The Evolutionary History of Ehrhartoideae, Oryzeae, and *Oryza*

Elizabeth A. Kellogg

Abstract Rice is a member of the genus *Oryza*, which has a history extending back into the Miocene. *Oryza* is in turn a member of the tribe Oryzeae, which along with the tribe Ehrharteae is included in the subfamily Ehrhartoideae. This paper reviews current knowledge of the genus, tribe and subfamily, and places rice in the larger evolutionary context of the entire grass family. The morphological characteristics of rice are an amalgam of characters that have originated at different times in its long evolutionary history. Increasingly, genomic characteristics are also being placed in a broad evolutionary context and it is becoming possible see which are characteristic of all grasses and which are more restricted to the genus *Oryza* or even to rice itself.

Keywords Rice · Phylogeny · Evolution · Genome ·
Grasses · Ehrharta · Oryza

The position of *Oryza sativa* within the angiosperms

This review will begin with the big picture of rice evolution, locating rice in successively smaller clades of plants. In the process I will describe a few of the characteristics that rice shares with other plants at each taxonomic level and what we know about the placement of rice within each successive group. This paper can thus be viewed as traveling through evolutionary time from the common ancestor of flowering plants (angiosperms) to *O. sativa* and its subspecies (Fig. 1). Much of the data in this first section is taken from Stevens (2001) and extensive references therein.

By placing rice in this very broad evolutionary context, it is possible to assess the level of generality of different aspects of the plant. Thus, for example, studies on megagametophyte development, or double fertilization mechanisms, might be expected to shed light on those processes elsewhere in the angiosperms. Conversely, a study of lemma epidermal development might be applicable only to the genus *Oryza* itself. Every organism is the product of all of its evolutionary history, both ancient and recent, with different characters having appeared at different times. This paper thus presents the time of acquisition of various aspects of the rice plant and the organisms to which it can be compared.

Rice is an angiosperm (flowering plant). With respect to the characteristics that it shares with all other angiosperms it thus can serve as a model for all flowering plants. Like all other angiosperms, rice produces its megagametophytes inside ovules, which are in turn enclosed in an ovary. It also shares with other angiosperms the process of double fertilization, whereby the microgametophyte (in the pollen) produces two sperm, one of which fertilizes the egg cell and one of which fertilizes the central cell to produce endosperm. It also shares with many other angiosperms the ability to produce vessels in the xylem, in addition to tracheids.

Within the angiosperms, rice belongs to the large group known as the monocotyledons, and is thus similar to onions, lilies, orchids, and the thousands of other monocot species. It is only distantly related to the models *Arabidopsis* and tomato, which are eudicotyledons. Rice thus provides a model for investigating characteristics that appear in monocots and not eudicots. Like other monocots, rice has a single cotyledon, whereas the eudicots have two. The

E. A. Kellogg (✉)
Department of Biology, University of Missouri-St. Louis,
One University Boulevard,
St. Louis, MO 63121, USA
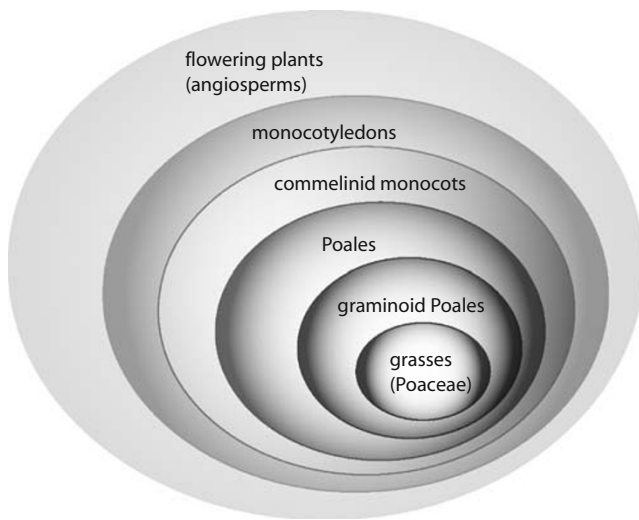e-mail: tkellogg@umsl.edu

Fig. 1 Venn diagram showing the taxonomic hierarchy of major groups in which rice is contained.

primary root is unbranched and short-lived and is replaced early in plant development by a profusion of stem-borne (adventitious) roots (Stevens 2001). The eudicots in contrast generally have a long-lived and branched primary root. Rice and the other monocots also have characteristic vascular patterning. Vascular bundles are scattered in the stem (vs. arranged in a ring in the eudicots). The major veins of the leaf are arranged parallel to each other (vs. in a reticulate pattern). Intermediate leaf veins always connect with other veins, rather than ending blindly in the mesophyll as in eudicots.

Within the monocots, rice falls within the group known informally as the commelinid clade. Other members of this large group are palms (Arecaceae), pineapples (Bromeliaceae), gingers, and bananas (Zingiberales), as well as grasses, sedges, and rushes. Like all other commelinids and unlike all non-commelinid angiosperms (monocot or dicot), rice produces ferulic acid in its cell walls, which makes them fluoresce under ultraviolet light. Most commelinids also produce silica (SiO_2) bodies in their leaves and have bracteate inflorescences. Commelinid stomata generally have two subsidiary cells with the long axis parallel to the stomate (paracytic stomata), or may have an additional two short subsidiaries at each end of the stomate, perpendicular to the long axis (tetracytic). All commelinids except the palms have abundant starch in their endosperm, an ancient characteristic that may have originated about 120 million years ago (Janssen and Bremer 2004). When early human agriculturalists exploited this character, they were thus taking advantage of a tissue that had been fixed by natural selection tens of millions of years earlier.

Limited comparative genomic data hints that the genomes of commelinids may have some similarities. A comparison of a region of almost two megabases (Mb) between two species of banana (*Musa*; Zingiberales) and rice found that both *Musa* and *Oryza* had relatively high GC content (39% and 43%, respectively), when compared either to onions (a non-commelinid monocot) or to *Arabidopsis* (a eudicot; 36%)(Lescot et al. 2008). Even though only a few BACs were sequenced from *Musa*, multiple regions were found with microsynteny to rice, including five to 11 syntenic genes.

Within the commelinids, rice falls in the order Poales, which includes 17 families all of which accumulate silica in the epidermis and have endosperm in which multiple nuclear divisions occur before cell walls are formed (nuclear endosperm). A clade comprising seven of these families (Anarthriaceae, Restionaceae, Centrolepidaceae, Flagellariaceae, Joinvilleaceae, Ecdeiocoleaceae, and Poaceae) is sometimes known as the graminoid Poales (Campbell and Kellogg 1987; Kellogg and Linder 1995) (Fig. 2). It includes plants with distichous (two-ranked) sheathing leaves, monoporate annulate pollen, a single anatropous ovule per carpel, and plumose stigmas. The primary cell wall includes (1-3,1-4)-ß-D-glucans and the sieve tube plastids include cuneate crystals. Most members of the clade (including all Poaceae) are wind pollinated, and as expected for wind-pollinated species, the flowers are small, with inconspicuous tepals, and are often imperfect. Reduction to a single ovule per fruit is common.
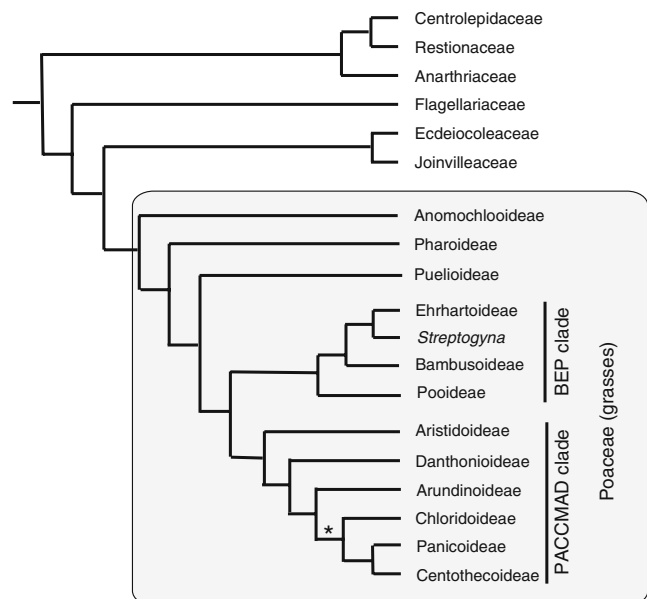

Fig. 2 Phylogeny of the graminoid Poales. Grasses are in the *gray box*. Relationships among grass subfamilies based on data from Vicentini et al. (2008); the position of subfamily Microairoideae is unresolved within the PACCMAD clade, and is not included. *Asterisk* indicates a clade that is poorly supported by the available data. Relationships among the grasses and the other poalean families are based on Marchant and Briggs (2007).

Within the Poales, rice is a member of the Poaceae, the grass family, a family of flowering plants that includes over 10,000 species, including all the major cereal crops (Clayton and Renvoize 1986; Grass Phylogeny Working Group 2001; Watson and Dallwitz 1992). The grasses are distinguished from their poalean relatives by the structure of their embryo, their fruit, and their pollen. In most monocots, including the graminoid Poales, the embryo is scarcely differentiated when the seed is mature (Campbell and Kellogg 1987; Rudall et al. 2005). In the grasses, in contrast, the relative timing of seed maturation and embryo development are shifted such that the grass embryo has a clear shoot and root apical meristem, and two or more differentiated leaves. This could have occurred developmentally either by accelerating embryo development and keeping the seed maturation time constant, or by delaying seed maturation until the embryo has differentiated further (Kellogg 2000). Although clearly derived from a three-carpellary ovary, the grass ovary has a single locule with a single seed; the seed coat generally is appressed or even fused to the inside of the ovary wall. This unique fruit is called a caryopsis or grain.

All but two genera of Poaceae have flowers borne in spikelets ("little spikes") (Grass Phylogeny Working Group 2001). The spikelet is subtended by two sterile bracts, the glumes, and contains one or more flowers. The flowers are sometimes called florets to distinguish them from more conventional monocot flowers. The floret is subtended by a bract-like structure (the lemma). Above the lemma on the floral axis is an adaxial structure, the palea, which is commonly interpreted as a prophyll (Arber 1934), although evidence is accumulating that it might be a highly modified outer perianth whorl (Malcomber et al. 2006; Preston and Kellogg 2007). Above the palea on the floral axis and enclosed within it are two (rarely three) flap-like structures, the lodicules, that clearly represent highly modified inner tepals (Ambrose et al. 2000; Malcomber et al. 2006; Whipple et al. 2004, 2007). Rather than serving an attractive or protective function like other inner perianth whorls, the lodicules become turgid at anthesis and force the lemma and palea apart, permitting exsertion of the stamens and stigmas. Stigma number is ancestrally three, but in most grasses (including rice) has been reduced to two (GPWG 2001). Stamen number is ancestrally six, a number that also appears in rice.

Almost all species of Poaceae fall into one of two large clades, one including the Bambusoideae, Poooideae, and Ehrhartoideae (the BEP clade), and the other including Panicoideae, Arundinoideae, Centothecoideae, Chloridoideae, Micrairoideae, Aristidoideae, and Danthonioideae (the PACCMAD clade)(Bouchenak-Khelladi et al. 2008; Clark et al. 1995; Grass Phylogeny Working Group 2001; Sánchez-Ken et al. 2007) (Fig. 2). The PACCMAD clade is certainly

monophyletic. In contrast, support for the BEP clade, which includes rice, was not strong in early phylogenies (Kellogg and Linder 1995). However, the two most recent phylogenies of the grasses (Bouchenak-Khelladi et al. 2008; Vicentini et al. 2008) both find strong support for monophyly of the clade.

Because rice is in the BEP clade and maize is in the PACCMAD clade, the date of the divergence of these two clades is also the maize–rice divergence date. A recent study places the maize–rice divergence at 51.6±7.9 million years ago (Vicentini et al. 2008). This estimate is quite similar to the 50-million-year date commonly attributed to the comprehensive review of grass evolution provided by Gaut (2002), who based the date in turn on a citation from Stebbins (1981).

Within the BEP clade, monophyly of the included subfamilies is unequivocal, but different studies have resolved different relationships among them. The most recent phylogenies either place Ehrhartoideae as sister to Streptogyna and the two then sister to Bambusoideae (e.g. (Grass Phylogeny Working Group 2001; Vicentini et al. 2008)), or sister to a clade comprised of Bambusoideae plus Pooideae (Bouchenak-Khelladi et al. 2008). The two studies are not comparable, however, in that Bouchenak-Khelladi et al. (2008) included only limited data for Streptogyna and thus its placement was provisional.

Vicentini et al (2008) estimated the divergence of Ehrhartoideae from Streptogyna at 43.1±7.2 million years ago, during the Eocene. This was a period of earth history when global climate was considerably warmer than it is now and there were no polar ice sheets (Zachos et al. 2001).

The discrepancies among different molecular phylogenetic studies can be ascribed in part to taxon sampling. If a major group (e.g. Ehrhartoideae) is represented by only one species (e.g. rice), then DNA sequence(s) from that species will contain many mutations relative to all other species in the study. The large number of mutations effectively randomizes the sequence relative to all the other sequences, such that the rice sequence will fit about equally well (or equally badly) with multiple other sequences. Thus, the differences among phylogenies with few taxa may be simply random (see for example, Kellogg and Linder 1995). With sufficient taxon sampling and enough nucleotides, however, the BEP clade appears to be monophyletic (see for example Bouchenak-Khelladi et al. 2008).

All members of Poaceae retain the signature of a whole genome duplication (Paterson et al. 2004; Wang et al. 2005). Genomic data from a variety of cereals confirm that the duplication occurred before the common ancestor of maize and rice. Data from the early diverging grasses in subfamilies Anomochlooideae, Pharoideae, and Puelioideae are more limited, but phylogenies of many individual gene families (see e.g., Malcomber et al. 2006) indicate that the

duplication occurred shortly before the common ancestor of the grass family (70.9±9.2 Ma)(Vicentini et al. 2008). As expected for duplicated genes, the duplicates have generally diverged in expression pattern and presumably function (Lin et al. 2008; Malcomber et al. 2006; Preston and Kellogg 2007).

Subfamily Ehrhartoideae

Rice belongs to subfamily Ehrhartoideae within the BEP clade. Ehrhartoideae includes the three tribes Oryzeae, Ehrharteae, and Phyllorachidae (Grass Phylogeny Working Group 2001). The placement of the latter is based solely on morphological similarity to the other two tribes; it has not been included in any phylogenetic analysis to date. Most of what we know about the subfamily is thus based on Oryzeae and Ehrharteae. The two tribes are estimated to have diverged 34.5±6.8 million years ago (Vicentini et al. 2008), in the late Eocene or early Oligocene when the earth as a whole was becoming cooler and global CO_2 concentration was dropping to near its pre-industrial levels. Guo and Ge (2005) cite a fossil member of Oryzeae from this time period (Litke 1968), consistent with the molecular date, but Thomasson (1987) referred this fossil simply to an unidentified grass, not specifically oryzoid.

Figure 3 shows the phylogeny of the 13 genera of Ehrhartoideae, created by grafting the phylogeny of *Ehrharta* from Verboom et al. (2003) to that of Oryzeae (Ge et al. 2002; Guo and Ge 2005). No one has undertaken a single comprehensive phylogenetic analysis for the entire subfamily, although this would be straightforward given the appropriate plant material. The tribe Ehrharteae may be divided into the four genera *Ehrharta*, *Tetrarrhena*, *Microlaena*, and *Zotovia* (= *Petriella*) (Watson and Dallwitz 1992), although some authors combine the four into a single genus, *Ehrharta*, with 35 species (Clayton and Renvoize 1986; Willemse 1982). As currently circumscribed, *Ehrharta*, *Tetrarrhena*, and *Microlaena* are polyphyletic.

The ancestral spikelet structure for Ehrhartoideae is a three-flowered spikelet in which the lower two flowers are sterile; this spikelet structure is uniquely derived (synapomorphic) for the subfamily (Grass Phylogeny Working Group 2001). Members of tribe Ehrharteae exhibit this structure clearly, with two glumes, two proximal sterile lemmas (remnants of the lower two flowers), and a single hermaphrodite floret. The internodes between these structures (rachilla) may be more or less elongated. Many members of the tribe Oryzeae, including *O. sativa* itself, retain this same ancestral structure, except that the glumes are minute ("rudimentary glumes" (Bommert et al. 2005)) (Fig. 3). In other members of Oryzeae, the glumes are reduced to a cup-like rim. In addition, the sterile lemmas

may be small or missing. In some species in which glumes and/or sterile lemmas are missing, the elongated internode then appears simply as a long stalk, sometimes called a stipe, above the glumes or above the sterile lemmas (Watson and Dallwitz 1992).

Ample evidence indicates that the rice spikelet is like that of other Ehrhartoideae in having three flowers, of which only the uppermost (distal) is fertile, an interpretation first suggested by Stapf (1917). In addition to the argument from comparative morphology and evolutionary relationship, outlined in the previous paragraph and also discussed by Malcomber et al. (2006), genetic studies show that overexpression of floral genes affects the sterile lemmas but not the true glumes (Komatsu et al. 2003; Prasad et al. 2001). Nonetheless, the interpretation of the rice spikelet as having only one flower persists in the literature and some authors (e.g. Terrell et al. 2001) continue erroneously to refer to the sterile lemmas as "empty glumes".

Members of the subfamily differ in the number of stamens. The ancestral number is clearly six, a number that is retained in *Oryza*. However, one clade of *Microlaena* and *Tetrarrhena* species includes plants with only four stamens, *Ehrharta delicatula* and *Ehrharta triandra* have stamen number reduced to three, and another clade that includes *Zotovia*, two species of *Microlaena* and *Ehrharta avenacea* shares a reduction of stamen number to two (Verboom et al. 2003)(Fig. 3). Independently, stamen number is reduced to one in *Chikusichloa* (Watson and Dallwitz 1992), whereas an increase in stamen number is reported for some species of *Luziola* (Watson and Dallwitz 1992).

The leaf anatomy of Ehrhartoideae was studied in detail by Tateoka (1963) and later by Watson and Dallwitz (1992). All species of both tribes have a double bundle sheath around the veins, as is true in all C_3 grasses, with the outer sheath parenchymatous and the inner with thick walls. All species investigated have microhairs somewhere on the leaf surface; these are generally bicellular and elongate, as is common in most grasses, but Watson and Dallwitz (1992) report that the microhairs in some species appear to be unicellular. In all Oryzeae, except the aquatic genus *Chikusichloa*, mesophyll cells have deeply invaginated cell walls (arm cells), which are otherwise found only in bamboos, some early diverging grasses, and a handful of members of the PACCMAD clade; arm cells are lacking in Ehrharteae (Fig. 3). In addition, large open cells in the
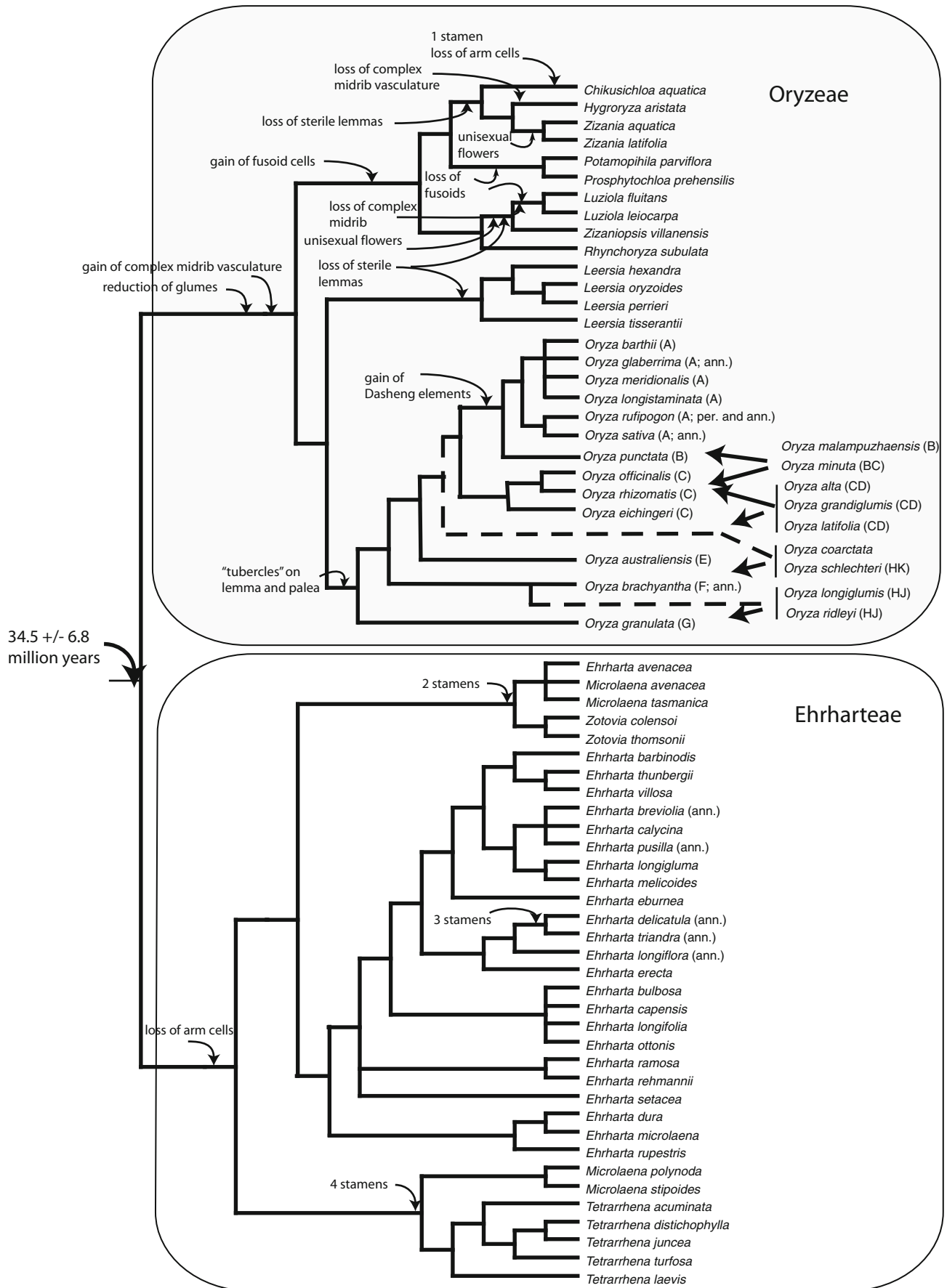
mesophyll (fusoid cells) occur in *Zizania*, *Zizaniopsis*, *Chikusichloa*, *Hygroryza*, and *Rhynchoryza* (Tateoka 1963). Fusoid cells are also known only in bamboos and in the early-diverging grasses (Grass Phylogeny Working Group 2001). The function of these cells is unknown. The midrib of Ehrharteae species is unremarkable, but midribs in most species of Oryzeae are complex, with two or more vascular bundles placed both adaxially and abaxially (i.e. top and bottom of the leaf), and often with air spaces (Tateoka 1963) (Fig. 3). However, *Luziola* and *Hygroryza* have a simple midrib vascular structure, more similar to that in Ehrharteae. Silica bodies are present in the epidermis, particularly in the costal region (over the veins); these are mostly dumbbell shaped and elongate either perpendicular to the axis of the leaf (oryzoid type) or parallel to the axis (dumbbell type).

Most Ehrhartoideae are perennials, and this appears to be the ancestral life history for the group. The annual habit has arisen independently among the A genome species of *Oryza*, in *Oryza brachyantha* (F genome), and at least twice in *Ehrharta* (Fig. 3). In an F_2 mapping population of *O. sativa* crossed with *Oryza longistaminata*, Hu et al. (2003) have identified multiple QTL that correlate with the presence of rhizomes, which are often correlated with the perennial habit. Two of these loci overlap with similar loci in sorghum suggesting that a few major regulators control growth form in many species of grasses. Because the switch from perennial to annual appears to occur frequently in evolutionary time, it will be interesting to learn if it reflects selection on a particular small set of loci.

Like rice, most members of Ehrhartoideae have a chromosome base number of $x=12$. The only exceptions are in the genera *Zizania* and possibly *Microlaena*. Members of *Zizania* have $2n=30$ or 34, numbers that appear to reflect expansion of the genome and some segmental duplications. A genome map of *Zizania palustris* shows that the base number of $x=15$ reflects a duplication of portions of rice chromosomes 1, 4, and 9 (Kennard et al. 2000). As in most other groups of grasses, polyploidy is common. Although Watson and Dallwitz (1992) report a chromosome base number of 10 for some species of *Microlaena*, recent data find that *Microlaena avenacea*, *Microlaena carsei*, *Microlaena polynoda*, and *Microlaena stipoides* have $2n=48$ chromosomes, so are tetraploids with a base number of $x=12$ (Murray et al. 2005).

Members of Ehrhartoideae are widely distributed, but most species occur in the Old World. Mapping geographic distribution data on the phylogeny of the diploid species in Fig. 3, using a parsimony algorithm in Mesquite (Maddison and Maddison 1997), yields the pattern illustrated in Fig. 4. As is common in such analyses, estimates of ancestral distributions are more ambiguous at deeper (i.e. older) nodes in the phylogeny. Thus, the common ancestor of the

subfamily may have occurred either in Australia (defined here to include New Zealand and Tasmania, for simplicity) or in tropical Africa. The common ancestor of Ehrharteae is inferred to have been Australian and the common ancestor of Oryzeae either African or Australian.

The common ancestor of *Oryza* and *Leersia* occurred in tropical Africa, and species in other parts of the world arrived there by independent dispersals. For example, the ancestor of *Oryza granulata* dispersed to India and China from Africa. A separate west-to-east dispersal resulted in the arrival of *Oryza eichingeri* and *Oryza rhizomatis* in Sri Lanka, and *Oryza officinalis* in India and eastern Asia. A third such event led to dispersal of *Oryza rufipogon* to India and eastern Asia. *Oryza meridionalis* and *Oryza australiensis* represent two independent dispersal events to Australia. Among other members of the tribe, *Prosphytochloa* and *Potamophila* represent another African-Australian pair. The phylogeny does not permit assessment of the direction of dispersal, but west to east is certainly plausible. Wide disjunctions have been documented for many aquatic plants (Les et al. 2003), so the broad distribution of many Oryzeae is not a surprise.

The phylogeny suggests dispersal from Australia to southern Africa in Ehrharteae. However, this could be an analytical artifact. The tree presented here is a grafted tree, created by simply pasting the phylogeny of Verboom et al. (2003) on to that of Guo and Ge (2005). If the Ehrharteae subclade were in fact rooted along the internode leading to the Cape Province *Ehrharta* species, then the ambiguity of the ancestral node would be removed and an African origin would be inferred. In addition, the parsimony optimization used here does not include any estimation of branch lengths, and thus does not take in to account the time available for long distance dispersal. Finally, all biogeographic analyses are painfully sensitive to taxon sampling and tree topology; inclusion of more species and/or minor changes in the phylogeny could easily affect the inference. A more rigorous analysis of the entire subfamily will be required to develop a better estimate of biogeographic history.

The phylogeny indicates at least three independent colonizations of North America: one each in *Leersia*, *Zizania*, and *Luziola*. It is likely that each occurred via a different route. For example, *Chikusichloa* occurs as far north as Japan, as does *Zizania*, suggesting that dispersal between Asia and North America may have occurred via a northern route. Conversely, *Luziola* is suggested to be the result of northward migration of South American species and *Leersia* represents a connection between Africa and North America. Taxon sampling in all three genera is poor outside of North America and many species have yet to be included in molecular phylogenetic analyses. Fortunately, a morphological phylogeny and monograph of the nine
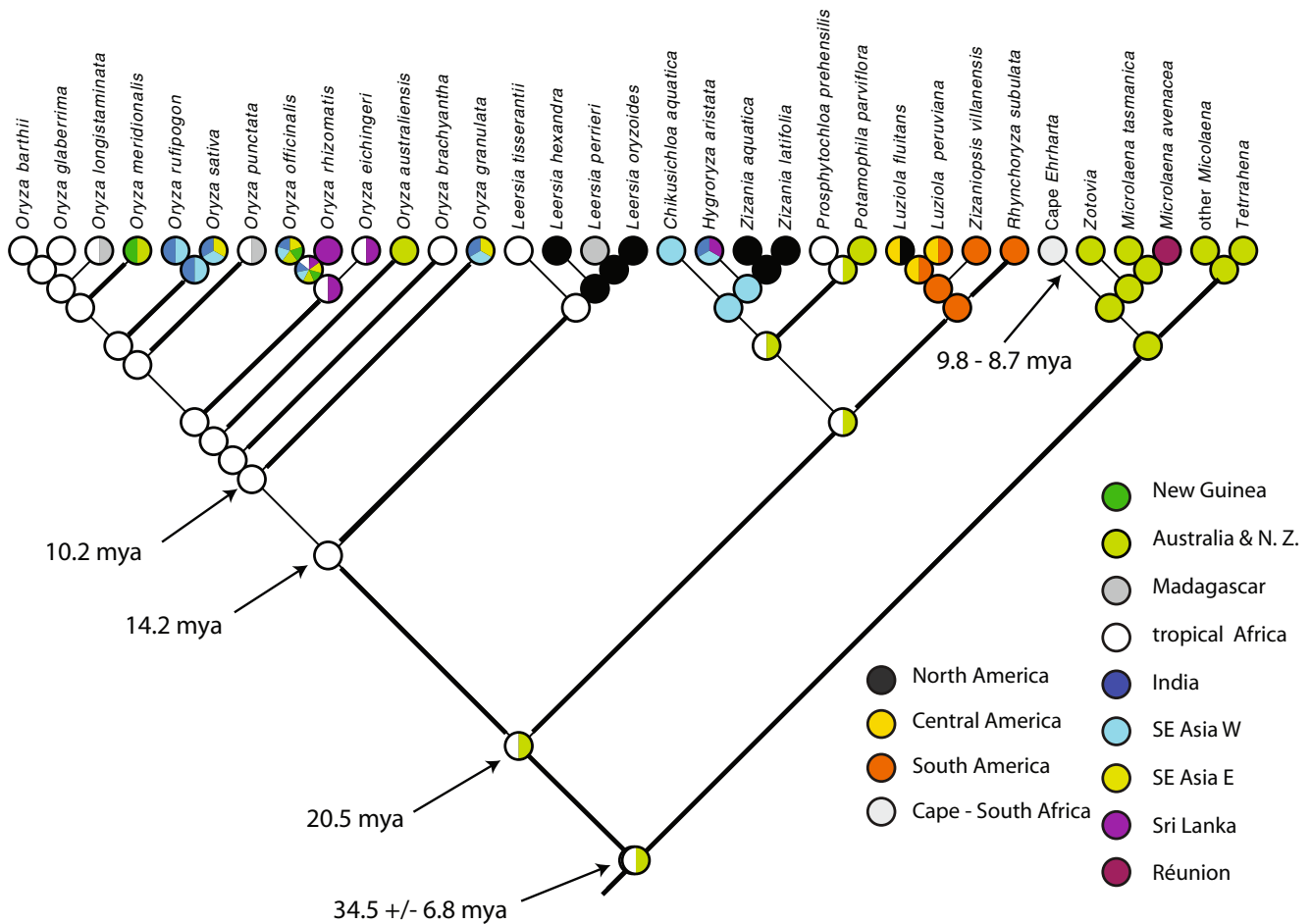
Fig. 4 Geographic distribution and approximate dates of divergence for members of Ehrhartoideae. Phylogeny as in Fig. 3, except that the Cape species of *Ehrharta* are represented as a single branch. *Luziola*
peruviana is synonymous with *L. leiocarpa*, following Martínez-y-Pérez et al. (2008).

species of *Luziola* have been published recently, setting the stage for molecular phylogenetic work (Martínez-y-Pérez et al. 2005, 2008). A full understanding of the history of Oryzeae will require investigation of all species within these widespread genera.

All evidence points to the ancestor of Ehrhartoideae being a plant of moist habitats. All members of Oryzeae grow in damp areas, many occupy flooded sites, and still others (e.g. *Hygroryza*) are truly aquatic. The ancestors of Ehrharteae are reconstructed as being plants of high rainfall areas (689–1,400 mm per annum) (Verboom et al. 2003). *Ehrharta* species later radiated in the much drier habitats of the fynbos of South Africa, probably during the late Miocene. This represents a shift from habitats with year-round rainfall to those with winter rainfall only. Many grasses are plants of moist shade (e.g. all members of Anomochlooideae, Pharoideae, and Puelioideae, many Bambusoideae) and many species grow at least partially submerged in water (e.g. *Glyceria* in subfamily Pooideae, *Phragmites* in Arundindoideae, *Spartina* in Chloridoideae,

to name just a few), so the Ehrhartoideae are not particularly unusual in this aspect of their biology.

The tribe Oryzeae

Oryzeae includes nine or ten genera and 60 to 70 species, and is very likely monophyletic (Ge et al. 2002; Guo and Ge 2005; Kellogg and Watson 1993). In all studies to date, the number of outgroup taxa included is not large enough to rule out completely the possibility of paraphyly, but available morphological and molecular data suggest that this is unlikely.

The phylogeny of Oryzeae has been addressed by Ge et al. (2002) and by Guo and Ge (2005). Both nuclear and chloroplast genes find that *Oryza* and *Leersia* are sister genera (Fig. 3). Sister to the *Oryza*/*Leersia* clade is a clade composed of the genera *Prosphytochloa*, *Potamophila*, *Chikusichloa*, *Rhynchoryza*, *Zizania*, *Zizaniopsis*, and *Luziola*. The position of *Hygroryza* is uncertain; it may be

sister to *Oryza*/*Leersia*, or sister to the clade comprised of *Rhynchoryza*, *Zizania*, *Zizaniopsis*, and *Luziola*. (Somewhat unhelpfully, species of *Zizania* are known in North America as wild rice, whereas in most other regions of the world the name "wild rice" generally refers to *O. rufipogon*.) The two major clades are recognized and named as subtribes (Oryzinae and Zizaniinae) by some authors (e.g. (Guo and Ge 2005)). However, the rank of subtribe is not required by the International Code of Botanical Nomenclature (Vienna Code) (2006), and formal subdivision seems unnecessary for a tribe with only a dozen genera.

The crown node of Orzyeae (divergence of *Oryza–Leersia* from the other genera) was placed at 20.5 mya by Guo and Ge (2005) and the divergence of *Oryza* from *Leersia* at 14.2 (Fig. 4). (Note that the methods used by Guo and Ge were different from those applied by Vicentini et al. (2008), and did not permit assessment of confidence intervals on dates.)

Several species in Oryzeae are so morphologically distinctive that previous taxonomists have given each its own genus. Molecular phylogenetic data, however, have placed these odd species in other genera. For example, *Hydrochloa caroliniensis* is the sole representative of the genus *Hydrochloa*, which is unusual in being a floating aquatic plant, with unisexual flowers (Fig. 3). The staminate and pistillate flowers are in separate inflorescences, with the staminate ones terminal on the plant and the pistillate ones axillary; the inflorescences are often reduced to a single spikelet (Watson and Dallwitz 1992). The inflorescence morphology of the genus *Luziola* is similar; the flowers are also unisexual and in separate inflorescences, which led Terrell and Robinson (1974) to suggest that *H. caroliniensis* is simply an odd species of *Luziola*. Accordingly, they moved *H. caroliniensis* to *Luziola* as *L. fluitans*. This classification is supported by molecular data (Guo and Ge 2005).

The taxonomic history of *Porteresia* is analogous to that of *Hydrochloa*, in that it is a name given to a slightly odd member of a genus. The species *Oryza coarctata* was originally described by Roxburgh in 1832 (Roxburgh 1832) (p. 206) based on a specimen collected in India in the Ganges Delta. Tateoka (1965) later concluded that *O. coarctata* was different from all other species of *Oryza* because it has leathery (coriaceous) leaf blades that are strongly ribbed. Curiously, each rib has a major vascular bundle near the abaxial side and a minor bundle just above (adaxial) to it. Tateoka (1964) also reported that the species has an unusually large embryo with a large free epiblast and a distinct cleft between the coleorhiza and the scutellum. (Embryo characters were relied on heavily in the pre-DNA era as evidence of phylogenetic relationship (see for example, Reeder 1957), and often do correlate with current

views of phylogeny (Grass Phylogeny Working Group 2001), but not always.) Accordingly, he created the new genus *Porteresia* to accommodate the species and to recognize the differences formally.

Molecular phylogenies based on the chloroplast genes *maturaseK* (*matK*) and the transfer RNA Leucine (*trnL*) intron, the mitochondrial gene NADH dehydrogenase (*nad1*), and the nuclear genes *alcohol dehydrogenase1* (*Adh1*), *alcohol dehydrogenase2* (*Adh2*), and G protein alpha subunit1 (GPA1) (Ge et al. 1999, 2002; Guo and Ge 2005) show that *P. coarctata* is derived from within *Oryza*. Therefore, the characteristics that Tateoka observed do not indicate a separate evolutionary origin for *O. coarctata*. The species was returned to its original position as a member of *Oryza* by Lu and Ge (2004).

Placing *O. coarctata* back in *Oryza* leaves open the question of the origin of its unusual morphology, habitat (wet saline sites), and distribution (coastal India). The species is tetraploid. It may share the H genome with other tetraploids such as *Oryza schlechteri* and *Oryza ridleyi* (Fig. 3). One molecular phylogenetic study suggests a close relationship with *O. schlechteri*, but the latter is native to New Guinea and has leaf morphology more similar to that of other *Oryza* species. The composition of repeat sequences in the genome of *O. coarctata* is also unusual (see below). The leaf morphology of *O. coarctata* is reminiscent of that of some species of *Ehrharta*, which have radiated into dry environments. Because saline environments also create physiological challenges similar to drought, it is perhaps not surprising that the leaf morphology of *O. coarctata* might have converged with that of more distantly related species.

The polyploid event that led to the formation of *O. coarctata* may provide a clue to its odd morphology and distribution. Polyploidy may have led to extensive genomic change that affected many aspects of the genome, including expression of genes that could be selected for a novel environment (Hegarty and Hiscock 2008). Polyploid species often have geographic ranges that are broader than or quite distinct from those of their diploid ancestors, so this would not be an unexpected outcome. A second more radical hypothesis is that *O. coarctata* is formed from a wide cross between an ancestor similar to rice and another ancestor more similar to another member of the tribe or subfamily. Such very wide crosses have been documented in wheat relatives (e.g. (Mason-Gamer 2004)) and are unusual but not unprecedented. Deep sequencing of the genome of *O. coarctata*, as well as broader phylogenetic studies in other Oryzeae, could help address this question.

Taxonomists have disagreed on the number of genera that should be recognized to accommodate the five species variously assigned to *Potamophila*, *Maltebrunia*, and *Prosphytochloa*. *Potamophila* was originally described

by Robert Brown (1810), *Maltebrunia* by Kunth (1829), and *Prosphytochloa* by Schweickerdt (1961). However, Duistermaat (1987) thought that this division was a mistake, given the morphological similarities among them. The molecular phylogeny of Guo and Ge (2005) supports the close relationship, at least between *Potamophila* and *Prosphytochloa*, and suggests that the two could be combined.

Monoecy has arisen twice in Oryzeae (Fig. 3). *Zizaniopsis* has staminate and pistillate flowers in separate inflorescences on the same plant, a characteristic that also appears in seven of the nine species of *Luziola* (including *Hydrochloa*). In the other two species of *Luziola* (*L. caespitosa* and *L. brasiliensis*), staminate and pistillate flowers are in the same inflorescence, with the staminate ones distal to the pistillate. *Zizania* also has staminate and pistillate flowers in the same inflorescence, but is somewhat peculiar among monoecious species in that the pistillate flowers are borne in the distal branches whereas the staminate ones are more proximal (lower).

Zaitchik et al. (2000) investigated the development of staminate and pistillate flowers in *Zizania*, and found that all florets initiated both an androecium and a gynoecium; in this respect, flowers of *Zizania* are similar to all other known unisexual flowers in the grasses (LeRoux and Kellogg 1999). Stamens in pistillate flowers ceased development soon after the outline of the anthers became visible. In staminate flowers, however, gynoecium development was arrested only after formation of the ovule and integuments and early development of the stylar arms. This is appreciably later than gynoecial abortion in maize or other panicoid grasses, and points to the possibility of a different underlying mechanism (Malcomber and Kellogg 2006).

Genera other than *Oryza* have received little attention from phylogeneticists, geneticists, or developmental biologists. Notably, *Zizania*, *Zizaniopsis*, *Leersia*, and *Luziola* all have broad geographic distributions and need to be investigated thoroughly throughout their ranges.

The genus *Oryza*

Oryza includes 20 to 24 species (the precise number depends on taxonomic preferences), distributed in tropical and subtropical regions of the world, and with a common ancestor dating between 10.2 and 8.8 mya (Guo and Ge 2005)(Fig. 4). An excellent web-based resource on the species of *Oryza* is provided by the International Rice Research Institute (IRRI), at http://www.knowledgebank. irri.org/wildRiceTaxonomy/default.htm. At this site are a key to the species, along with individual species descriptions, photos, and maps.

All species of *Oryza* share distinctive epidermal outgrowths or "tubercles," otherwise unknown among grasses, on the upper (fertile) lemma and palea (Terrell et al. 2001). The tubercles thus constitute a synapomorphy that provides evidence for monophyly of the genus. These outgrowths appear to be under the control of the SEPALLATA-like protein, LEAFY HULL STERILE1 (= OsMADS1); when *LHS1* is over-expressed, not only do the sterile lemmas enlarge and become more like the fertile lemma, but they produce epidermal tubercles as well (Prasad et al. 2001).

Molecular phylogenies show that each of the genomic groups in *Oryza* is monophyletic, as expected from cytogenetic studies (Ge et al. 1999; Zou et al. 2008). The diploid species with the A genome—*O. sativa*, *O. rufipogon*, *Oryza barthii*, *Oryza glaberrima*, *O. longistaminata*, and *O. meridionalis*—all form a clade, with *O. rufipogon* as the sister taxon and presumed wild progenitor of *O. sativa*. Likewise, the C genome diploids, *O. officinalis*, *O. rhizomatis*, and *O. eichingeri* are sisters. The B, E, F, and G genomes are each represented at the diploid level by a single species, *Oryza punctata*, *O. australiensis*, *O. brachyantha*, and *O. granulata* respectively.

A recent phylogenomic study of the diploid species (Zou et al. 2008) used 124,079 bp in 142 genes to confirm that the A and B genome species are sisters and the C genome clade is sister to that (Fig. 3). Some of the genes suggested alternative relationships, which the authors ascribed to incomplete lineage sorting among some of the genes. If this is true, then future population samples of the various species should yield other alleles that produce the "correct" phylogeny.

The polyploid species are allotetraploids and combine the genomes of two disparate progenitors. In a few cases, the progenitor genomes can be identified as being quite similar to those of extant diploids (e.g. *Oryza minuta* and tetraploid *O. officinalis* with the B and C genomes) (Fig. 3). (Note that some authors consider tetraploid *O. officinalis* to be a separate species, *Oryza malampuzhaensis* (Thomas et al. 2001).) In other cases, however, one or both of the progenitor genomes is extinct, or enough genomic rearrangement has occurred since polyploidization that similarities to extant diploids are lost. Thus, *Oryza alta*, *Oryza grandiglumis*, and *Oryza latifolia* contain the C and D genomes and are likely derived from an ancestor similar to one of the C genome diploids, but chromosome pairing data indicate that the D genome is distinct from known diploids, a result supported by genomic in situ hybridization (Li et al. 2001). However, phylogenetic analyses suggest that the D genome may be derived from an ancestor similar to the E genome of *O. australiensis* (Bao and Ge 2004).

Chloroplast phylogenies have helped to place the maternal parent of some of the polyploids. The chloroplasts

of the polyploids *O. alta*, *O. grandiglumis*, and *O. latifolia* are closely related to those of the C genome diploids (*O. officinalis*, *O. rhizomatis*, and *O. eichingeri*), indicating that the maternal parent of the CD polyploids had a C genome (Bao and Ge 2004). The chloroplasts in the polyploids *O. coarctata* and *O. schlecteri* (the latter with the H and K genomes) are closely related to each other, and share a recent common ancestor with the ancestral chloroplast of the A, B, and C genomes (Ge et al. 2002) (Fig. 3, dotted line). Likewise the chloroplast of *Oryza longiglumis* (HJ) appears closely related to that of *O. brachyantha* (F genome) (Ge et al. 2002) (Fig. 3, dotted line). Figure 3 places the other HJ polyploid, *O. ridleyi*, in this position as well, but this is by inference from *O. longiglumis*, and needs to be checked by DNA sequence data.

The genus *Oryza* has excellent genomic resources which are providing an unprecedented view into the highly dynamic nuclear genome (Kim et al. 2008). Recently genomic libraries and partial sequences have become available for many of them (the OMap project; http://www.omap.org).

Many lineages include characteristic retrotransposons; bursts of amplification of these transposons have led to considerable expansion of genome size. Members of *Oryza* share a group of closely related Ty3-gypsy retrotransposons, called the RWG family, whose common ancestor was at least as old as the genus itself (Ammiraju et al. 2007), although it is somewhat similar to the Grande elements of *Zea mays* (Ohtsubo et al. 1999). Although similar in their sequences, different lineages of the transposons have amplified independently in the various lineages of *Oryza*. Each of these independent amplifications has been given a separate name, such that the clade of transposons that occurs in *O. sativa* is known as RIRE2 (Ohtsubo et al. 1999), that in *O. australiensis* is Wallabi (Piegu et al. 2006), and that in *O. granulata* is called Gran3. In addition, a non-autonomous derivative of an RWG element, known as Dasheng, originated in the common ancestor of the A, B, and C genomes (Ammiraju et al. 2007) (Fig. 3).

Amplification of retrotransposons correlates with genome size. Among the diploid species, the largest genomes are in *O. officinalis* and *O. granulata* (Zuccolo et al. 2007). A project sequencing random sheared genomic libraries found that over 66% of the reads in *O. officinalis* involved repetitive sequences of some sort and the majority of these were retrotransposons. To make the data comparable among species, the authors divided the number of repeats by the total number of megabase pairs (Mbp) sequenced. With this calculation, *O. granulata* and *O. australiensis* (with large genomes) had the highest number of retrotransposons per Mbp, whereas *O. brachyantha* (with a small genome) had the lowest.

O. coarctata, despite being polyploid, has an unusually low percentage of repeat sequences, with few that are comparable to those in other species of *Oryza* (Zuccolo et al. 2007). Zuccolo et al. (2007) suggested that this might argue for recognition of the genus *Porteresia*. However, because genera are defined phylogenetically and because *O. coarctata* is clearly derived from within *Oryza*, it is preferable to consider *O. coarctata* as an unusual and intriguing member of the genus *Oryza*. Nonetheless, it would be of interest to know whether the unusual leaf morphology of the species were causally connected to the unusual lack of repeat elements in the genome. Because the leaf morphology of *O. coarctata* is similar to that of some species of *Ehrharta*, one might speculate that it represents a reversion to an ancestral type caused by a genome-scale change in gene expression.

All species of *Oryza* investigated to date have an array of genes encoding ribosomal RNA (rDNA, or nucleolar organizing region, or NOR) on chromosome 9, as observed with fluorescent in situ hybridization, although data are not available for the F and G genome species (Chung et al. 2008). Most taxa also have an rDNA array on chromosome 10. However, the chromosome 10 rDNA was not observed in *O. sativa* ssp. *japonica* cv. Nipponbare, in *O. glaberrima*, or in *O. australiensis*. Because of the phylogenetic distribution of the chromosome 10 array, the absence of the array in Nipponbare and *O. glaberrima* is most simply interpreted as a loss, whereas the lack of the array in *O. australiensis* could either be a loss peculiar to the *O. australiensis* lineage, or could be the ancestral condition. *Oryza punctata* (B genome) and its polyploid derivatives, *O. punctata* and *O. minuta* (BC) have an additional array on chromosome 4 and *O. officinalis* (C genome) and some of its derivatives have arrays on chromosome 5.

Early genome mapping studies in *Oryza sativa* identified a duplication of the distal part of chromosomes 11 and 12 (Rice Chromosomes 11 and 12 Sequencing Consortia 2005; Wu et al. 1998). The duplicated region is approximately 2 Mb in length and based on divergence between orthologous genes appears to have occurred approximately six mya (Jiang et al. 2007). If this date is correct, then the duplication is placed before the divergence of the A, B, and C genomes. The exact timing could be tested phylogenetically by assembling gene trees for genes shared in the duplicated segments. As expected with duplicate genes, the rate of molecular evolution is accelerated, particularly in the orthologues on chromosome 12 (Jiang et al. 2007). In addition, the rate of rearrangements is also elevated, again particularly on chromosome 12, such that inversions are more common than elsewhere in the genome.

Centromere evolution in *Oryza* is also active. Like all known centromeres, those of rice are made up of centromere-specific repeat sequences interspersed with a

variety of retrotransposons (Nagaki et al. 2004). The most common centromeric repeat in *Orzya* is known as CentO, which is similar to the corresponding sequence in maize. However, novel repeat sequences were found in the centromeres of *O. rhizomatis* (C genome) and *O. brachyantha* (F genome). The latter, called CentO-F, is strikingly different from all other known centromeric repeat sequences in the grasses (Lee et al. 2005). Curiously CentO-F has not only replaced most of the (presumably ancestral) CentO sequences, but also the repeated retrotransposons. In addition, the centromeric region of chromosome 8 in *O. sativa* appears to be inverted relative to *O. officinalis* and *O. brachyantha* (Ma et al. 2007)

Combined studies of molecular and genome evolution point to an unusual history of *O. brachyantha* (F genome). Phylogenetic studies place *O. brachyantha* at the end of a longer branch (Zou et al. 2008), indicating a higher mutation rate than in other species. In addition, members of the species have the smallest genome reported in the genus, which must reflect either a marked loss of retrotransposons, or failure of retrotransposons to amplify. The novel centromere repeat and loss of centromeric retrotransposons may reflect a single underlying genomic mechanism. *O. brachyantha* is also an annual plant, and hence has a short life cycle. It is tempting to speculate that all these observations might be related.

The species *O. sativa*

The history of *O. sativa* itself has been thoroughly reviewed elsewhere (Vaughan et al. 2008a, b), and will only be summarized here. *Oryza sativa* is comprised of two subspecies, *indica* and *japonica*, both of which are domesticated from a wild ancestor similar to *O. rufipogon*. Subspecies *japonica* is informally divided into tropical and temperate groups, the latter presumed to be derived from the former. Additional informal groups, such as aus, aromatic, ashwina, and rayada, are recognized and apply to plants with particular characteristics of a local environment.

An unresolved question in the history of *O. sativa* is the number of domestication events involved. Evidence has accumulated in the literature for various evolutionary histories, summarized in tabular form by Vaughan et al. (2008b).

If *O. sativa* were the result of a single domestication event, then all alleles in *O. sativa* should trace back to the alleles present in a single locality in *O. rufipogon*, producing a pattern similar to that in Fig. 5a. Conversely, if *indica* and *japonica* were domesticated independently, then the *indica* alleles should be derived from a set of alleles in one part of the range of *rufipogon* and the *japonica* alleles from a set in another part of the range, as
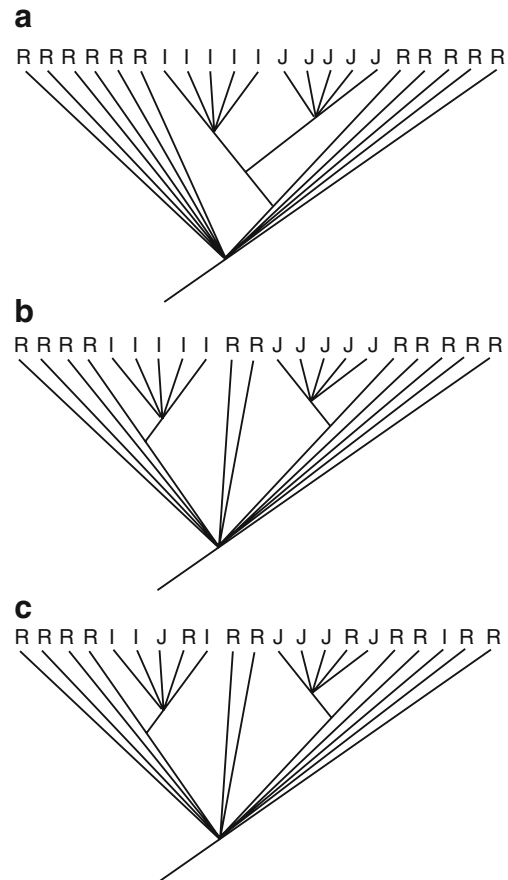

Fig. 5 Hypothetical phylogenetic patterns expected under different scenarios of rice domestication. **a** A single domestication event, followed by divergence of indica and japonica. **b** Two independent domestication events. **c** Two independent domestication events followed by gene flow. *I* alleles from *indica*, *J* alleles from *japonica*, *R* alleles from *rufipogon*.

shown in Fig. 5b. The distinction between the two patterns would disappear rapidly, however, if there is any recent or current gene flow between the groups, and if humans have selected for particular alleles. With gene flow and selection, the patterns will be more complex, with *indica*-like alleles appearing in groups otherwise composed of *japonica* alleles, and vice versa, and even *O. rufipogon* alleles entering the mix, as in Fig. 5c. The pattern also may be specific to particular genetic loci. For example, if a locus is under strong selection, it is quite possible that the selected locus would exhibit a pattern like that in Fig. 5a, even while non-selected loci would show patterns such as in Fig. 5b.

Methods for assessing the pattern of evolution include phylogenetic analysis of sequence data, as well as tests for selection. The latter tests compare the observed patterns of sequence variation with the patterns expected if mutation were purely random and no selection were occurring (a neutral model). If mutations occur randomly (i.e. none is

favored by selection), most sequence variants (haplotypes) will each appear in only a few individual plants. Conversely, a small number of haplotypes will each appear in a large number of plants. In other words, a survey of single nucleotide polymorphisms (SNPs) in a population will find a few that are very common and many that are rare. A deficiency of rare SNPs or an excess of common ones indicates that the population is not evolving neutrally. However, there are many possible reasons why a population might depart from neutral expectation.

A recent analysis of single nucleotide polymorphisms in material of *O. sativa indica*, *japonica*, aromatic, aus, and *O. rufipogon* found a pattern similar to that in Fig. 5b (Caicedo et al. 2007). In addition, too many loci had SNPs that were at a high frequency in *O. sativa*. The authors concluded that the pattern could be explained by a sharply reduced population size (bottleneck) during domestication, combined with a selective sweep. Alternatively, the pattern might indicate that the ancestral population had been subdivided, but that subsequent gene flow had occurred among the populations.

An analogous study using microsatellite loci found a pattern similar to that in Fig. 5a, with a single domestication origin and some apparent gene flow between the two subspecies of *O. sativa* (Gao and Innan 2008). The microsatellite study included fewer plants and a smaller number of markers than the sequence-based study (Caicedo et al. 2007). Nonetheless the discrepancies between the two studies indicate how difficult it may be to arrive at a definitive answer to the question of rice domestication.

All studies agree that domestication reduced the population size of *O. sativa* relative to that of *O. rufipogon*, as expected. Because domestication selects only a sample of plants from the population of the wild ancestor, genetic diversity is necessarily reduced. Beyond this, however, there is little agreement among the studies. It seems unlikely that the issue of one versus two origins of *O. sativa* will be firmly resolved with current analytical methods and data. In the presence of gene flow and selection, both of which are occurring, the two patterns in fact cannot be distinguished. (See also discussions in Vaughan et al. (2008a, b))

For a few types of domesticated rice, the history appears easier to assess. For example, fragrant rice is a result of a mutation at the *Badh2* locus (Chen et al. 2008), which encodes the gene for betain aldehyde dehydrogenase. In most rice varieties the gene product prevents formation of 2-acetyl-1pyrroline. Fragrant varieties, in contrast, contain a loss of function mutation in the gene, thus permitting synthesis of the volatile compound, which provides the aroma for which the rice is known. Because all fragrant varieties seem to contain the same mutant haplotype, fragrant rice is inferred to have arisen once.

Summary and conclusions

Like any living organism, rice carries with it the legacy of its evolutionary history. Some characteristics were acquired deep in the distant past of eukaryotic evolution, whereas other aspects were acquired more recently along with the common ancestors of other angiosperms, other monocots, or other grasses. An understanding of evolutionary history permits us to assess the generality of aspects of rice biology and conversely to understand which aspects of other plants might be comparable. For example, the pathways that permit accumulation of ferulic acid in the cell wall are likely common to all the commelinid monocots, so that data from rice could reasonably be extrapolated to cell wall structure in pineapples or bananas, and conversely. On the other hand, formation of tubercles (outgrowths) on the lemma and palea (hull) is a characteristic unique to the genus *Oryza*; an understanding of this aspect of hull architecture can be extrapolated to all congeneric species, but not to other Oryzeae or other cereal crops.

Use of rice as a model for other grass species or for other monocots requires extensive comparative study. Only by determining the taxonomic distribution of species that share particular characteristics will we be able to translate information from rice to other plants, both wild and cultivated. Such comparative studies can be guided by the extensive phylogenetic information now available, and will permit targeted investigation of plants likely to illuminate the generality of results from studies of rice.

References

Ambrose BA, Lerner DR, Ciceri P, Padilla CM, Yanofsky MF, Schmidt RJ. Molecular and genetic analyses of the *Silky1* gene reveal conservation in floral organ specification between eudicots and monocots. Mol Cell 2000;5:569–79.

Ammiraju JSS, Zuccolo A, Yu Y, Song X, Piegu B, Chevalier F, et al. Evolutionary dynamics of an ancient retrotransposon family provides insights into evolution of genome size in the genus *Oryza*. Plant J 2007;52:342–51.

Arber A. The Gramineae: A study of cereal, bamboo, and grass. Cambridge: The University Press; 1934.

Bao Y, Ge S. Origin and phylogeny of *Oryza* species with the CD genome based on multiple-gene sequence data. Plant Syst Evol 2004;249:55–66.

Bommert P, Satoh-Nagasawa N, Jackson D, Hirano HY. Genetics and evolution of inflorescence and flower development in grasses. Plant Cell Physiol 2005;46:69–78.

Bouchenak-Khelladi Y, Salamin N, Savolainen V, Forest F, van der Bank M, Chase MW, et al. Large multi-gene phylogenetic trees of the grasses (Poaceae): progress towards complete tribal and generic taxon sampling. Mol Phylog Evol 2008;47:488–505.

Brown R. Prodromus Flora Novae Hollandiae et insulae van Diemen. London: Richard Taylor; 1810.

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fiedel-Alon A, York TL, et al. Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS Genet 2007;3:1745–56.

Campbell CS, Kellogg EA. Sister group relationships of the Poaceae. In: Soderstrom TR, Hilu KW, Campbell CS, Barkworth ME, editors. Grass systematics and evolution. Washington, DC: Smithsonian Institution; 1987. p. 217–24.

Chen S, Yang Y, Shi W, Ji Q, He F, Zhang Z, et al. *Badh2*, encoding betaine aldehyde dehydrogenase, inhibits the biosynthesis of 2-acetyl-1-pyrroline, a major compnent in rice fragrance. Plant Cell 2008;20:1850–61.

Chung M-C, Lee Y-I, Cheng Y-Y, Chou Y-J, Lu C-F. Chromosomal polymorphism of ribosomal genes in the genus *Oryza*. Theor Appl Genet 2008;116:745–53.

Clark LG, Zhang W, Wendel JF. A phylogeny of the grass family (Poaceae) based on *ndhF* sequence data. Syst Bot 1995;20:436–60.

Clayton WD, Renvoize SA. Genera graminum. London: Her Majesty's Stationery Office; 1986.

Duistermaat H. A revision of *Oryza* (Gramineae) in Malesia and Australia. Blumea 1987;32:157–93.

Gao L, Innan H. Nonindependent domestication of the two rice subspecies, *Oryza sativa* ssp. *indica* and ssp. *japonica*, demonstrated by multilocus microsatellites. Genetics 2008;179:965–76.

Gaut BS. Evolutionary dynamics of grass genomes. New Phytol 2002;154:15–28.

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 U S A 1999;96:14400–5.

Ge S, Li A, Lu B-R, Zhang S-Z, Hong D-Y. A phylogeny of the rice tribe Oryzeae (Poaceae) based on *matK* sequence data. Amer J Bot 2002;89:1967–72.

Grass Phylogeny Working Group. Phylogeny and subfamlial classification of the Poaceae. Ann Missouri Bot Gard 2001;88:373–457.

Guo Y-L, Ge S. Molecular phylogeny of Oryzeae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genome. Amer J Bot 2005;92:1548–58.

Hegarty MJ, Hiscock SJ. Genomic clues to the evolutionary success of polyploid plants. Curr Biol 2008;18:R435–44.

Hu FY, Tao DY, Sacks E, Fu BY, Xu P, Li J, et al. Convergent evolution of perenniality in rice and sorghum. Proc Natl Acad Sci U S A 2003;100:4050–4.

International Code of Botanical Nomenclature (Vienna Code). 2006;146:A. R. G. Gantner KG.

Janssen T, Bremer K. The age of major monocot groups inferred from 800+ *rbcL* sequences. Bot J Linn Soc 2004;146:385–98.

Jiang H, Liu D, Gu Z, Wang W. Rapid evolution in a pair of recent duplicate segments of rice. J Exp Zool (Mol Dev Evol) 2007;308B:50–7.

Kellogg EA. The grasses: a case study in macroevolution. Ann Rev Ecol Syst 2000;31:217–38.

Kellogg EA, Linder HP. Phylogeny of Poales. In: Rudall PJ, Cribb PJ, Cutler DF, Humphries CJ, editors. Monocotyledons: systematics and evolution. Kew: Royal Botanic Gardens; 1995. p. 511–42.

Kellogg EA, Watson L. Phylogenetic studies of a large data set. I. Bambusoideae, Andropogonodae, and Pooideae. Bot Rev 1993;59:273–343.

Kennard WC, Phillips RL, Porter RA, Grombacher AW. A comparative map of wild rice (*Zizania palustris* L. $2n=2x=30$).. Theor Appl Genet 2000;101:677–84.

Kim H, Hurwitz B, Yu Y, Collura K, Gill N, SanMiguel P, et al. Construction, alignment and analysis of twelve framework physical maps that represent the ten genome types of the genus *Oryza*. Genome Biol 2008;9:R45.

Komatsu M, Chujo A, Nagato Y, Shimamoto K, Kyozuka J. FRIZZY PANICLE is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. Development 2003;130:3841–50.

Kunth KS. Révision des Graminées 1:6. Paris: Librairie-Gide; 1829.

Lee H-R, Zhang W, Langdon T, Jin W, Yan H, Cheng Z, et al. Chromatin immunoprecipitation cloning reveals rapid evolutionary patterns of centromeric DNA in *Oryza* species. Proc Natl Acad Sci U S A 2005;102:11793–8.

LeRoux LG, Kellogg EA. Floral development and the formation of unisexual spikelets in the Andropogoneae (Poaceae). Amer J Bot 1999;86:354–66.

Les DH, Crawford DJ, Kimball RT, Moody ML, Landolt E. Biogeography of discontinuously distributed hydrophytes: a molecular appraisal of intercontinental disjunctions. Int J Plant Sci 2003;164:917–32.

Lescot M, Piffanelli P, Ciampi AY, Ruiz M, Blanc G, Leebens-Mack J, et al. Insights into the *Musa* genome: syntenic relationships to rice and between *Musa* species. BMC Genomics 2008;9:58.

Li C-B, Zhang D-M, Ge S, Lu B-R, Hong D-Y. Differentiation and inter-genomic relationships among C, E, and D genomes in the *Oryza officinalis* complex (Poaceae) as revealed by multicolor genomic in situ hybridization. Theor Appl Genet 2001;103:197–203.

Lin H, Ouyang S, Egan A, Nobuta K, Haas B, Zhu W, Gu X, Silva JC, Meyers BC, Buell CR. Characterization of paralogous protein families in rice. BMC Plant Biol 2008;8:18.

Litke R. Uber den Nachweis tertiaer Gramineen. Monats Deutschen Akad Wissenschaft Berlin 1968;10:462–71.

Lu B-R, Ge S. *Oryza coarctata*: the name that best reflects the relationships of *Porteresia coarctata* (Poaceae: Oryzeae). Nordic J Bot 2004;23:555–8.

Ma J, Wing RA, Bennetzen JL, Jackson SA. Evolutionary history and positional shift of a rice centromere. Genetics 2007;177:1217–20.

Maddison DR, Maddison WP. MacClade 4: Analysis of phylogeny and character evolution. Release version 4.08. Massachusetts: Sinauer Associates, Sunderland; 2005.

Maddison W, Maddison D (1997–2008) Mesquite version 2.5 (build 77). http://mesquiteproject.org

Malcomber ST, Kellogg EA. Evolution of unisexual flowers in grasses (Poaceae) and the putative sex-determination gene, *TASSELSEED2* (*TS2*). New Phytol 2006;170:885–99.

Malcomber ST, Preston JC, Reinheimer R, Kossuth J, Kellogg EA. Developmental gene evolution and the origin of grass inflorescence diversity. Adv Bot Res 2006;44:425–81.

Marchant AD, Briggs BG. Ecdeiocoleaceae and Joinvilleaceae, sisters of Poaceae (Poales): evidence from *rbcL* and *matK* data. Telopea 2007;11:437–50.

Martínez-y-Pérez JL, Mejía-Saulés T, Sosa V. Phylogenetic relationships of *Luziola* (Poaceae: Oryzeae) and related genera from aquatic habitats. Can J Bot 2005;84:1753–70.

Martínez-y-Pérez JL, Mejía-Saulés T, Sosa V. A taxonomic revision of *Luziola* (Poaceae: Oryzeae). Syst Bot 2008;33:702–18.

Mason-Gamer RJ. Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. Syst Biol 2004;53:25–37.

Murray BG, DeLange PJ, Ferguson AR. Nuclear DNA variation, chromosome numbers and polyploidy in the endemic and indigenous grass flora of New Zealand. Ann Bot 2005;96:1293–305.

Nagaki K, Cheng ZK, Ouyang S, Talbert PB, Kim M, Jones KM, et al. Sequencing of a rice centromere uncovers active genes. Nature Genet 2004;36:138–45.

Ohtsubo H, Kumekawa N, Ohtsubo E. RIRE2, a novel gypsy-type retrotransposon form rice. Genes Genet Syst 1999;74:83–91.

Paterson AH, Bowers JE, Chapman BA. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci U S A 2004;101:9903–8.

Piegu B, Guyot R, Picault N, Roulin A, Saniyal A, Kim H, et al. Doubling genome size without polyploidization: Dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. Genome Res 2006;16: 1262–9.

Prasad K, Sriram P, Kumar CS, Kushalappa K, Vijayraghavan U. Ectopic expression of rice *OsMADS1* reveals a role in specifying the lemma and palea, grass floral organs analogous to sepals. Devel Genet Evol 2001;211:281–90.

Preston JC, Kellogg EA. Conservation and divergence of *APETALA1/ FRUITFULL-like* gene function in grasses: evidence from gene expression analyses. Plant J 2007;52:69–81.

Reeder JR. The embryo in grass systematics. Amer J Bot 1957;44:756–68.

Rice Chromosomes 11 and 12 Sequencing Consortia. The sequence of rice chromosomes 11 and 12, rich in disease resistance genes and recent gene duplications. BMC Biol 2005;3:20.

Roxburgh W. Flora indica; or, descriptions of Indian Plants. London: Serampore, Calcutta; 1832.

Rudall PJ, Stuppy W, Cunniff J, Kellogg EA, Briggs BG. Evolution of reproductive structures in grasses inferred by sister-group comparisons with their putative closest living relatives. Amer J Bot 2005;92:1432–43.

Sánchez-Ken JG, Clark LG, Kellogg EA, Kay EE. Reinstatement and emendation of subfamily Micrairoideae. Syst Bot 2007;32:71–80.

Schweickerdt HG. Graminearum species nova et genus novum. Zur Kenntnis einiger südafrikanischer gramineen. Züchter 1961;31: 193–5.

Stapf O, (ed). Gramineae. Flora of tropical Africa. London: Lowell Reeve & Co.; 1917.

Stebbins GL. Coevolution of grasses and herbivores. Ann Missouri Bot Gard 1981;68:75–86.

Stevens PF. Angiosperm phylogeny website. Version 9, June 2008 [and more or less continuously updated since]. 2001 onwards. http://www.mobot.org/MOBOT/research/APweb/.

Tateoka T. Notes on some grasses. XIII. Relationship between Oryzeae and Ehrharteae, with special reference to leaf anatomy and histology. Bot Gaz 1963;124:264–70.

Tateoka T. Notes on some grasses. XVI. Embryo structure of the genus *Oryza* in relation to the systematics. Amer J Bot 1964;51:539–43.

Tateoka T. *Porteresia*, a new genus of Graminae. Bull Nat Sci Mus, Tokyo 1965;8:405–6.

Terrell EE, Robinson H. Luziolinae, a new subtribe of oryzoid grasses. Bull Torr Bot Club 1974;101:235–45.

Terrell EE, Peterson PM, Wergin WP. Epidermal features and spikelet micromorphology in *Oryza* and related genera (Poaceae: Oryzeae). Smithsonian Contr Bot 2001;91:1–50.

Thomas G, Sreejayan, Joseph L, Kuriachan P. Genetic variation and population structure in *Oryza malampuzhaensis* Krish. et Chand. endemic to Weatern Ghats, South India. J Genet 2001;80:141–8.

Thomasson JR. Fossil grasses: 1820–1986 and beyond. In: Soderstrom TR, Hilu KW, Campbell CS, Barkworth ME, editors. Grass systematics and evolution. Washington: Smithsonian Institution; 1987.

Vaughan DA, Lu B-R, Tomooka N. The evolving story of rice evolution. Plant Sci 2008a;174:394–408.

Vaughan DA, Lu B-R, Tomooka N. Was Asian rice (*Oryza sativa*) domesticated more than once? Rice 2008b;1:16–24.

Verboom GA, Linder HP, Stock WD. Phylogenetics of the grass genus *Ehrharta*: evidence for radiation in the summer-arid zone of the South African Cape. Evolution 2003;57:1008–21.

Vicentini A, Barber JC, Aliscioni SA, Giussani LM, Kellogg EA. The age of the grasses and clusters of origins of C_4 photosynthesis. Global Change Biol 2008;14:2963–77.

Wang X, Shi X, Hao B, Ge S, Luo J. Duplication and DNA segmental loss in the rice genome: implications for diploidization. New Phytol 2005;165:937–46.

Watson L, Dallwitz MJ. Grass genera of the world. Wallingford: CAB International; 1992.

Whipple CJ, Ciceri P, Padilla CM, Ambrose BA, Bandong SL, Schmidt RJ. Conservation of B-class floral homeotic gene function between maize and Arabidopsis. Development 2004;131:6083–91.

Whipple CJ, Zanis M, Kellogg EA, Schmidt RJ. Conservation of B class MADS-box gene expression in the second whorl of a basal grass and outgroups links the origin of lodicules and petals. Proc Natl Acad Sci U S A 2007;104:1081–6.

Willemse LPM. A discussion of the Ehrharteae (Gramineae) with special reference to the Malesian taxa formerly included in *Microlaena*. Blumea 1982;28:181–94.

Wu J, Kurata N, Tanoue H, Shimokawa T, Umehara Y, Yano M, et al. Physical mapping of duplicated genomic regions of two chromosome ends in rice. Genetics 1998;150:1595–603.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 2001;292:686–93.

Zaitchik BF, leRoux LG, Kellogg EA. Development of male flowers in *Zizania aquatica* (North American wild-rice; Gramineae). Int J Plant Sci 2000;161:345–51.

Zou X-H, Zhang F-M, Zhang J-G, Zang L-L, Tang L, Wang J, et al. Analysis of 142 genes resolves the rapid diversification of the rice genus. Genome Biol 2008;9:R49.

Zuccolo A, Sebastian A, Talag J, Yu Y, Kim H, Collura K, et al. Transposable element distribution, abundance and role in genome size variation in the genus *Oryza*. BMC Evol Biol 2007;7:152.