Paleoecology, Ploidy, Paleoatmospheric Composition, and Developmental Biology: A Review of the Multiple Uses of Fossil Stomata1[OPEN]

Jennifer C. McElwain* and Margret Steinthorsdottir

Earth Institute, O’Brien Centre for Science, and School of Biology and Environmental Science, University College Dublin, Belfield, Dublin 4, Ireland (J.C.M.); Department of Palaeobiology, Swedish Museum of Natural History, SE-104 05 Stockholm, Sweden (M.S.); and Bolin Centre for Climate Research, Stockholm University, SE-104 05 Stockholm, Sweden (M.S.)

ORCID IDs: 0000-0002-1729-6755 (J.C.M.); 0000-0002-7893-1142 (M.S.).

The presence of stomata is a diagnostic trait of all living and extinct land plants with the exception of liverworts. They are preserved widely in the fossil record from anatomically pristine stomatal complexes on permineralized and charcoalfied stems of the earliest land plants dating back >400 million years to isolated guard cell pairs in quaternary aged palynological samples. Detailed study of fossil stomatal complexes has been used to track the evolution of genome size and to reconstruct atmospheric composition, to circumscribe new species to science, and to bring ancient landscapes to life by providing both habitat information and insights on fossil plant ecophysiological function and life form. This review explores how fossil stomata can be used to advance our understanding of plant, environment, and atmospheric evolution over the Phanerozoic. We compare the utility of qualitative (e.g. presence/absence of stomatal structures) versus quantitative stomatal traits (e.g. amphistomaty ratio) in paleoecological reconstructions. A case study on Triassic-Jurassic Ginkgoales is provided to highlight the methodological difficulty of teasing apart the effect of genome size, ploidy, and environment on guard cell size evolution across mass extinction boundaries. We critique both empirical and mechanistic stomatal-based models for paleoCO2 reconstruction and highlight some key limitations and advantages of both approaches. Finally, we question if different stomatal developmental pathways have ecophysiological consequence for leaf gas exchange and ultimately the application of different stomatal-based CO2 proxy methods. We conclude that most studies currently only capture a fraction of the potential invaluable information that can be gleaned from fossilized stomata and highlight future approaches to their study that better integrate across the disciplinary boundaries of paleobotany, developmental biology, paleoecology, and plant physiology.

The fossil record of land plants (embryophytes) dates back unequivocally to the Middle Ordovician (~460 million years ago [mya]). This is supported by the presence of spore tetrads contained within an enveloping sporangium (Wellman et al., 2003). Since Wellman’s discovery, the fossil spore record has revealed older and older spores of various morphologies (naked, enveloped) and configurations (singular, paired, etc.) that may eventually push back even further the accepted date of the oldest land plant (Wellman and Strother, 2015). This early phase in land plant evolution is complex to interpret, however, since no stomata have been discovered so far on the earliest fossilized land plants, suggesting perhaps that they may have been absent, as is the case for the early land plants’ algal predecessors. As soon as sheets of fossilized cuticle with true stomata started to appear in Silurian aged (443–419 mya) sediment samples, our ability to taxonomically separate charophyacean algae from land plants based on fragmentary fossil evidence improved greatly because the presence of stomata is the defining anatomical trait of all living and extinct land plant sporophytes with the exception of liverworts. Therefore, it is unsurprising that the use of fossilized stomata for taxonomic purposes and to elucidate the phylogeny of land plants as revealed by the fossil record has a long history in paleobotany.

Stomatal traits that are considered of utility for fossil plant taxonomy and systematics are numerous, including stomatal presence or absence, size, geometry and orientation, and association with subsidiary cells (Table I), whether they are sunken, raised, or flush with epidermal cells or plugged with wax, are kidney or dumbbell shaped, are overarched by papillate subsidiary/epidermal cells, or

1 This work was supported by a Science Foundation Ireland Principal Investigator Award 11/PI/1103, by European Research Council Award ERC-2011-StG 279962 to J.C.M., and by Swedish Research Council Starting Grant (VR NT-7 2016 04905) and the Bolin Centre Award ERC-2011-StG 279962 to J.C.M., and by Swedish Research Council Principal Investigator Award 11/PI/1103, by European Research Council.
are completely encircled by a ring of fused subsidiaries (Cleal and Zodrow, 1989; Hill and Pole, 1992; Carpenter and Jordan, 1997; Denk and Velitzelos, 2002; Krings et al., 2003; Carpenter et al., 2005; Kerp et al., 2006; Cleal, 2008; Pole, 2008; Hernandez-Castillo et al., 2009; Pott and McLoughlin, 2009; Bomfleur and Kerp, 2010; Cleal and Shute, 2012). Guard cell lignification (Lacourse et al., 2016), striations (Barclay et al., 2007), and the presence of two size classes of stomata, including giant stomata (Fišer Pecník et al., 2012), have also been examined for taxonomic purposes. The problem of using fossil stomatal traits in taxonomy and systematics is that some traits “show some genetically uncontrolled variability” (Baranova, 1992) with specific traits showing greater variability than others (Barclay et al., 2007; Cleal and Shute, 2012; Jordan et al., 2014; Lacourse et al., 2016).

Over the past three decades, since publication of Woodward’s (Woodward, 1987) seminal article on the inverse relationship between stomatal density and atmospheric CO₂ concentration, the “uncontrolled variability” in stomatal traits such as stomatal density (SD) and stomatal index (SI) has been seized on by paleobiologists as an opportunity to extract meaningful paleoclimatic and paleoenvironmental information from fossil plant stomata (McElwain et al., 2005; Roth-Nebelsick, 2005; Wagner et al., 2005; Küsters et al., 2008; Lammertsma et al., 2011; Steinthorsdottir et al., 2011b; Franks et al., 2014; Maxbauer et al., 2014; Bai et al., 2015; Montañez et al., 2016; Steinthorsdottir et al., 2016a, 2016b). This has led to a subtle tension in the field of paleobotany, where at one extreme some studies have focused almost exclusively on the taxonomic and systematic utility of stomata with insufficient consideration of environmentally driven variability, while at the other extreme some reconstructions of paleoatmospheric composition have been undertaken using fossil stomatal traits without due consideration for taxonomic determination of the fossils used. The aim of this article is to briefly review both long-established and novel uses of fossil stomata, including their use, and (1) to infer the paleoecology of ancient landscapes from fossil plant assemblages, (2) to elucidate the genomic history of embryophytes, (3) to reconstruct paleoatmospheric trends in carbon dioxide concentration (pCO₂) over the past ~400 mya, and (4) to gain insights into plant developmental biology, in particular that of the leaf. We provide a case study demonstrating how genome size and ploidy can be estimated from fossil guard cell size of Triassic-Jurassic Ginkgoales. Furthermore, we highlight two paleo-pCO₂ cross-calibration examples for the Late Pennsylvanian-Early Permian (311–296 mya) and Late Triassic-Early Jurassic (~209–199 mya), where pCO₂ estimates are derived by applying different stomatal models and calibration approaches to the same stomatal datasets. A secondary, but equally important, aim of this review is to highlight potential future avenues for synthesis of these myriad uses of fossil stomata in order to maximize the genotypic and phenotypic information that can be gathered from their frequency, size, geometry, distribution, developmental pathway, and association with neighboring and/or subsidiary cells.

### PALEOECOLOGY AND FOSSIL STOMATA

Paleobotanists as a community use all tools available to reconstruct the climate and ecology of ancient landscapes from fossil plant assemblages. Briefly, these tools can include sedimentological information (Farrish et al., 1982); assessment of fossil taxa co-occurrence (Bowman et al., 2014); behavioral information such as leaf herbivory damage (Carvalho et al., 2014); census information on paleodiversity, relative abundance, density, and fossil plant preservation type (Cúneo et al., 2003; Wilf et al., 2005;
Mander et al., 2013; Willis and McElwain, 2014; Falcon-Lang, 2015), factors that could bias fossilization potential (Mander et al., 2012); and of course organ, tissue (Royer et al., 2007; Belcher et al., 2010), and cellular level traits (Haworth and McElwain, 2008) that provide unique ecological information.

The extraction of precise paleoecological information from fossil plant stomata and their associated neighboring and subsidiary cells is not however without its pitfalls. Some of the most enduring traits of the stomatal complex considered to have paleoecological value include the presence of papillae on neighboring or subsidiary cells that overarch the stomatal aperture, sunken guard cells, and the presence of deep stomatal crypts. Traditionally, all of these modifications to the stomatal complex have been interpreted functionally as anatomical adaptations to aridity, since they reduce stomatal conductance and leaf level transpiration (Hill, 1998; Krings et al., 2003). Haworth and McElwain (2008) directly challenged this narrow interpretation, arguing that many of these morphological traits also occur in very humid and wet environments, with a function to repel liquid water from the leaf surface, as well as in self-cleaning of atmospheric particulates, spores and other irritants (Barthlott et al., 2017). They cautioned that an antitranspirant function of papillae and sunken stomata must be interpreted in conjunction with other paleoecologically significant information. A detailed stomatal trait-climate study on Proteaceae further confirmed that while deep stomatal crypts are systematically associated with climatic aridity, the presence of sunken stomata and papillae on cells of the stomatal complex are not uniquely associated with ecological aridity (Jordan et al., 2008). The living fossil, *Sciadopitys* (Japanese umbrella-pine), for example, possesses a deep stomatal crypt yet it occurs in modern day wet warm temperate forest.

An additional stomatal trait that is sometimes used in paleoecological studies is stomatal distribution (Bomfleur and Kerp, 2010; Cleal and Shute, 2012). Amphistomaty, where stomata occur on both leaf surfaces, is often associated with high irradiance (Fitter and Peat, 1994), high elevation (Woodward, 1986), as well as open aquatic or open desert environments (Mott et al., 1982). It facilitates increased stomatal and mesophyll conductance and is usually associated with higher photosynthetic rates (Mott et al., 1982). In Proteaceae, amphistomaty is strongly associated with open vegetation (Jordan et al., 2014), leading to the suggestion that amphistomaty may be a good proxy for an open habitat. However, Muir (2015) argues that the adaptive significance of this relatively rare stomatal distribution (<10% of modern global flora are amphistomatic) relates most strongly to life history traits such as relative growth rate. Annuals, biennials, and perennials were shown to have a much greater frequency of species that were amphistomatous than shrubs and trees across 599 species and 94 families (Muir, 2015).

Therefore, it appears that simple binary data regarding the presence or absence of particular stomatal traits, including crypts, sunken stomata, papillate overarching subsidiary cells, etc., cannot be used by themselves to interpret paleoecology across all phylogenetic groups but may have great utility when applied in phylogenetically restricted studies, for example, in Proteaceae (Jordan et al., 2014). However, other aspects of stomatal development, such as amphistomaty or specifically the ratio of abaxial to adaxial stomatal density *sensu* Muir, 2015, look more promising for broader application to the fossil record to interpret paleoecology, including habit and habitat. Herbaceous taxa are believed to be very poorly represented in the fossil record yet many fossil plant assemblages over the past 400 million years are characterized by an unusually high proportion of amphistomatous taxa compared with the modern flora (e.g. *Dicroidium* dominated floras of the Late Triassic; Bomfleur and Kerp, 2010). Were paleo-amphistomatous taxa fast growing herbaceous plants or woody shrubs...
and trees from open habitats? Extant herbaceous species with amphistomatic leaves are typically dorsiventral with differentiated mesophyll tissue layers, whereas extant sclerophyllous species of arid environments with amphistomatic leaves are typically isobilateral with undifferentiated mesophyll. Permineralized or charcoalified fossil leaf preservation will therefore likely be required to examine mesophyll tissue differentiation in addition to stomatal distribution in order to assess if fossil taxa were dorsiventral or isobilateral and correctly interpret their paleoecology.

Furthermore, quantitative stomatal traits such as anatomical \( g_{\text{max}} \) can be used in conjunction with other anatomical data such as vein density and chemistry (e.g. carbon isotopic composition) to infer important ecophysiological characteristics such as transpiration rate (Steinhorsdottir et al., 2012), assimilation rate (Brodribb et al., 2007), and water use efficiency from fossilized leaves and cuticle fragments (Franks and Beerling, 2009a; Assouline and Or, 2013; Wilson et al., 2015; McElwain et al., 2016b; Montañez et al., 2016). This opens up the possibility of using stomata to make quantitative paleoecological comparisons between cohabiting fossil taxa (Montañez et al., 2016), which may be more fruitful than grouping all fossil taxa to broadly classify the climatic preference of an entire assemblage or collection locality. The former approach was used to provide a better mechanistic explanation underlying vegetation dynamics in response to glacial-interglacial cycles of the great Carboniferous ice age (between \( \sim 312 \) and 299 mya) and across the Triassic-Jurassic mass extinction boundary (\( \sim 201 \) mya; Wotzlaw et al., 2014; Steinhorsdottir et al., 2012). In both cases, \( g_{\text{max}} \) calculations based on the density and size of stomata in different fossil taxa were used to model paleo-transpiration rates and compare water use efficiencies of the taxa that suffered local extinction (giant lycopsid in the Late Pennsylvanian and Bennettitales in the latest Triassic) to those that survived and subsequently proliferated (Pecopteris and Ginkgo in the Late Pennsylvanian and Early Jurassic, respectively; Steinhorsdottir et al., 2012; Montañez et al., 2016).

Where the phylogeny of the fossil taxon under study is well known and detailed trait-environment relationships have been established, it may also be possible to characterize ecological habit using stomatal traits. For example, guard cell length appears to be tightly correlated with habit in Proteaceae with open vegetation displaying significantly longer guard cells than forested (Jordan et al., 2014).

**DETERMINING GENOME SIZE AND PLOIDY FROM FOSSIL STOMATA**

In rare and exceptionally fossilized specimens, genome size has been estimated from direct morphometric analysis of intact nuclei preserved within permineralized cells (Bomfleur et al., 2014). Scaling relationships between nuclear envelope size and DNA content then allowed estimation of genome size (Bomfleur et al., 2014) from which ploidy can be interpreted. Where such remarkable preservation is not available, however, alternative proxies for genome size are required. Two such proxies that have received increasing interest include pollen/spore size (Kürschner et al., 2013) and guard cell size (Wagner et al., 2000; Beaulieu et al., 2008; Lomax et al., 2009; Brodribb et al., 2013; Lomax et al., 2014), as both cell types have relatively conserved dimensions and provide an approximation of nuclear content and hence genome size (for review, see Lomax et al., 2014). Genome size is strongly positively correlated (\( r^2 = 0.59 \)) with guard cell size across 101 angiosperm species (Beaulieu et al., 2008; Lomax et al., 2014; Fig. 1A), and an increase in ploidy is associated with an increase in guard cell size in the gymnosperm species *Ginkgo biloba* (Šmarda et al., 2016). The fossil record of stomata is long-ranging, dating back almost to the earliest embryophyte, and they are preserved in many styles (permineralized, charcoalified, and dispersed cuticle) that enable approximation of guard cell dimensions. Stomata are therefore considered to offer a unique archive with which to study genome size evolution in plants (Lomax et al., 2009; Franks et al., 2012). However, the genome size-guard cell size relationship is by no means universal, as exceptions have been identified such as in the Proteaceae (Jordan et al., 2015) where large changes in guard cell size are not accompanied by similar magnitude changes in their genome.

As we have seen in previous sections of this review, stomatal size (guard cell size) is also influenced by the environment. In the case of fossils, is it possible to tease apart the effect of paleoenvironment and paleoecology from genome size with conviction? Lomax et al. (2009) addressed this issue for the extant model plant *Arabidopsis* (*Arabidopsis thaliana*) and demonstrated a very narrow range of environmental driven variability in guard cell length following exposure to a range of abiotic stressors including high UV-B radiation, super-elevated CO\(_2\) (2,000 ppmV), pathogen attack, and water deficit. Since then, it has been suggested that a change in genome size offers a means of changing cell size in plants to optimize gas exchange (Franks et al., 2012; Brodribb et al., 2013) but that a substantial amount of environmental driven change in guard cell size can occur over evolutionary time after a genome duplication event (Jordan et al., 2015). Put simply, it appears that the size of fossil stomata provide both valuable genotypic (e.g. genome size and potentially ploidy) and phenotypic data (e.g. maximum pore area for gas exchange), the latter of which can be determined by genome size and/or environment. As argued by Jordan et al. (2015), it seems that changes in cell size such as guard cells are adaptive and that change in genome size is one of several ways in which changes in cell size can occur. However, there is not always a simple relationship between genome size and ploidy because increased genome size can be caused by amplification of repetitive DNA sequences without whole-genome duplication (Lomax et al., 2014). Careful studies assessing guard cell size variability, together...
with independent paleoecological and paleoclimate data, are therefore required to assess the relative contribution of each driver.

As a demonstration of the complexity of interpreting changes in guard cell size in fossils, we reanalyzed an extensive Ginkgoales stomatal size data set from Steinhorskoditt et al. (2012) that spans the Triassic-Jurassic transition ~201.36 ± 0.17 mya (Wotzlaw et al., 2014), marked by transient global warming and the end-Triassic mass extinction event. The data set comprises mean stomatal pore length measurements (n = ~50 per leaf sample) from 47 discrete fossil leaf samples derived from three ginkgoalean genera: Ginkgo, Baeira, and Sphenobaiera (Supplemental Table S1). Guard cell lengths (gcl) for all samples were estimated from pore length (pl) measurements by assuming conservatively that gcl = 1.5(pl). Variability in fossil guard cell length over evolutionary time (an estimated ~8 million years) was assessed simplistically under two extreme scenarios; the first “environment only” assumes stasis in nuclear genome size of Triassic and Jurassic fossil ginkgoalean taxa, where observed variability in guard cell length is assumed to be driven exclusively by the extreme environmental change (McElwain et al., 1999; van de Schootbrugge et al., 2007; Richoz et al., 2012; Bacon et al., 2013) associated with the Triassic-Jurassic mass extinction event (Fig. 1A). Here, the fossil genome size is assumed to be equivalent to that of extant diploid G. biloba (2C DNA = 18.4 Gbp; Smarda et al., 2016; Fig. 1A). Under the second scenario, “genome only,” it is assumed that nuclear genome size is the predominant control on guard cell length across the Triassic-Jurassic mass extinction interval with no influence from environmental factors (Fig. 1B). It is of course more likely that both environment and genome size influenced the evolution of guard cell size and that our analysis masks likely genome size differences that occur between the different Ginkgoales genera. However, the demonstration is useful as a thought exercise as it highlights the potential impact of the two end-member scenarios. For instance if the “environment only” scenario is correct, our results suggest that, while variability in guard cell length exposed to abiotic stressors may be very constrained in some taxa and under short-term modern experimental exposure (e.g. Arabidopsis gcl data set of Lomax et al., 2009), the range of environmentally driven variability in other taxa is high given sufficient time for evolutionary adaptation (e.g. Ginkgoales gcl data set of Steinhorskoditt et al., 2012; Fig. 1A).

The “genome only” scenario is particularly interesting because the trends in estimated 2C DNA content hint at the presence of possible tetraploid Ginkgoales among individual samples of the fossil taxon Sphenobaiera spectabilis (Nathorst) florin from Early Jurassic aged fossil plant bed 8 from Astartekløft, East Greenland. Two samples of S. spectabilis are identified that have an estimated 2C DNA amount of ~47.1 Gbp (FMNH47853) and 46.9 Gbp (FMNH 47863), respectively (Supplemental Table S1), which both exceed the known 2C DNA amounts in
modern tetraploid *G. biloba* (37.4 Gbp; Šmarda et al., 2016). If the presence of tetraploids within the genus *Sphenobaiera* can be independently verified, it would add support to the suggestion by Fawcett et al. (2009) that mass extinction intervals in Earth history trigger polyploidization (whole-genome duplication) events or expose vacant niches into which existing polyploid taxa can diversify. It would also provide considerable support to long-standing, but mostly untested, hypotheses that a high prevalence of polyploidization events in land plants compared with animals is a key mechanism for generic and family level resilience and survivorship of plants at mass extinction boundaries (Willis and McElwain, 2014). High variability in the size and morphology of conifer pollen tetrads collected from Triassic-Jurassic aged sediments have been interpreted as evidence of unreduced gametes, indicating polyploidization in the extinct conifer Cheirolepidadeae (Kürschner et al., 2013). A similar but higher resolution study on size variability in fossil Ginkgoales pollen across the Triassic-Jurassic is now needed to test the hypothesis arising from fossil guard cell length data (Fig. 1B) that *S. spectabilis* may be represented by both diploid and tetraploid individuals within the Early Jurassic forests of East Greenland. Alternatively, increased 2C DNA amount may just reflect increased genome size in a diploid due to amplification of repetitive DNA sequences (Lomax et al., 2014).

**RECONSTRUCTING PALEOATMOSPHERIC CO₂ USING FOSSIL STOMATA**

**Assumptions and Limitations**

Fossil stomata have been extensively used to reconstruct atmospheric composition through Earth history, in particular atmospheric CO₂ concentration (*pCO₂*). Paleo-CO₂ reconstruction methods that use stomata as proxies include both empirical (McElwain and Chaloner, 1996; Wagner et al., 1996; Kürschner et al., 2008; Barclay et al., 2010; Maxbauer et al., 2014) and mechanistic (Wynn, 2003; Konrad et al., 2008, 2017; Franks et al., 2014) approaches. All approaches operate on the same underlying assumption—that the density and/or size of stomata are the primary control of leaf level gas exchange in vascular plants that is optimized to balance CO₂ uptake for photosynthesis against water loss through transpiration (Woodward, 1987; Katul et al., 2010; Manzoni et al., 2011). All empirical proxy CO₂ methods are also underpinned by the assumption of evolutionary conservatism. For example, CO₂ proxies that use an inverse relationship between stomatal density (or index) and atmospheric CO₂ make the implicit assumption that the slope and sign of the SD-∗pCO₂* relationship has not evolved significantly over geological time. Jordan (2011) suggested that high species variability in the magnitude and sign of stomatal density response to CO₂ between different species weakens the reliability of stomatal-based CO₂ proxies because it breaks the key assumption of evolutionary conservatism. We agree that this is the case for a number of families that have been investigated (Kelly and Beerling, 1995; Kürschner et al., 1997; Haworth et al., 2010); however, other families show strong conservatism in the SD-∗pCO₂* relationship with species belonging to the same genera, and even at the family or order level, clustering together with similar SD/SI values, as well as displaying similar response directions and magnitudes to changes in *pCO₂* (McElwain et al., 2002; Barclay et al., 2010; Haworth et al., 2011; Steinthorsdottir et al., 2011a; Steinthorsdottir et al., 2011b; Steinthorsdottir et al., 2016b). This is particularly important when operating in the pre-quaternary fossil record, which does not typically offer fossil plants that are conspecific with modern plants. The key to the selection and application of the most effective and accurate paleo-∗pCO₂* proxy is thus better characterization of how different taxa, both extant and extinct, control their gas exchange and whether they do indeed conform to the principles of evolutionary conservatism in the SD/SI-∗pCO₂* response required for empirical based CO₂ proxies to be appropriately applied.

Part of the reason why there is such a high apparent variability in the slope and sign of SD-∗pCO₂* relationships among different plant species is due to the fact that plants control gas exchange using multiple mechanisms including developmental responses that set the maximum anatomically possible stomatal conductance (∗gₚ₅₅₅₅₅₅ себе) and/or through physiological control of the stomatal aperture (Haworth et al., 2013; Brodribb and McAdam, 2017). ∗gₚ₅₅₅₅₅₅ себе is determined by both stomatal density and stomatal pore area and depth (Parlange and Waggner, 1970; Lawson and Morison, 2004; Franks and Beerling, 2009b), and although taxa use on average 25% of this maximum potential anatomical conductance—referred to as the operational stomatal conductance (∗gₚ₅₅₅₅₅₅ себе)—considerable species level differences in ∗gₚ₅₅₅₅₅₅ себе/∗gₚ₅₅₅₅₅₅ себе are apparent (Dow et al., 2014; Franks et al., 2014; McElwain et al., 2016b). Therefore, it follows that species-specific differences in the direction and magnitude of the SD/SI response to *pCO₂* can be compensated for by alteration to stomatal pore size and depth (Lammertsma et al., 2011) and/or by physiological control of aperture (in taxa with abscisic acid-mediated regulation of guard cell aperture; see Brodribb and McAdam, 2017), resulting in optimized control of gas exchange (Haworth et al., 2013) and a strong mechanistic underpinning for stomatal-∗pCO₂* proxies that will be discussed later in this section.

Empirical approaches use an inverse relationship between SD or SI with *pCO₂* to infer paleo-∗pCO₂* from SD and SI calculations on fossil leaves (for review, see Roiyer, 2002; Roth-Nebelsick, 2005; Jordan, 2011; Barclay and Wing, 2016). This type of proxy is simple to apply but limited by the fact that it ideally should only be applied to fossil taxa that are known to control gas exchange by predominantly developmental control of SD/SI, termed morphological responders by Haworth et al. (2013) (e.g. *G. biloba*), rather than via active
hormone-mediated physiological control of stomatal aperture (Brodribb and McAdam, 2011, 2017), termed physiological responders by Haworth et al. (2013). This simple dichotomy is complicated by the fact that the mode of stomatal response to CO₂ may be time-dependent, with short-term CO₂ fumigation experiments eliciting only physiological responses that reduce stomatal conductance with little or no significant change in SD (Roth-Nebelsick 2005), whereas long-term herbarium data sets of the same species show strong stomatal developmental changes that result in a reduction in SD in response to increasing CO₂ (Barclay and Wing, 2016).

Mechanistic approaches to pCO₂ reconstruction use SD plus pore length measurements to compute maximum theoretical stomatal conductance (g_{max}; Box 1) together with other important factors influencing gas exchange, such as boundary layer and mesophyll conductance. Total estimated conductance from fossils is then used in conjunction with carbon isotopic data (used to estimate Ci) and the Farquhar photosynthesis model to estimate paleo-pCO₂ (Wynn, 2003; Konrad et al., 2008, 2017; Franks et al., 2014; McElwain et al., 2016a, 2017; Franks and Royer, 2017; Box 1). These methods take into account both developmental (from g_{max} measurements) and physiological control (from carbon isotopic composition) of plant gas exchange and therefore have wide application in the fossil record, particularly in relation to extinct taxa, for which the stomatal control of gas exchange is unknown. They represent significant improvements because they are not merely correlational, but provide a mechanistic basis for interpreting changes in SD and stomatal pore length. Furthermore, they do not require an assumption of evolutionary conservatism in the SD-CO₂ relationship, which is a prerequisite of empirical approaches. Just like all empirical stomatal-pCO₂ models, however, the newly developed mechanistic models (Konrad et al., 2008, 2017; Franks et al., 2014) are limited by a requirement to cross-calibrate with extant nearest living relative (NLR) or nearest living equivalent (NLE) taxa (McElwain et al., 2017). For example, the stomatal ratio approach of McElwain and Chaloner (1996) estimates paleo-pCO₂ by comparing fossil SD or SI with that of an NLE using one of two possible calibrations, the carboniferous standardization: pCO₂ (paleo) = SD_{NLE}/SD_{Fossil} × 600, and the recent standardization, pCO₂ (paleo) = SD_{NLE}/SD_{Fossil} × pCO₂_{2NLE}. The Franks et al. (2014) mechanistic model needs to be parameterized with fossil plant net assimilation rate under modern ambient pCO₂ (termed A₀), a task that is difficult to do without cross-comparison with modern taxa (McElwain et al., 2017). The Konrad optimization models require parameterization with multiple photosynthetic variables taken from NLRs or NLEs.

Stomatal pCO₂ Proxies and the Role of Carbonic Anhydrase

A likely mechanism supporting the stomatal proxy (explaining how plants adjust their stomatal densities in response to changes in atmospheric pCO₂) has been proposed: Plants use carbonic anhydrase (CA) located principally in the guard cells to detect (sense) pCO₂ enveloping their leaves (Hu et al., 2010, 2015; Chater et al., 2015; Engineer et al., 2016) and control initiation of stomatal development via a signal transduction pathway that is modulated by the HIC gene (Gray et al., 2000; Brownlee, 2001). It has further been shown experimentally that transpiration rates of mature leaves correspond with stomatal densities in developing leaves, suggesting a link between the short-term control of the stomatal aperture and the long-term regulation of stomatal development (Lake and Woodward, 2008). To date, however, this has only been demonstrated for angiosperms and other reviews within this special issue (Brodribb and McAdam, 2017) highlight the complexity of evolution of stomatal control in land plants as a whole. CAs are distributed among all three domains of life, being a group of enzymes that catalyze the rapid conversion of CO₂ and water to bicarbonate and protons (and back)—a fundamental reaction to all biological processes (Elleuche and Pöggeler, 2010; Cummins et al., 2014), suggesting that the process may be evolutionary highly conservured or ancestral (Frommer, 2010; Chater et al., 2015). Most of the molecular and genetic studies referred to above have involved the model plant Arabidopsis, and although CAs have also been shown to play a similar role in stomatal responses and photosynthesis in, e.g. Nicotiana (Hu et al., 2015), Flaveria (Ludwig, 2011), and even Olea (olive trees; Perez-Martin et al., 2014), much more work needs to be done to understand the relationship between the sensory and signaling mechanisms and the long-term responses of stomatal development to changing environmental conditions (Doheny-Adams et al., 2012). Future work should further illuminate the potential taxon-specific differences in stomatal response to pCO₂ and other environmental factors between plant groups (Brodribb and McAdam, 2013; Merilo et al., 2014).

The way mature leaves (early shoots) control stomatal development of younger leaves through long-distance signaling (Lake et al., 2001, 2002) may in part explain early failures to replicate the inverse relationship of stomata to pCO₂ in some modern CO₂ enrichment experiments (Royer, 2001; Reid et al., 2003; Tricker et al., 2005; Ainsworth and Rogers, 2007) as well as indicate that plants may need to be grown in atmospheric experimental conditions for longer time intervals than has typically been the case, in order to obtain a reliable SD/SI response to elevated pCO₂. Experimental studies to date have usually been short-term, limited mostly to a single growing season or less, probably in most cases not providing the proper conditions for stomatal responses to pCO₂ to be morphologically expressed. Strongly supporting this assumption is a recent growth chamber experiment, exposing Betula nana to three levels of pCO₂ (150, 450 and 800 ppm) over two successive growing seasons, finding that while some adjustment of stomatal parameters took place in the first growing season, amplified adjustment of stomatal properties, such as SD, occurred mostly in the experiments’ second year.
Uses of Fossil Stomata

**BOX 1. Definition of Terms**

**Assemblage:** Collection of fossils leaves/wood/reproductive structures from a single locality or fossil plant bed.

**Empirical stomatal model:** A palaeoCO₂ proxy model that is constructed using observed correlation between stomatal density (SD) or stomatal index (SI) and atmospheric CO₂ concentration (pCO₂) in historical herbarium specimens and/or in plants exposed to a range of experimentally controlled pCO₂ in growth chambers. Inverse regression analysis is used to estimate unknown pCO₂ from fossil SD or SI. For example, Barclay and Wing (2016) propose the following revised empirical stomatal model for Ginkgo biloba based on a simple univariate regression between G. biloba herbarium specimen SI values and atmospheric pCO₂ ranging from 290 to 429 ppm: pCO₂ = 9920.9(SI⁻¹) [ppm].

**Maximum theoretical stomatal conductance (gmax):** The maximum calculated potential conductance of stomata to water (g(mv)) or CO₂ (g(mv)) assuming all stomata on the leaf surface are fully open in the geometry of a circle or ellipse and assuming no heterogeneity in stomatal behaviour: gmax is calculated (Parlane and Waggone, 1970; Franks and Beering, 2009b) as follows:

$$g_{\text{max}} = \frac{d_{w} \cdot \text{SD} \cdot p_{\text{max}}}{p_{\text{d}} + \frac{\pi}{2} \sqrt{p_{\text{max}}/\pi}}$$

where $d_{w}$ is diffusivity of water vapor at 25°C (0.0000249 m² s⁻¹), $v$ is molar volume of air (0.0224 m³ mol⁻¹) at both constants, SD is stomatal density (m⁻²), $p_{\text{max}}$ is maximum stomatal pore area (m²) calculated as an ellipse using stomatal pore length (L in m) as the long axis and L/2 as the short axis, and $p_{\text{d}}$ is stomatal pore depth (m) considered to be equivalent to the width of an inflated, fully turgid guard cell.

**Mechanistic stomatal model:** A palaeoCO₂ proxy model that uses an underlying understanding of leaf total diffusivity and photosynthetic responses to CO₂ in modern extant taxa to reconstruct unknown pCO₂ from estimates of fossil plant gas exchange and photosynthetic characteristics. For example, the model of Franks et al. (2014) is used to estimate pCO₂ ($c_{2}$) from estimated fossil leaf assimilation rate ($A_{s}$), total conductance to CO₂ diffusion ($g_{\text{leaf}}$) and the ratio of pCO₂ inside ($c$) and outside ($c_{2}$) the leaf.

$$c_{2} = \frac{A_{s} \cdot g_{c}(10^c) \cdot (1 - \frac{c}{c_{2}})}{g_{c}(10^c) \cdot (1 - \frac{c}{c_{2}})}$$

**Neighbouring cells:** Cells immediately adjacent to the guard cells which are morphologically and/or chemically indistinguishable from epidermal cells.

**Operational conductance (gσ):** The measured stomatal conductance of a leaf by porometry or infrared gas analysis that is representative of typical daily stomatal conductance ($g_{q}$) to water and CO₂ which occurs under non-drought-stressed field conditions and that avoids midday depression in $g_{q}$ if present in the taxa. See McElwain et al. (2016b) for measurement protocols.

**Stomatal complex:** Also known as the stomatal apparatus, this is made up of two guard cells together with neighbouring cells and/or subsidiary cells.

**Subsidiary cells:** Cells immediately adjacent to guard cells that are morphologically distinct in form, size, structure, and/or chemistry from epidermal cells. The ontogenetic definition of subsidiary cells is not considered here because it is not of use in fossil stomatal studies.
(Hincke et al., 2016). An additional reason why so few growth chamber enrichment studies have documented a reduction in SD in response to elevated pCO2 may be counting error (Barclay and Wing, 2016). Barclay and Wing’s reanalysis of elevated pCO2 grown G. biloba from growth chamber experiments revealed that 45% of stomata were malformed and likely lacked functionality compared with 35% of ambient chamber grown plants. The malformed stomata represent cells that have arrested at various stages in the developmental pathway to guard cell pair formation and therefore should not be included in a stomatal density or index count as they do not contribute to the leaf’s overall diffusivity.

Insights from Stomatal-pCO2 Model Cross-Comparison Studies

A recent high-resolution paleo-pCO2 reconstruction spanning 16 million years of the Carboniferous and Early Permian (~312–296 mya) uses three methods: the paleosol-pCO2 proxy (Breecker, 2013; Montañez, 2013), an empirical stomatal-based proxy model (McElwain and Chaloner, 1995, 1996), and a mechanistic stomatal based model (Franks et al., 2014) to derive a consensus pCO2 record (Montañez et al., 2016). Glacial-interglacial fluctuations in atmospheric pCO2 were reconstructed that correlate well with inferred sea level and modeled polar ice volume records (Montañez et al., 2016). Interestingly, both empirical and mechanistic paleo-pCO2 models applied to the same fossil taxa (Medullosa, an extinct gymnosperm) yield very comparable pCO2 estimates throughout the Carboniferous (Fig. 2; \( r^2 = 0.4399, \; P = 0.00076 \)). The generally good correlation between pCO2 estimates based on the different calibration approaches confirms that SD (and SI) is likely the major control of gas exchange in the medullosan seed ferns studied. Furthermore, the good consistency between the two stomatal based methods supports the robustness of the simpler stomatal-ratio method that only requires SD or SI measurements from fossil leaves to estimate pCO2. Unlike application of the Franks et al., (2014) model, it does not require detailed measurements of stomatal pore length and depth, which can be difficult to observe in all fossil preservation types and in fossil taxa with sunken guard cells.

Similar congruence in pCO2 estimates are observed in a cross-comparison of different empirical based calibrations (e.g. McElwain, 1998; Barclay and Wing, 2016) applied to the same fossil Ginkgoales leaves spanning the Triassic-Jurassic mass extinction boundary (Fig. 3). McElwain’s (1998) NLE method yields almost identical pCO2 estimates to Barclay and Wing’s (2016) updated SI-CO2 transfer function based on modern G. biloba applied to extinct Ginkgo, Baiera, and Sphenobaiera (Fig. 2). Could these two case study examples for the Carboniferous (Fig. 2) and Triassic-Jurassic (Fig. 3) suggest a high degree of stasis in the evolutionary relationship between stomatal frequency and pCO2 giving confidence to the assumption of uniformitarianism that underpins the stomatal-CO2 proxy method?

In other cases, however, empirical and mechanistic methods do not fully agree, even when based on the same fossil stomatal dataset. This has been the case for some Cenozoic pCO2 reconstructions, where, for example, estimates for Late Eocene and Early and Late Oligocene pCO2, although broadly comparable quantitatively, do not always agree on pCO2 trends through time (see Roth-Nebelsick et al., 2012; Grein et al., 2013; Steinthorsdottir et al., 2016a). Part of the issue here may be that the optimization model of Konrad et al. (2008, 2017) requires calibrating pCO2 using multiple contemporaneous (overlapping) species to derive a best estimate of pCO2, making it difficult to compare to single-species databases.

Much of the variability in pCO2 estimates derived using various stomatal proxy CO2 models when examined more broadly stems from differences in stratigraphic position and/or resolution, differences in fossil sample number/replication, and due to different methodological approaches to raw data collection and subsequent calibration. In addition, although SD-pCO2 calibration models are continuously being updated with improved data sets and methodological approaches (Ginkgo in Barclay and Wing, 2016), pCO2 estimates based on outdated calibration models remain part of the literature (e.g. Retallack, 2001; Royer et al., 2001). Very few true cross-calibration studies are available to test the performance of different stomatal models and calibrations using the same fossil stomatal datasets. Detailed cross-comparison studies of the different stomatal based methods are now urgently required for a range of phylogenetic groups and a range of geological time intervals to improve our understanding of the weaknesses and strengths inherent in all methods where stomatal data are used in whole or in part to propagate paleo-CO2 estimates. The challenge for the future is to develop a truly taxon independent paleo-pCO2 proxy that is not hampered by phylogenetic differences in the rate, magnitude, and sign of stomatal, isotopic, and photosynthetic responses to atmospheric pCO2. In the meantime, the robustness of paleo-pCO2 estimates based on fossil stomata can be greatly improved by taking a multiproxy approach (Jordan, 2011) that incorporates both mechanistic and empirical models, through careful cuticle preparation and counting protocols (Barclay and Wing, 2016), and by undertaking calibration experiments that span at least two rather than one growth season where empirical approaches are used.

DEVELOPMENTAL INSIGHTS FROM THE FOSSIL STOMATAL COMPLEX

Studies of fossil stomata and their associated subsidiary and neighboring cells can provide insights into the origin and evolution of developmental pathways of the stomatal complex (Barbacka and Bóka, 2000; Barclay et al., 2007; Bomfleur and Kerp, 2010; Rudall et al., 2013).
Fossil fingerprints” may in the future help to identify the stomatal developmental pathways for the many extinct lineages that remain unknown (Rudall et al., 2013). Fingerprints or developmental markers of a mesogenous stomatal developmental pathway include (Rudall et al., 2013) (1) nonrandom stomatal orientation in relation to the leaf long axis, (2) two size classes of stomata present on the same leaf, and (3) variability in epidermal cell size, including alteration of long and short epidermal cells. Identifying these fingerprints in fossils requires excellent anatomical preservation and positional information along a fossil leaf. For example, Bennettitiales show strong morphological disparity in all of these traits moving from the basal to apical portion of the leaf (Fig. 3 in Steinhorscdottir et al., 2011a).

Of interest from the perspective of this review is the likely physiological significance of different developmental pathways, if any. Paleo-pCO2 proxies based on empirical relationships between SD and SI responses to pCO2 are limited by the fact that some taxa are CO2 sensitive in their developmental response (Haworth et al., 2011), while others are not (Haworth et al., 2011). To date, no study has investigated if the stomatal density response to CO2 is modulated in some way by the different developmental pathways of the stomatal complex. The study of Barclay and Wing (2016) illustrated that elevated CO2 increased malformation of the stomatal complex but their focus was on guard cell formation rather than on subsidiary and/or neighboring cells within the stomatal complex. Equally, the debate regarding the evolution of active versus passive stomatal control (Brodribb and McAdam, 2011; 2017) is ongoing, yet no study to our knowledge has attempted to correlate physiological responsiveness of the stomatal pore to abiotic stimuli with the developmental pathway of the stomatal complex. If the magnitude of stomatal aperture opening and closing responses to abiotic stimuli or the sensitivity of SD response to pCO2 could be meaningfully linked to perigenous versus mesogenous stomatal complex development, this would open up the possibility of using fossil developmental fingerprints to categorize the utility of fossil taxa as paleoatmospheric proxies and assess the likely stomatal control of extinct lineages. An example working hypothesis is that mesogenous taxa are highly responsive to CO2 via a stomatal density response because asymmetrical divisions may provide greater flexibility to alter stomatal spacing without altering total cell number.

Another interesting avenue would be to explore the energetic cost of different stomatal complex types in extinct and extant taxa using de Boer et al.’s (2016) proxy measure of the “cost” of stomata (termed fgc), which is calculated as the fraction of epidermal cells allocated to stomata. Perhaps the fgc function should be expanded to include all mesogenous cells within the stomatal complex or all subsidiary cells whatever their developmental pathway if they serve a supporting functional role to guard cells. The wide variability in

This is more than just an academic exercise as a deeper understanding of the origin, evolution, and diversity of stomatal developmental pathways in both extant and extinct lineages underpins genetic engineering programs for altered stomatal conductance, assimilation rates, and leaf cooling capacities in modern crop plants. This is because traits that are physiologically important to plant carbon acquisition and water loss, such as maximum anatomical stomatal conductance (Franks and Beerling, 2009b; de Boer et al., 2016; McElwain et al., 2016b), are all ultimately controlled by stomatal development. According to Rudall et al. (2013), the perigenous stomatal complex was likely ancestral in land plants, whereby neighboring and subsidiary cells adjacent to guard cells developed from different meristemoids than those from which guard cells developed. An alternative developmental pathway called mesogenous, where guard 2017 and subsidiary cells developed from the same meristemoid, likely developed later (for review, see detail in Rudall et al., 2013). A combination pathway called mesoperigenous, where subsidiary and/or neighboring cells derive from both perigenous and mesogenous routes is also evident in some fossil lineages, for example, in the extinct gymnosperm lineage Bennettitales.
mature stomatal complex morphology observed in fossils is also of interest (Table I). Why do some taxa have many obvious subsidiary cells, while others have neighboring cells that cannot be distinguished from normal epidermal cells? In part, the disparity may be due to the style of fossil preservation (cuticle versus permineralization etc.) and the type of microscopy used to observe fossil stomata (e.g. bright-field, epifluorescent, scanning electron microscopy). For example, while bright-field microscopy may not always be able to distinguish subtle differences in the morphology of cells adjacent to guard cells from normal epidermal cells, epifluorescence can readily distinguish these subsidiary cells from neighboring cells via differences in their autofluorescence.

Distinct functional roles have been identified in the subsidiary cells of modern plants, including sequestration and isolation of metal ions, such as nickel in the Brassicaceae species Thalaspi montanum var. siskiyouense (Heath et al., 1997), providing mechanical support to guard cells (Franks and Farquhar, 2007) and providing a source of water and/or ions to guard cells. It is also hypothesized that the role of subsidiary cells is to balance differences in growth rate between normal epidermal pavement cells and guard cells (Rudall et al., 2013). Perhaps then different subsidiary cell arrangements are related to whole plant relative growth rate? Cleal and Shute (2012) observed a wide variety of different subsidiary cell arrangements in Carboniferous Medullosales (Table I) with an apparent trend in mature stomatal complex types through time from simple anomocytic to stephanocytic and paracytic. Could these differences observed within Medullosales stomatal complex type reflect changes in the relative growth rate of the fronds belonging to different species? It certainly appears as though detailed assessment of fossil stomatal complex developmental pathways (Rudall et al., 2013) together with quantitative paleoecological traits such as the ratio of amphistomaty (SD adaxial surface/SD abaxial surface; Muir, 2015) have the potential to yield important insights on plant habit in the future.

CONCLUSION

What has emerged from this brief review is that fossil stomata are being underutilized as tools in most studies and have the potential to provide integrated insights on paleoecology, taxonomy, paleatmospheric composition, genome size, and developmental biology (see Outstanding Questions). For best application of empirical based proxies for paleo-pCO2 reconstruction, we recommend the use of multiseason elevated pCO2 experiments and/or multiseason herbarium data sets to generate appropriate transfer functions and calibrations.
OUTSTANDING QUESTIONS

- Can the fossil record of stomata provide insights into the ongoing debate on the origin of active versus passive stomatal control?
- Is there a universal scaling relationship between $g_{w}$ and $g_{max}$ that transcends phylogenetic, biogeographic, climatic, and functional differences between taxa?
- Can different ontogenetic pathways of stomatal complex development be recognized in the fossil plant record using chemical mapping techniques, such as FTIR, or by differences in the autofluorescence properties of subsidiary cells compared with epidermal cells?

Mechanistic models for paleo-CO$_2$ reconstructions represent very significant advances in the field. Careful assessment of whether stomata have been aborted or are fully developed and functional is also a key prerequisite of obtaining accurate stomatal density and index counts from fossils and neotaxonomic specimens. Where possible, the application of multiple paleo-pCO$_2$ proxies to fossil stomatal data sets is also recommended as these will improve the robustness of pCO$_2$ estimates by reporting the consensus pCO$_2$. We suggest that trends in guard cell length data are interpreted from two end member perspectives: an environment-only and a genome-only scenario unless independent data sets allow the relative impact of environment and genome to be assessed with confidence. We recommend moving away from attempts to characterize the entire ecological setting of a fossil assemblage using presence/absence traits associated with stomata. Instead, we advocate the use of quantitative traits that provide comparative estimates of water use efficiency, maximum gas exchange capacity, and leaf level transpiration of co-occurring fossil taxa within a fossil assemblage. Finally, of all the possible stomatal-based traits that can be observed in fossils, it appears that quantitative assessment of stomatal distribution across both leaf surfaces (Muir’s stomatal distribution ratio; Muir, 2015) together with estimated genome size (Beaulieu et al., 2008) will yield meaningful inferences on relative growth rate and plant habit and help to assess hypotheses on the likely growth form of extinct taxa.

Supplemental Data

The following supplemental materials are available.

Supplemental Table S1. Fossil Ginkgoales stomatal data used to estimate genome size (2C DNA amount).

ACKNOWLEDGMENTS

We thank C. Evans-Fitzgerald and A. Porter (University College Dublin) for providing feedback on the manuscript. We also thank G. Jordan and two anonymous reviewers for feedback and comments that greatly improved the manuscript.

Received February 10, 2017; accepted May 10, 2017; published May 11, 2017.

LITERATURE CITED


Beaulieu JM, Leitch IJ, Patel S, Pendharkar A, Knight CA (2008) Genome size is a strong predictor of cell size and stomatal density in angiosperms. New Phytol 179: 975–986


Brodribb TJ, McAdam SAM (2013) Unique responsiveness of angiosperm stoma to elevated CO2 explained by calcium signalling, PLoS One 8: e82057


Richoz S, Van De Schootbrugge B, Pross J, Pütthmann W, Quan TM, Kollist H (2016a) Comment on “Was atmospheric CO₂ caged at 1000ppm over the past 300 million years?”. Palaeargeogr Palaecolimol Palaeeoc 441: 653–658
Richoz S, Van De Schootbrugge B, Pross J, Pütthmann W, Quan TM, Kollist H (2016b) To open or to close: stomatal responses to simultaneously applied opposing environmental factors. New Phytol 202: 499–508
Raw_text_19: Uses of Fossil Stomata


© 2017 American Society of Plant Biologists. All rights reserved.
quantified: calibration, Eocene case study, and implications. Paleobiology 33: 574–589


Steinthorsdottir M, Bacon KL, Popa ME, Bochner L, McElwain JC (2011a) Bennettitalean leaf cuticle fragments (here Anomozamites and Pterophyll) can be used interchangeably in stomatal frequency-based palaeo-CO(2) reconstructions. Palaeontology 54: 867–882


