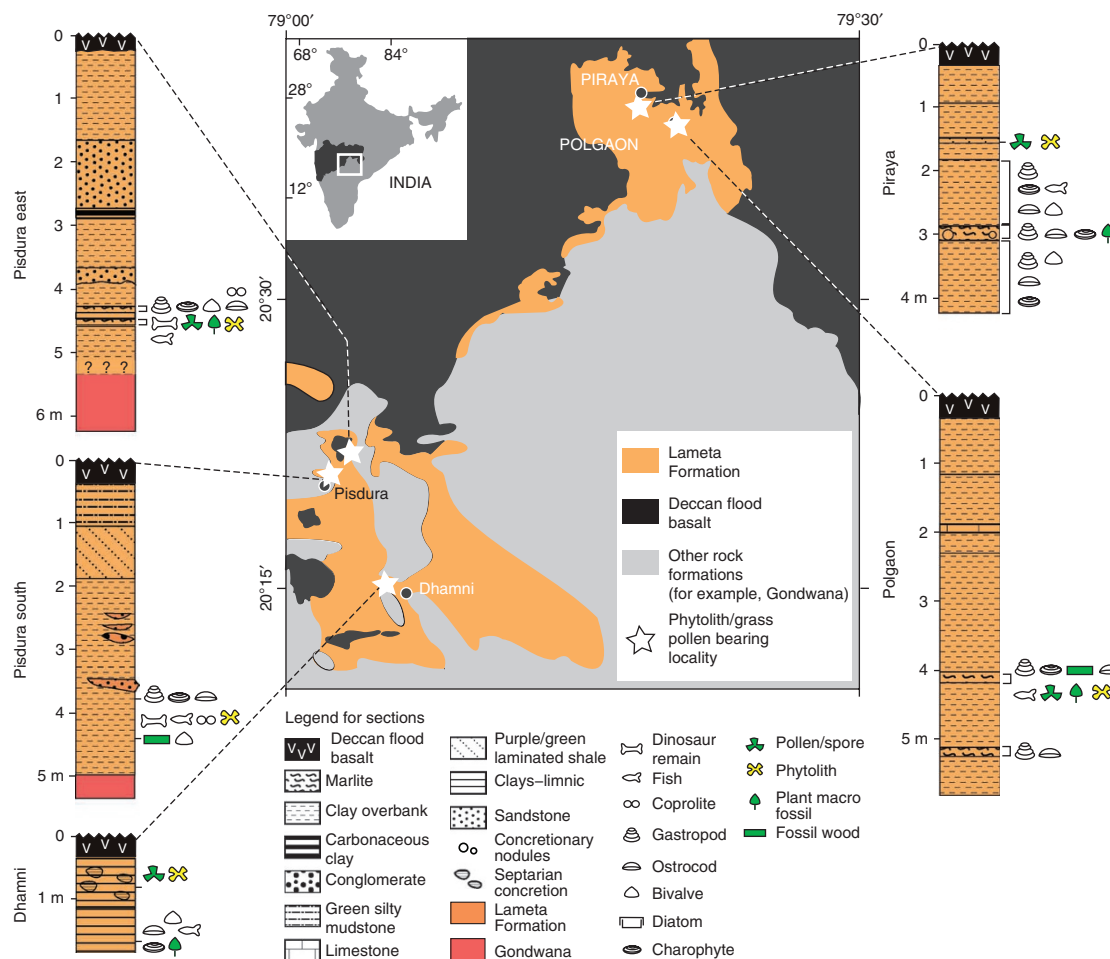
**Fossil phytolith morphotype.** *Tateokai deccana* (Fig. 1b)

**ICPN name.** Vertical bilobate with scooped-end epidermal short cell + rectangular sinuate multi-papillate epidermal long cells.

**Etymology.** Genus named in honour of Tuguo Tateoka, who has done extensive work on the morphology, anatomy, and systemat-



**Figure 1 | Fossil cuticle with phytoliths.** (a) *Changii indicum*, (b) *Tateokai deccana*. Scale bar, 10 µm. Bl = bilobate; El = epidermal long cells; Sn = sinuate margin; Sc = scooped; Sr = silicified ribbed thickening; Pa = papillae; Ps = papillae on subsidiary cells; Tr = trichome.



**Figure 2 | Map and stratigraphy of localities.** The localities have yielded grass epidermis and phytoliths in coprolites (Pisdura East and South) and in sediment (Dhamni, Piraya and Polgaon) of the Late Cretaceous (Maastrichtian) Lameta Formation in Central India<sup>26</sup>. The Polgaon and Pisdura East sections have also produced grass pollen.

ics of the rice tribe and relatives; species referring to the Deccan intertrappean beds.

**Holotype.** One piece of epidermal tissue, SCPGNU 1 (C-48-3).

**Locality and horizon.** Septeria concretions in clays Lameta Formation, Dhamni (Fig. 2).

**Generic Diagnosis.** Costal area of grass leaf consisting of four rows of 14–17 closely spaced bilobate GSSC arranged on rows of epidermal long cells. Remaining epidermal long cells have numerous, thin-walled, densely distributed, simple and bifurcated, long papillae. The subsidiary cells have 2–4 papillae. The GSSC are highly regular in their morphology with well-developed shanks and lobes. The bilobates have deeply scooped ends and are oriented with their long axis at right angles to the papillate epidermal long cells (oryzoid pattern).

**Specific diagnosis.** As for genus.

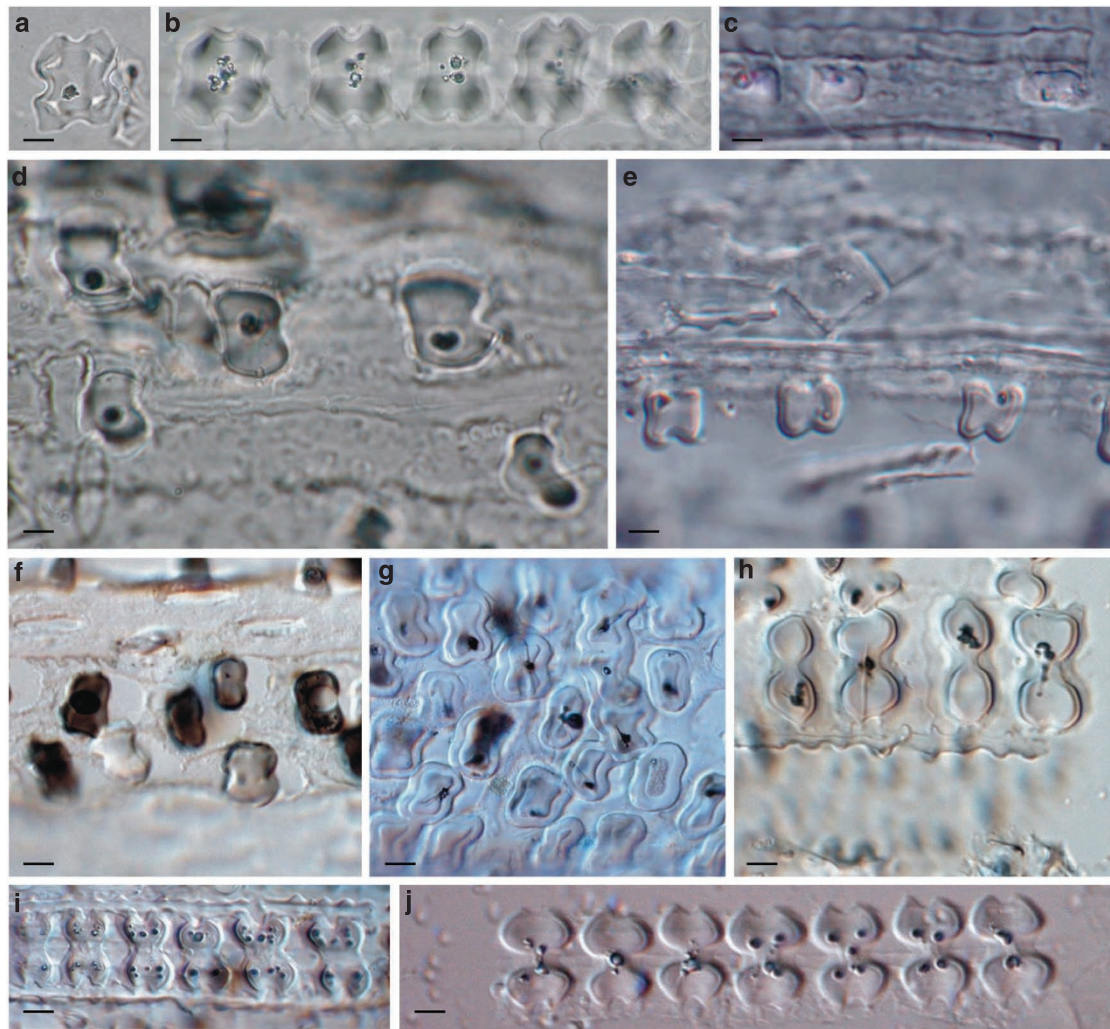
**Additional fossil phytolith morphotypes.** Apart from these more complete specimens, a variety of phytolith morphotypes with more or less clear affinity to Ehrhartoideae based on the distribution of anatomical traits across Poaceae in our analysis (see below), have been found in addition to those reported in Prasad *et al.* (2005). These include an isolated row of vertical scooped end bilobate/cross and attached papillate epidermal long cells (BSIP 1380, L-48-2), isolated scooped end bilobate (GSSC) referred as *Matleyites indicum* (BSIP 13162, K-2; BSIP 13162, K-56-4; BSIP 13153, T-19-2)], and one specimen of papillate long

epidermis along with two associated keystone bulliform morphotypes characteristic of Bambusoideae and Ehrhartoideae clades (BSIP 13153, P-26-3).

**Character analysis of modern reference samples.** The Joinvillaceae, one of the putative sister groups of grasses, share with the Poaceae the unique division into long cells and short cells of the epidermis, with at least some short cells containing silica<sup>13</sup>; the other closely related taxa (for example, Ectocoleaceae and Flagellariaceae) lack this arrangement or any well-defined phytoliths<sup>29</sup>. Silica short cells with distinctive bilobate or cross shape, organized in costal rows, such as those found in *Changii indicum* and *Tateokai deccana* are exclusive to the grass family<sup>29</sup> (C.A.E.S., unpublished data), firmly establishing the affinity of the fossils to the Poaceae.

Although many of the grasses included in the overall Poaceae dataset have been previously investigated<sup>22</sup>, some of the character state scores differ in the current study. These changes are a result of further research, and slightly different and more rigorous definition of characters and character states. For example, the genera *Lithachne*, *Olyra* and *Pariana* (all commonly placed in the Olyreae) are reported in the literature as having ‘oryzoid’ bilobates, meaning that they are oriented with their long axis perpendicular to the epidermal long cells<sup>28</sup>, and they were scored as such in Prasad *et al.*<sup>22</sup> as part of the conservative, inclusive approach to character state assignment taken in that paper. However, on careful inspection of phytoliths from these taxa, we find that their costal short cells are





**Figure 3 | Examples of samples from modern reference collection material used in this study.** (a,b) Cross variant 3 and 8 with a nodule at each lobe (*Lithachne pauciflora*); single bilobate (a), in a row (b). (c) Row of horizontally oriented oblong bilobates (*Ehrharta* sp.). (d) Row of bilobates grading into polylobate or crenate short cells (*Pseudosasa usawai*). (e) Row of horizontal bilobates (*Potamophila* sp.). (f,g) Vertical bilobates with rounded ends arranged randomly in the leaf tissue; *Zizaniopsis* sp. (f), *Rhynchoriza* sp. (g). (h) Row of vertical bilobates with rounded to slightly concave ends (*Zizania* sp.). (i,j) Row of vertical bilobates-crosses with deeply concave (scooped) ends; *Leersia hexandra* (i), *Oryza sativa* (j). Scale bar, 5  $\mu$ m.

very different from the scooped bilobates of the Oryzeae (Fig. 3). They are most commonly cross-shaped and characterized by a ‘top’, consisting of a more or less flat plate with square lobes at the corners, with the ‘base’ demarcated by four large bumps or nodules at the lobes (Fig. 3a,b). Piperno<sup>30–32</sup> recognized this distinct shape as ‘cross variant 3 and 8’. Similarly, although *Pseudosasa* is not included in the phylogenetic analysis, it can be noted that its bilobates do not have well-developed shanks of typical Oryzeae bilobates and many have an extra lobe along the shank, making them more similar to polylobate or crenate forms (Fig. 3d).

**Fossil placement and estimation of ancestral character states.** The phylogenetic analysis showed that the fossils form a clade with in the Oryzeae (posterior probability=0.8) and more inclusively within the Ehrhartoideae (posterior probability=0.88) and display a combination of vertical bilobates in the costal row (character 2; Table 1) with scooped shape (character 3), papillate long cells (character 5), and papilla on the subsidiary cells of the stomata (character 6) (Fig. 4). Bayesian estimation of ancestral character states for characters 2 and 5 indicates that these character states evolved in the Oryzeae and, hence, are derived with respect to the Ehrhartoideae.

An anatomical survey of many of the species and genera used in phylogenetic analysis of Oryzeae<sup>10</sup> confirms the presence of these traits in all studied Oryzeae genera (Supplementary Table S1). Papilla on long cells appear to have evolved independently in the Bambusoideae as well as in the Puelioideae and certain PACMAD taxa (represented by *Distichlis*, *Amphipogon* and *Micraira*) (Fig. 4). Our analysis and survey of Oryzeae taxa further suggest that vertical bilobates with scooped shape (character 3) are unique to taxa nested in the Oryzeae (Supplementary Table S1). In addition, papillate stomatal subsidiary cells (character 6), seem to have evolved in the closely related genera (*Oryza* and *Leersia*) within the Oryzeae tribe, as well as in certain bambusoid taxa (*Chusquea* and *Olyreae*, represented by *Lithachne* and *Olyra*) (Fig. 4; Supplementary Table S1). Remaining characters (1, 4 and 7) are either highly homoplastic (character 1) or have too much missing data (4 and 7) to be informative with respect to character evolution and phylogenetic placement of the fossil taxa.

**Dating analysis.** Because of the relatively low branch support (0.8) for the preferred placement of the fossils as part of Oryzeae, we considered three alternative placements of the fossils for the dating

analysis: at the base of Ehrhartoideae (H<sub>1</sub>), at the base of a clade consisting of PACMAD and Pooideae (H<sub>2</sub>), and at the base of the (PACMAD, Ehrhartoideae) clade (H<sub>3</sub>) (Supplementary Fig. S1). Dating using BEAST showed that inclusion of the new fossils affects dating of all nodes regardless of where they are placed (preferred placement, H<sub>1</sub>–H<sub>3</sub>) (Table 2; Supplementary Fig. S2). However, the

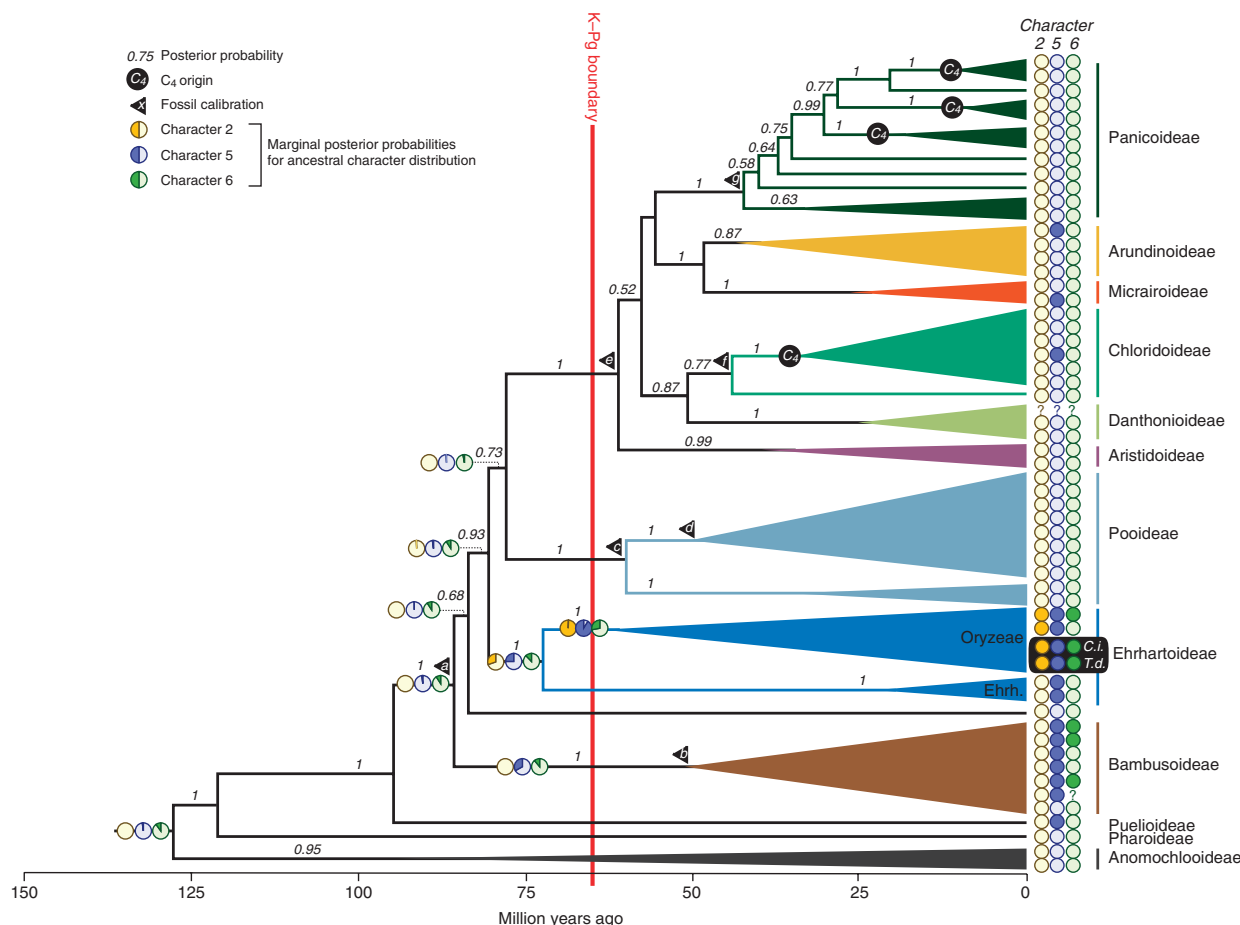
effect is greatest when the fossils are placed in Oryzeae (preferred placement) and is larger both in absolute terms (millions of years) and proportionally for deeper nodes. For example, for the (Bambusoideae, PACMAD) clade, the most conservative hypothesis for the placement of the fossils, H<sub>3</sub>, results in a >6% older date (76 Ma as opposed to 72 Ma without the new fossils), whereas the preferred placement results in a 20% difference (86 Ma). The corresponding difference for the much younger clade of C<sub>4</sub> grasses within the Paniceae X=9 group is a <6% difference (19 Ma instead of 18 Ma) for H<sub>3</sub> and 13% (20 Ma) for the preferred placement of the fossils. The largest effect is on the date for the origin of the clade that includes Oryzeae and the fossils under the preferred placement, with a difference in age of 124% (63 Ma instead of 28 Ma). The preferred placement produces an age for the root of Poaceae of 128 Ma, 20% older than when the new fossils are excluded; ages for H<sub>1</sub>–H<sub>3</sub> range from 115 to 121 Ma.

**Table 1 | Morphological characters and character states.**

Character number	Character	Character state number	Character state description
1	Costal rows of GSSC	0	Absent
		1	Present
2	Vertical bilobates in costal row	0	Absent
		1	Present
3	Scooped shape of vertical bilobates	0	Absent
		1	Present
4	Numerous glandular cells and air spaces in the epidermal long cells	0	Absent
		1	Present
5	Papilla on epidermal long cells	0	Absent
		1	Present
6	Papilla on subsidiaries of stomata	0	Absent
		1	Present
7	Silicification of cells that bear costal GSSC	0	Absent
		1	Present

**Discussion**

Whereas some leaf anatomical and phytolith characters analysed in this study are less phylogenetically informative because they are either highly homoplastic (arrangement of GSSCs) or too much data are currently missing (for example, silicification of cells that bear GSSCs), several characters, when carefully defined, are clearly useful for placing well-preserved fossils in the phylogeny. The Bayesian estimation of ancestral character states across Poaceae, supplemented by our survey of grasses within the Ehrhartoideae, indicate



**Figure 4 | Character state evolution of three of the seven leaf epidermal and phytolith characters across Poaceae.** Poaceae timetree estimated using BEAST version 1.6.1. Placement of the fossils are based on MrBayes phylogenetic analysis based on molecular and morphological data (Supplementary Fig. S1). Pie charts indicating marginal posterior probability distributions (dark shades = presence; light shades = absence of trait). Preferred placement of grass fossils emphasized by black box. Cretaceous–Paleogene (K–Pg) boundary marked with red vertical line. C.i. = *Changii indicum*; T.d. = *Tateokai deccana*. Characters: 2 = vertical bilobates in costal row; 5 = papilla on epidermal long cells; 6 = papilla on subsidiaries of stomata.

**Table 2 | Dating analysis, showing that divergence times differ substantially when the new fossils are included (Supplementary Fig. S9).**

Dating scenario	Summary Statistic (Ma)*	Poaceae	Clade (Pharus, PACMAD)	Clade (PACMAD, Bambusoideae)	Clade (Ehrhartoideae, PACMAD + fossils)	Clade (Pooideae, PACMAD + fossils)	PACMAD	Pooideae	Ehrhartoideae + fossils	Oryzaeae + fossils	Bambusoideae	'Stipoid pooid' (nested in Pooideae)	Chloridoideae	Panicoideae	C <sub>4</sub> d (Chloridoideae)	C <sub>4</sub> a (Paniceae X=9)	C <sub>4</sub> c (Paniceae X=10)	C <sub>4</sub> b (Andropogoneae)
Excluding new fossils	Mean	107.44	101.97	71.69	68.34	67.19	53.12	52.64	49.12	27.98	44.66	44.43	40.92	37.14	30.76	17.80	8.74	8.41
	95% HPD lower	82.75	78.97	58.61	55.67	54.37	43.31	43.37	32.19	13.37	33.49	35.85	31.22	27.94	22.47	10.57	3.50	3.56
	95% HPD upper	134.44	126.09	84.71	81.15	80.24	63.46	62.54	65.48	44.19	55.78	53.32	50.65	46.47	39.37	25.71	14.73	13.94
Fossils at base of Oryzaeae (MrBayes preferred)	Mean	128.87	121.95	86.17	81.33	79.28	61.06	60.27	72.38	62.58	49.26	49.71	45.52	42.09	34.10	20.10	10.03	9.70
	95% HPD lower	99.48	97.56	73.42	69.38	66.79	50.44	49.72	61.40	52.83	37.17	40.23	34.89	32.00	25.48	11.48	3.78	3.89
	95% HPD upper	159.81	149.25	99.65	94.02	92.09	72.86	70.87	83.67	72.47	61.94	59.92	56.99	53.49	43.51	29.56	16.63	15.97
H <sub>1</sub> : Fossils at base of Ehrhartoideae	Mean	121.31	115.40	81.64	77.06	75.25	58.85	57.80	67.14	35.60	47.40	47.93	44.64	40.85	33.64	19.38	9.89	9.26
	95% HPD lower	95.86	92.23	69.55	65.80	63.94	48.93	48.23	56.85	15.80	36.46	39.00	33.88	31.29	24.54	10.89	4.00	3.70
	95% HPD upper	148.87	141.02	93.83	88.58	87.14	69.10	67.59	76.98	56.77	59.72	57.01	55.00	50.79	42.54	28.10	16.32	15.50
H <sub>2</sub> : Fossils at base of (PACMAD, Pooideae)	Mean	116.10	109.79	76.85	73.30	72.21	56.55	55.42	52.86	29.80	46.27	46.13	43.04	39.09	32.43	18.51	9.25	8.93
	95% HPD lower	92.53	88.17	66.14	63.13	62.31	47.31	46.15	34.46	13.26	35.75	38.00	32.82	29.92	24.17	10.61	3.56	3.56
	95% HPD upper	144.47	134.16	87.94	83.47	82.39	66.24	64.70	69.22	46.98	57.68	54.86	52.54	48.08	40.68	26.27	15.48	14.98
H <sub>3</sub> : Fossils at base of (PACMAD, Ehrhartoideae)	Mean	115.06	109.15	76.27	72.96	71.64	55.90	55.22	52.65	29.19	45.72	45.93	42.65	39.29	31.99	18.79	9.36	8.83
	95% HPD lower	91.07	87.22	65.39	63.26	61.43	46.71	46.30	35.37	13.19	34.81	37.59	33.08	30.08	24.00	10.81	3.87	3.63
	95% HPD upper	142.36	133.08	87.25	83.38	82.00	65.17	64.50	68.82	47.34	57.29	54.37	52.25	49.68	40.27	27.11	15.51	14.65

\*HPD = highest posteriority density interval.

that, among extant grasses, vertically oriented bilobate GSSCs with scooped shape are found exclusively in members of tribe Oryzaeae of the Ehrhartoideae subfamily. Previous workers have lumped vertically oriented ('oryzoid') silica bodies produced by several grass genera, for example, *Eriachne*, *Psuedosasa* and the Olyreae (for example, *Pariana*, *Lithachne* and *Olyra*), with those typical of Oryzaeae<sup>28,33</sup>. As a result, these types of GSSCs have been considered only moderately useful from a phylogenetic perspective<sup>34</sup>. In contrast, through the more in-depth study of this character, we found that the bilobate and cross-shaped silica bodies in grasses outside the rice tribe are very different in three-dimensional morphology from the vertical bilobates that signify members of Oryzaeae (Fig. 3), and therefore not directly comparable. The substantial phylogenetic distance separating the taxa that produce different types of vertical silica bodies also supports the hypothesis that this aspect of GSSC morphology is a result of convergence, not common ancestry. Papillate epidermal long cells and papillate stomatal subsidiaries also appear to be diagnostic of certain members of the Oryzaeae, although these traits are also found in a few taxa outside this group (Fig. 4; Supplementary Table S1).

The distribution of character states across Poaceae and the combination of character states in the grass fossils indicate that both *Changii indicum* and *Tateokai deccana* belong in the Oryzaeae, although the reconstruction yields relatively low support. Alternative placements suggested by the phylogenetic analysis nevertheless make both fossils derived with respect to modern subclades such as Bambusoideae. This placement of the Late Cretaceous fossils has radical implications for the amount of diversification that had taken place within the Poaceae at the close of the Mesozoic era. In the phylogeny used herein, it can be inferred that, by 65 Ma (K–Pg boundary), the Ehrhartoideae had split from the Pooideae + PACMAD lineage<sup>18</sup>, or else from the other clades within the BEP<sup>13,16,17</sup>. In addition, the Oryzaeae must have split from the Ehrhartoideae within the Ehrhartoideae. In our dating analysis, placing the fossils as part of the Oryzaeae results, not surprisingly, in extremely old dates for the

Ehrhartoideae and Oryzaeae (72 Ma and 63 Ma, respectively). These dates contrast sharply with the divergence times yielded when only the seven published fossils are used to calibrate the tree (49 Ma and 28 Ma, respectively), and with dates proposed previously based on molecular clock work, namely a late Middle or even Late Eocene (40–35 Ma) origin of the Ehrhartoideae, and a Late Eocene–Early Oligocene (~25 Ma) origin of the Oryzaeae<sup>16–18</sup>. Similarly, the preferred placement of the fossils leads to dates for the origin of the Pooideae and the PACMAD clade that are >20 myr older than published estimates of 39–44 Ma and 38–44 Ma, respectively<sup>16–18</sup>. These differences are in large part due to the inclusion of the fossils described in this paper, but also to some extent a result of consideration of recently published fossils from Patagonia, Argentina (Supplementary Table S2). In addition, our use of certain fossil calibration points differs from previous studies. For example, we use the *Chusquea*-like fossils described from the Late Eocene of North America as a calibration point for the Bambusoideae, whereas Vicentini *et al.*<sup>16</sup> used it to calibrate the BEP–PACMAD split. Note that we did not include the fossils described by Prasad *et al.*<sup>22</sup> because of the uncertainty in the taxonomic affinities of these fossils.

Two sets of dates inferred from our dating analysis are of particular interest. First, under the preferred position of the fossils, we place the oldest origin of C<sub>4</sub> grasses, in the Chloridoideae, in the latest Eocene as opposed to the Early Oligocene (< 32 Ma), as deduced in former studies<sup>16–18,35</sup>. This early timing is consistent with the discovery of C<sub>4</sub> grass pollen from the earliest Oligocene of Spain<sup>36</sup>, and indicates that some C<sub>4</sub> grasses evolved under the high (> 1,000 p.p.m.) atmospheric CO<sub>2</sub> of the Eocene and earliest Oligocene. It has been generally assumed that the C<sub>4</sub> photosynthetic pathway, which today confers an advantage to plants in climates characterized by high temperatures, aridity, summer precipitation, as well as at low pCO<sub>2</sub> (ref. 37) could have evolved only after CO<sub>2</sub> levels had dropped well below 1,000 p.p.m. after the earliest Oligocene (< 32 Ma)<sup>16,18,35,38</sup>. The possible appearance of grasses with C<sub>4</sub> photosynthesis before this time means that a mechanism other than low CO<sub>2</sub> must be



invoked for the evolution of this trait, for example, aridity combined with the relatively elevated temperatures in the Late Eocene<sup>36,39</sup>.

Second, our analyses yield markedly older dates for the origin of the grass family than previous work—regardless of the placement of the new fossils. The estimated age of the Poaceae inferred herein range from 107 Ma (new fossils excluded) to 129 Ma (preferred placement of fossils), compared with published ages of ~70–95 Ma (refs 16–18,40). These new, very old dates for the Poaceae would likely also result in more ancient ages for the monocotyledons and angiosperms as a whole. Molecular clock analyses constrained by the known fossil record of angiosperms have typically provided ages of, respectively, 120–130 Ma and 140–190 Ma, for these groups<sup>41–43</sup>. However, several recent molecularly based dating analyses of land plants point to significantly older dates, with monocotyledons emerging in the Late Jurassic period (161–145 Ma)<sup>44</sup> and angiosperms in the Late Triassic period (235–202 Ma)<sup>44,45</sup>. Our estimated age for the Poaceae is more consistent with these latter analyses. To reconcile these wildly disparate outcomes for the ages of grasses and flowering plants, it may be necessary to re-examine presently employed rate models in molecular clock dating and ways of implementing fossil calibration points<sup>44,45</sup>, as well as to expand and refine our knowledge of Late Cretaceous grasses through paleontological fieldwork and phylogenetic analysis. Our study, therefore, reinforces the importance of the growing record of fossil phytoliths<sup>22,46,47</sup> for shedding light on the hitherto obscure evolutionary history of grasses.

Grasses are today notoriously widespread and attempts at reconstructing the ancestral area of the group using current biogeographic ranges have remained inconclusive, although a Gondwanan origin is widely assumed<sup>17,48</sup>. The fossils described herein, representing a very early stage in the history of the Poaceae, combined with the recalibrated grass phylogeny have the potential of shedding some light on this question. Our recalibration suggests that the grass clade had diversified substantially by the K–Pg transition (Fig. 4); the question is, did this diversification occur on the Indian subcontinent or elsewhere? Because India was biogeographically isolated during the time when the Lameta Formation was deposited<sup>49</sup>, one possibility is that the Poaceae underwent most of its initial taxonomic radiation in India during this isolation. Dispersal of the major grass lineages would have taken place only after India made contact with the South Asian margin in the Early Eocene period (~55 Ma)<sup>49,50</sup>. An alternative scenario that is more concordant with our new, ancient dates for major splits within Poaceae is one in which diversification and dispersal of grass lineages among at least some Gondwanan continents preceded the loss of a land bridge between India and the rest of Gondwana at ~80 Ma (refs 51,52). This hypothesis would also agree better with a recent phylogenetic effort to reconstruct Poaceae biogeographic history that proposes an African origin for both BEP + PACMAD and BEP<sup>17</sup>, because it would imply that these lineage divergences occurred while Africa and India were still in contact. Such a biogeographic scenario remains to be tested in the fossil record; so far, sediment samples from Africa (Madagascar) of Late Cretaceous age have not yielded any (grass or other) phytoliths (Strömberg, unpublished data).

## Methods

**Extraction process.** Coprolites collected from the Pisdura area (Fig. 2) were processed by V.P. After thorough cleaning of the surface, ~2 g of material was taken from the inside of the coprolites. The samples were treated with 40% HCl to remove carbonates, 40% H<sub>2</sub>O<sub>2</sub> to oxidize excess organic material and 5% sodium hexa-meta-phosphate to remove clays. The Lameta sediment samples from Pisdura, Piraya and Dhamni locality were collected and processed using similar methods by B.S. The residue was mounted in Canada Balsam for viewing under a light microscope at ×400 and ×630 magnification and immersion oil.

**Description of fossil phytolith morphotypes.** Fossil cuticles have been given unique names for reference. Protocols for phytolith description and assigning nomenclature are the same as carried out in a previous study<sup>22</sup>. Slides with fossil phytoliths

are housed in the repository of the Birbal Sahni institute for palaeobotany (BSIP), Lucknow, and Postgraduate Department of Geology, Rashtasant Tukadoji Maharaj Nagpur University, Nagpur, India. For each specimen, the number (for example, BSIP 13160) refers to the slide; this is followed by England Finder coordinates.

**Modern reference samples.** Leaves of fourteen modern Ehrhartoideae species from nine genera were sampled in the Laboratory of Systematic and Evolutionary Botany, Chinese Academy of Sciences, Beijing, China; *Phyllostachys* material was collected at the National Botanical Research Institute. In addition, phytoliths extracted from 46 grass taxa (Fig. 4; Strömberg, unpublished data) were studied and complemented by information from the literature<sup>28,32,33</sup>. Because the sister groups of the Poaceae (Joinvillaceae, Ecteiocoleaceae and Flagellariaceae) lack well-defined GSSCs<sup>29</sup> (C.A.E.S., unpublished data), we did not include these in the modern reference set (see Discussion). The material was scored for seven epidermal/phytolith characters (Table 1). Character states were assigned using what appear to be natural breaks in variation; future studies will attempt to refine this scheme through quantitative, statistically based subdivision of character states which was not possible in this study.

**Phylogenetic analysis.** A phylogeny for the Poaceae was estimated through Bayesian inference using MrBayes v3.1.2 (ref. 53) with a combination of plastid molecular data (*rbcL* and *ndhF*) from Christin *et al.*<sup>18</sup> (Supplementary Table S3) and a new morphological dataset. The Christin *et al.* dataset was reduced to 56 taxa for which anatomical data were available, with Anomochloideae (*Anomochloa*, *Streptochaeta*) used as the outgroup. The *rbcL* and *ndhF* genes (3,463 bp) were partitioned and modelled using the GTR + I +  $\Gamma$  model of nucleotide substitution (all parameters were unlinked) and combined with a morphological dataset containing seven characters (Supplementary Table S1). The morphological data used the MK model (Markov k model<sup>54</sup>), and this dataset included the 56 taxa with molecular data as well as the two new fossil taxa (with no molecular data). As none of the morphological characters were constant, we used the commands 'coding = variable' and 'rates = gamma' in MrBayes. Four separate analyses were run; each initiated with a different random starting seed. The analyses used four heated Markov chains (using default heating values) that were run for 50 million generations with sampling intervals of 10,000 steps. Convergence was assessed by inspecting the cumulative posterior probabilities of clades and by comparing the posterior probability values across multiple runs using the online program 'Are we there yet?'<sup>55</sup> (Supplementary Figs S3–S6). Posterior probability values were obtained by summarizing the posterior distribution of trees (post burn-in) with a 50% majority-rule consensus tree (Supplementary Fig. S1).

**Ancestral character state estimation.** Ancestral character states were estimated using Bayesian inference in BayesTraits v1.0 (refs 56,57). Advantages of the Bayesian approach are that it uses reversible-jump Markov chain Monte Carlo to average over all *k*-state models of character evolution and accommodates uncertainty in the phylogeny and the rate parameters of the models<sup>58</sup>. Bayesian inference of ancestral character states was conducted using all post-burn-in trees from the posterior distribution of phylogenies obtained from the MrBayes analysis of the combined molecular and morphological data. Utilizing this posterior distribution of trees allowed us to include the new fossil taxa in the ancestral character state estimation procedure. The BayesTraits analyses were run for 10 million generations, sampling at intervals of 10,000 steps after a burn-in period of 100,000 steps. The rate deviation parameter was set, so that the resulting acceptance rates fell between 0.2 and 0.4 (generally 0.26–0.30). The average of the resulting posterior probability distributions was used to produce final estimates of the probabilities for ancestral state estimates (Supplementary Table S4). We repeated these analyses with a revised posterior distribution of trees that excluded the new fossil taxa to examine the impact of these taxa on ancestral state estimations.

**Dating analysis.** We estimated divergence times for the Poaceae using BEAST v1.6.1 (ref. 59). These analyses utilized the same molecular genetic data and partitioning strategy used in the MrBayes analysis (56 taxa, 3,463 characters, 2 gene partitions). We inferred divergence times by assigning seven fossil calibrations to the tree either in combination with or separately from the new fossil taxa described here (Table 2; Supplementary Table S2; Supplementary Figs S7 and S8). Uncertainties in the fossil calibrations were incorporated into the divergence dating analysis using lognormal prior distributions with soft bounds (no zero offset). We parameterized the log(mean) and log(standard deviation) values of the lognormal distributions using ParameterSolver v2.3 (<https://biostatistics.mdanderson.org/SoftwareDownload/>) so that the upper and lower absolute dates represented the 2.5 and 97.5% quantiles of the distribution. We also examined the impact of each of the seven fossil calibrations by systematically removing each one from the analysis and evaluating the change in the posterior distributions for divergence times (Supplementary Fig. S9). For all analyses, substitution rate variation among lineages was modelled using an uncorrelated lognormal distribution, and we ran the analyses under the birth–death branching prior. Monophyly was not enforced for any nodes. The MCMC analyses were run for 20 million generations, logged every 1,000 steps, and the divergence times for nodes of interest were sampled at the same frequency. We used Tracer v 1.5 to determine convergence, measure the effective sample size of each parameter, and calculate the mean and 95% highest posterior density (HDP) interval for divergence times<sup>60</sup>.

## References

- Delteil, A., Zhang, J., Lessard, P. & Morel, J.-B. Potential candidate genes for improving rice disease resistance. *Rice* **3**, 56–71 (2010).
- Ge, S., Sang, T., Lu, B.-R. & Hong, D.-Y. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proc. Natl Acad. Sci. USA* **96**, 14400–14405 (1999).
- Wang, Z.-F. *et al.* Genetic control of germination ability under cold stress in rice. *Rice Sci.* **16**, 173–180 (2009).
- Fuller, D. Q. & Qin, L. Water management and labour in the origins and dispersal of Asian rice. *World Archaeol.* **41**, 88–111 (2009).
- Zhao, Z. J. & Piperno, D. R. Late Pleistocene/Holocene environments in the Middle Yangtze River Valley, China and rice (*Oryza sativa* L.) domestication: the phytoliths evidence. *Geoarchaeology* **15**, 203–222 (2000).
- Molina, J. *et al.* Molecular evidence for a single evolutionary origin of domesticated rice. *Proc. Natl Acad. Sci. USA* **108**, 8351–8356 (2011).
- Saxena, A., Prasad, V., Singh, I. B., Chauhan, M. S. & Hasan, R. On the Holocene records of wild and cultivated rice from Ganga plain: evidence for rice-based agriculture. *Curr. Sci.* **90**, 1547–1552 (2006).
- Ge, S., Li, A., Lu, B.-R., Zhang, S.-Z. & Hong, D.-Y. A phylogeny of the rice tribe Oryzaceae (Poaceae) based on matK sequence data. *Am. J. Bot.* **89**, 1967–1972 (2002).
- Guo, Y.-L. & Ge, S. Molecular phylogeny of Oryzaceae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genome. *Am. J. Bot.* **92**, 1548–1558 (2005).
- Tang, L. *et al.* Phylogeny and biogeography of the rice tribe (Oryzaceae): Evidence from combined analysis of 20 chloroplast fragments. *Mol. Phylogenet. Evol.* **54**, 266–277 (2010).
- Zou, X.-H. *et al.* Analysis of 142 genes resolves the rapid diversification of the rice genus. *Genome Biol.* **9**, R49 (2008).
- Davis, J. I. & Soreng, R. J. A preliminary phylogenetic analysis of the grass subfamily Pooideae (Poaceae), with attention to structural features of the plastid and nuclear genomes, including an intron loss in GBSSI. *Aliso* **23**, 335–348 (2007).
- Grass Phylogeny Working Group. Phylogeny and subfamilial classification of the grasses (Poaceae). *Ann. Missouri Bot. Gard.* **88**, 373–457 (2001).
- Saarela, J. M. & Graham, S. W. Inference of phylogenetic relationships among the subfamilies of grasses (Poaceae: Poales) using meso-scale exemplar-based sampling of the plastid genome. *Botany* **88**, 65–84 (2010).
- Kellogg, E. A. The evolutionary history of Ehrhartoideae, Oryzaceae, and *Oryza*. *Rice* **2**, 1–4 (2009).
- Vicentini, A., Barber, J. C., Aliscioni, S. S., Giussani, L. M. & Kellogg, E. A. The age of the grasses and clusters of origins of C<sub>4</sub> photosynthesis. *Glob. Change Biol.* **14**, 2963–2977 (2008).
- Bouchenak-Khelladi, Y., Verboom, G. A., Savolainen, V. & Hodkinson, T. R. Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Bot. J. Linn. Soc.* **162**, 543–557 (2010).
- Christin, P.-A. *et al.* Oligocene CO<sub>2</sub> decline promoted C<sub>4</sub> photosynthesis in grasses. *Curr. Biol.* **18**, 37–43 (2008).
- Caetano-Anollés, G. Grass evolution inferred from chromosomal rearrangements and geometrical and statistical features in RNA structure. *J. Mol. Evol.* **60**, 635–652 (2005).
- Thomasson, J. R. *Archaeoleersia nebraskensis* gen. et sp. nov. (Graminae-Oryzaceae), a new fossil grass from the late Tertiary of Nebraska. *Am. J. Bot.* **67**, 876–882 (1980).
- Thomasson, J. R. in *Grass Systematics and Evolution* (eds Soderstrom, T. R., Hilu, K.W., Campbell, C.S. & Barkworth, M.E.) 159–167 (Smithsonian Institution, 1987).
- Prasad, V., Strömberg, C. A. E., Alimohammadian, H. & Sahni, A. Dinosaur coprolites and the early evolution of grasses and grazers. *Science* **310**, 1177–1180 (2005).
- Rana, R. S. & Wilson, G. P. New Late Cretaceous mammals from the Intertrappean beds of Rangapur, India and paleobiogeographic framework. *Acta Palaeontol. Pol.* **48**, 331–348 (2003).
- Wilson, G. P., Das Sarma, D. C. & Anantharaman, S. Late Cretaceous sudamericid gondwanatherians from India with paleobiogeographic considerations of Gondwanan mammals. *J. Vert. Paleontol.* **27**, 521–531 (2007).
- Prasad, G. V. R. & Sahni, A. Late Cretaceous continental vertebrate fossil record from India: Palaeobiogeographical insights. *B. Soc. Geol. Fr.* **180**, 369–381 (2009).
- Mohabey, D. M., Udhoji, S. G. & Verma, K. K. Palaeontological and sedimentological observations on nonmarine Lameta Formation (Upper Cretaceous) of Maharashtra, India: their palaeoecological and palaeoenvironmental significance. *Palaeoogeogr. Palaeoecolimatol. Palaeoecol.* **105**, 83–94 (1993).
- Samant, B. & Mohabey, D. M. Palynoflora from Deccan volcano-sedimentary sequence (Cretaceous-Palaeogene transition) of central India: implications for spatio-temporal correlation. *J. Biosci.* **34**, 811–823 (2009).
- Watson, L. & Dallwitz, M. J. *Grass Genera of the World: Descriptions, Illustrations, Identification, and Information Retrieval; including Synonyms, Morphology, Anatomy, Physiology, Phytochemistry, Cytology, Classification, Pathogens, World and Local Distribution, and References* (<http://delta-intkey.com/>) (1992).
- Prychid, C. J., Rudall, P. J. & Gregory, M. Systematics and biology of silica bodies in monocotyledons. *Bot. Rev.* **69**, 377–440 (2004).
- Piperno, D. R. A comparison and differentiation of phytoliths from maize and wild grasses: use of morphological criteria. *Am. Antiq.* **49**, 361–383 (1984).
- Piperno, D. R. *Phytolith Analysis, an Archaeological and Geological Perspective* (Academic, 1988).
- Piperno, D. R. & Pearsall, D. M. The silica bodies of tropical American grasses: morphology, taxonomy, and implications for grass systematics and fossil phytolith identification. *Smithsonian Contrib. Bot.* **85**, 1–40 (1998).
- Metcalfe, C. R. *Anatomy of the Monocotyledons, Volume I, Gramineae* (Oxford University Press, 1960).
- Kellogg, E. A. & Watson, L. Phylogenetic studies of a large data set. 1. Bambusoideae, Andropogonodae, and Pooideae (Gramineae). *Bot. Rev.* **59**, 273–343 (1993).
- Bouchenak-Khelladi, Y. *et al.* The origins and diversification of C<sub>4</sub> grasses and savanna adapted ungulates. *Glob. Change Biol.* **15**, 2397–2417 (2009).
- Urban, M. A. *et al.* Isotopic evidence of C<sub>4</sub> grasses in southwestern Europe during the Early Oligocene–Middle Miocene. *Geology* **38**, 1091–1094 (2010).
- Sage, R. F. The evolution of C<sub>4</sub> photosynthesis. *New Phytol.* **161**, 341–370 (2004).
- Edwards, E. J., Osborne, C. P., Strömberg, C. A. E. & Smith, S. A. The origins of C<sub>4</sub> grasslands: integrating evolutionary and ecosystem science. *Science* **328**, 587–591 (2010).
- Strömberg, C. A. E. The origin and spread of grass-dominated habitats. *Annu. Rev. Earth Pl. Sci.* **39**, 517–544 (2011).
- Janssen, T. & Bremer, K. The age of major monocot groups inferred from 800+ *rbcl* sequences. *Bot. J. Linn. Soc.* **146**, 385–398 (2004).
- Magallon, S. & Castillo, A. Angiosperm diversification through time. *Am. J. Bot.* **96**, 349–365 (2009).
- Moore, M. J., Bell, C. D., Soltis, P. S. & Soltis, D. E. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc. Natl Acad. Sci. USA* **104**, 19363–19368 (2007).
- Bell, C. D., Soltis, D. E. & Soltis, P. S. The age and diversification of the angiosperms re-visited. *Am. J. Bot.* **97**, 1296–1303 (2010).
- Smith, S. A., Beaulieu, J. M. & Donoghue, M. J. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proc. Natl Acad. Sci. USA* **107**, 5897–5902 (2010).
- Clarke, J. T., Warnock, R. C. M. & Donoghue, P. C. J. Establishing a time-scale for plant evolution. *New Phytol.* doi: 10.1111/j.1469-8137.2011.03794 (2011).
- Strömberg, C. A. E. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proc. Natl Acad. Sci. USA* **102**, 11980–11984 (2005).
- Strömberg, C. A. E. & McInerney, F. A. The Neogene transition from C<sub>3</sub> to C<sub>4</sub> grasslands in North America: Assemblage analysis of fossil phytoliths. *Paleobiology* **37**, 50–71 (2011).
- Bremer, K. Gondwanan evolution of the grass alliance of families (Poales). *Evolution* **56**, 1374–1387 (2002).
- McLoughlin, S. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Aust. J. Bot.* **49**, 271–300 (2001).
- Clyde, W. C., Khan, I. H. & Gingerich, P. D. Stratigraphic response and mammalian dispersal during initial India-Asia collision: evidence from the Ghazij Formation, Balochistan, Pakistan. *Geology* **31**, 1097–1100 (2003).
- Case, J. A. A new biogeographical model for dispersal of Late Cretaceous vertebrates into Madagascar and India. *J. Vert. Paleontol.* **22**, 42A (2002).
- Hay, W. W. *et al.* in *Evolution of the Cretaceous Ocean–Climate System Geological Society of America Special Paper 332* (eds Barrera, E. & Johnson, C. C.) 1–47 (Geological Society of America, 1999).
- Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
- Lewis, P. O. A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* **50**, 913–925 (2001).
- Nylander, J. A. A., Wilgenbusch, J. C., Warren, D. L. & Swofford, D. L. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics* **24**, 581–583 (2008).
- BayesTraits, v1.0. Available from <http://www.evolution.rdg.ac.uk/> (2009).
- Pagel, M., Meade, A. & Barker, D. Bayesian estimation of ancestral character states on phylogenies. *Syst. Biol.* **53**, 673–684 (2004).
- Pagel, M. & Meade, A. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am. Nat.* **167**, 808–825 (2006).
- Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).
- Rambaut, A. & Drummond, A. J. TRACER, v1.4.1. (2009) (<http://tree.bio.ed.ac.uk/software/tracer/>).



## Acknowledgements

We thank R. Olmstead and V. Soza for discussion about phylogenetic methods. P.-A. Christin for supplying the data and tree for the Poaceae phylogeny, and G.P. Wilson for constructive comments on earlier versions of this manuscript. We are also grateful to the International Rice Research Institute (Los Banos, Philippines) for providing seed and leaf samples. R.P. was funded by a PURSE Grant from the Department of Science and Technology, New Delhi. S.G. was supported by the National Natural Science Foundation of China (30990240). Contributions by D.M.M. and B.S. are under the existing MoU for Joint Collaborative work involving the Geological Survey of India and RTM Nagpur University. B.S. is thankful to the Department of Science and Technology, New Delhi (No. SR/FTP/ES-32/2003) for financial assistance.

## Author contributions

V.P. conceived of the project. D.M.M. and A.S. mapped and lithologged the sediments and collected the samples. V.P. and B.S. extracted and studied fossil samples. V.P. and

C.A.E.S. extracted and investigated modern reference samples. A.D.L., C.A.E.S. and R.P. conducted phylogenetic analyses. C.A.E.S. and V.P. wrote the paper with contributions from all other authors.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

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

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Prasad, V. *et al.* Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae. *Nat. Commun.* 2:480 doi: 10.1038/ncomms1482 (2011).