Quantifying the complexity of plant reproductive structures reveals a history of morphological and functional integration

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Proceedings B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>RSPB-2023-1810.R1</td>
</tr>
<tr>
<td>Article Type:</td>
<td>Research</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>07-Oct-2023</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Leslie, Andrew; Stanford University, Geological Sciences Department Mander, Luke; The Open University</td>
</tr>
<tr>
<td>Subject:</td>
<td>Palaeontology &lt; BIOLOGY, Evolution &lt; BIOLOGY</td>
</tr>
<tr>
<td>Keywords:</td>
<td>macroevolution, networks, gymnosperms, angiosperms, vascular plants</td>
</tr>
<tr>
<td>Proceedings B category:</td>
<td>Palaeobiology</td>
</tr>
</tbody>
</table>
Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:

Statement (if applicable):
CUST_IF_YES_ETHICS : No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:

Statement (if applicable):
All data and adjacency matrices, along with Python and R scripts to analyze them, are available on Dryad doi.org/10.5061/dryad.n02v6wx36. Part type lists for reproductive structures are available in the electronic supplementary material.

If yes, please tell us how your data or code can be accessed and provide a link to your data if it is in a repository for the editors and reviewers to use.
CUST_DATA_INITIAL_ACCESS : No data available.

If your data is in a repository, please provide any temporary (private) link/s for reviewers/editors to access here.
doi.org/10.5061/dryad.n02v6wx36

Conflict of interest

I/we declare we have no competing interests.

Statement (if applicable):
CUST_STATE_CONFLICT : No data available.

Use of AI

Please provide a statement of any use of AI technology in the preparation of the paper.

No, we have not used AI-assisted technologies in creating this article
CUST_IF_YESDECLARATION_OF_AI_USE : No data available.
Quantifying the complexity of plant reproductive structures reveals a history of morphological and functional integration

Andrew B. Leslie¹* and Luke Mander²

Affiliations
¹Geological Sciences Department, Stanford University; 450 Jane Stanford Way, Building 320, Room 118, Stanford, CA 94305, USA.
²School of Environment, Earth and Ecosystem Sciences, The Open University; Walton Hall, Milton Keynes, MK7 6AA, UK.

*Corresponding author. Email: aleslieb@stanford.edu

Keywords: macroevolution, networks, gymnosperms, angiosperms
Abstract

Vascular plant reproductive structures have undoubtedly become more complex through time, evolving highly differentiated parts that interact in specialized ways. But quantifying these patterns at broad scales is challenging because lineages produce disparate reproductive structures that are often difficult to compare and homologize. We develop a novel approach for analyzing interactions within reproductive structures using networks, treating component parts as nodes and a suite of physical and functional interactions among parts as edges. We apply this approach to the plant fossil record, showing that interactions have generally increased through time and that the concentration of these interactions has shifted towards differentiated surrounding organs, resulting in more compact, functionally integrated structures. These processes are widespread across plant lineages, but their extent and timing vary with reproductive biology; in particular, seed-producing structures show them more strongly than spore or pollen-producing structures. Our results demonstrate that major reproductive innovations like the origin of seeds and angiospermy were associated with increased integration through greater interactions among parts. But they also reveal that for certain groups, particularly Mesozoic gymnosperms, millions of years elapsed between the origin of reproductive innovations and increased interactions among parts within their reproductive structures.
Introduction

Vascular plants produce an extraordinary variety of reproductive structures, from the simple clusters of sporangia found in many ferns to the elaborate flowers of orchids (1-3). The complexity of structures such as flowers reflects not only their diverse and highly differentiated parts (3-5) but also the intricacy of interactions among these parts. Botanists have long noted, for example, how specialized floral organs are often borne in specific geometric arrangements that bring anthers, stigmas, and animals together to facilitate pollination (6-10). Floral complexity also reflects extensive fusion and coordinated growth among component parts (2,11-15). Such functional and developmental integration among reproductive traits has been most widely studied in angiosperms (16-20), but these phenomena are not restricted to the group; reproductive structures with some degree of functional integration are common in extant plant lineages from lycopsids with compact spore-producing strobili (21) to conifers with tightly interlocking seed cone scales (22,23). The evolution of diverse and specialized functional interactions among reproductive parts then appears to be an important aspect of plant evolutionary history.

Such interactions have undoubtedly increased over vascular plant evolutionary history, given that the earliest known reproductive structures consisted simply of solitary or loose aggregations of sporangia (24-26). Many later groups appear to have evolved more complex morphologies by co-opting organs to perform new functional roles in spore release, pollination, seed protection and seed dispersal (3,21,27,28), and novel interactions among reproductive parts are an important aspect of this process. For example, major reproductive innovations like the origins of seeds and angiospermy reflect close physical, developmental, and functional associations between reproductive organs and enclosing structures, involving either the
integument or seed coat alone (29, 30) or the integument plus the carpel in angiosperms (31).

Within seed plants, Mesozoic lineages evolved a wide diversity of cupules, scales, and compact strobili that cover seeds (see 25) and which are often assumed to protect them in some capacity (32-34).

That interactions among plant reproductive parts have increased over time and are associated with morphological diversification is intuitive, but the extent and tempo of these trends are difficult to evaluate. Although seed plants are widely regarded as having evolved more integrated and complex reproductive structures (e.g., 2-4,13), other diverse clades such as leptosporangiate ferns (Polypodiidae sensu 35) have maintained simple reproductive structures over their evolutionary history (3). Quantitatively assessing temporal patterns in reproductive part interactions across clades is challenging because they are difficult to consistently define and characterize across lineages with disparate reproductive structures, especially those in which homologies among parts are nonexistent or uncertain (see discussions in 36-38). In this study, we use a network approach (39) to overcome some of these difficulties. Networks have been used in previous studies of morphological integration (40), morphological complexity (41), and ecological complexity (42, 43). Here they allow us to create a flexible but consistent framework in which to characterize temporal patterns in part interactions and functional integration, regardless of the age, affinities, or specific homologies of the reproductive structures.

To create reproductive interaction networks, we divide reproductive structures into their basic component part types following a previous study (3), which we treat as nodes in a network. We then score five kinds of interactions among these parts and represent these interactions as edges linking network nodes. Interactions represent basic physical and geometric relationships among parts, including three types describing physical contact (attachment, fusion, and
conforming, which refers to closely adpressed parts), and two describing nested spatial
relationships (envelopment and enclosure of one or more part types by another). The simple
networks that we build from these interactions, which can often be clearly recognized in well-
preserved fossils, encode a suite of potential functional relationships among the basic units of a
plant reproductive structure. By analyzing networks from both extant and extinct plants, we ask
if interactions among reproductive parts have increased through time generally and within major
plant lineages specifically. We also use the topological structure of our networks to ask if the
location and concentration of part interactions has changed through time; in particular, we ask if
interactions have shifted away from the reproductive organs themselves and onto auxiliary
structures, as a greater variety of organs have become incorporated into reproductive structures
over plant evolutionary history (see 27, 28).

Materials and Methods

We used published literature descriptions supplemented with direct observations from
herbarium specimens to score 1461 reproductive structures (905 fossil, 556 extant) from 1306
vascular plant taxa (we treat conspecific micro- and megasporangiate structures as separate units
of analysis). A full list of taxa and reproductive structures can be found in the Supplementary
Materials, and sources are listed in files on Dryad (https://doi.org/10.5061/dryad.n02v6wx36).
The number of nodes for each reproductive structure is based on a list of part types derived from
a previous study (the 'morphological element types' of [3]), although modified in some cases to
include updated information. We defined part types as distinct geometric regions of reproductive
structures summed over ontogeny; in angiosperms, for example, the total number includes parts
at pollination and fruit maturity. It is important to emphasize that the number of network nodes
represents distinct part types rather than the absolute number of parts; for example, a single part
type such as a petal may be expressed many times in a reproductive structure. As in (3), we
sampled reproductive structures at the genus level to minimize uneven sampling because many
fossil reproductive structures belong to monotypic genera. Species within genera often do not
vary in part types present (i.e., node number) or part interactions as we define them, although in
cases with known variation we did include multiple representatives. For fossil taxa with a wide
stratigraphic range, we generally sampled one representative per geologic epoch. Among extant
angiosperms, we sampled one representative taxon per major clade (i.e., monophyletic families
following APG IV [44]) due to their extreme species richness (although not all families were
sampled). This dramatically undersamples extant angiosperms and our results should not be
taken to encompass their full range of diversity or complexity, although our analyses are not
specifically focused on angiosperms and exhaustive sampling of their diversity is not required
for the basic comparisons with other vascular plant groups that we discuss here. See
Supplementary Materials for a description of part types in each taxon as well as a more in-depth
discussion of sampling and scoring concepts).

For each reproductive structure, we scored five kinds of part interactions: 1) attachment,
2) fusion, 3) conforming, 4) envelopment, and 5) enclosure. Attachment describes organic or
anatomical connection among part types at a small point or thin edge, while fusion describes part
types that are attached for more than 1/3 of their length or surface area. We represent fusion in
our networks as two edges, one representing attachment (a necessary condition of being fused)
and one representing the fusion interaction itself (Fig. 1a,b). Conforming or adpression
interactions are represented by a single edge and refer to part types with regular, close contact
(although not direct anatomical attachment) for much of their ontogeny; this relationship creates
a closed surface as the structure grows and develops (Fig. 1a,b). Envelopment occurs when one part type surrounds another; this interaction most often involves a cup-like structure although it may also include parts that are sunken into subtending organs. An enveloping part need not physically contact internal part type(s) but as defined here must form a continuous wall or tissue groundmass surrounding them for half or more of their length or surface area. Enclosure is an extension of envelopment where internal part type(s) are completely sealed off from the outside environment throughout their ontogeny and only interact with it through mechanical fracture (e.g., seeds released from a dehiscent fruit) or a dedicated opening (e.g., a seed micropyle). Enclosure necessarily includes envelopment, but the distinction between these two is important because enclosure necessitates a functional change in how reproductive propagules (spores, pollen, or seeds) are released or captured. We therefore use two edges, one for envelopment and one for complete enclosure itself, to represent this interaction in our networks. We refer to fusion, conforming, envelopment, and enclosure interactions as “engagement edges” throughout this study because they add additional physical or geometric links beyond simple attachment, and we interpret them as indicating greater integration among reproductive structures. For detailed descriptions and examples of these interactions, as well as the full list of interactions for all taxa, see Supplementary Materials.

For each reproductive structure, we used the part types and their interactions to construct two kinds of networks (Fig. 1b): one with attachment edges only (‘backbone’) and one with attachment plus engagement edges (‘complete’). For both backbone and complete networks, we scored part types that interact with themselves (e.g., sporophylls that are imbricated or sporangia that are fused) as self-loops (Fig. 1b). We note that edges in our networks represent relationships between part types in aggregate rather than between every individual element belonging to those
types; for example, the seed integument in a conifer cone is commonly conformable with adjacent ovuliferous scales and the part types are therefore scored as exhibiting an engagement edge, even though not every ovuliferous scale in a cone produces seeds.

We represented each network by an adjacency matrix, which is a symmetric matrix that records the number of edges among pairs of nodes (39; see Fig. 1b; all adjacency matrices available on Dryad at https://doi.org/10.5061/dryad.n02v6wx36). In total, we analyzed 763 complete adjacency matrices (many reproductive structures are described by the same adjacency matrix) and their backbone equivalents. In each matrix, the sporangium or megasporangium was scored as the 0th node (Fig. 1a,b); all other parts were numbered arbitrarily. To analyze temporal patterns, we calculated basic metrics (see 39) including the number of nodes and edges in a network, as well as the maximum and average degree of the network (the degree of a node is the number of edges incident to it [39]). We divided reproductive structures into free-sporing, pollen-producing, and seed-producing categories, which correspond to major differences in reproductive biology and function, and binned them into geologic periods, although we separated the Devonian (D) and Cretaceous (K) into D1 (Early; 419-393 Ma), D2 (Middle to Late; 393-359 Ma), K1 (Early; 145-100 Ma), and K2 (Late; 100-66 Ma) intervals to capture major shifts occurring during those periods. Within each time bin, we used the number of engagement edges and average degree for each taxon to calculate minimum, maximum, interquartile range (IQR), and median values. We also performed a bootstrap resampling procedure to calculate 95% confidence intervals on the median number of engagement edges per taxon. For each interval, we resampled the number of engagement edges and average degree with replacement and calculated the median, repeated this 1000 times, and then took the 95% quantile range of this distribution. We also repeated these analyses for complete networks where self-loops had been removed.
because a self-interaction places two edges incident upon a node and may substantially increase maximum and average degree. The results of these analyses, however, show similar temporal patterns to those that included self-interactions (see Fig. S7).

For each taxon, we also analyzed how the concentration of interaction edges varied throughout the reproductive structure by determining how node degree changes with distance from the sporangium or megasporangium. We recorded distance as the number of attachment edges that separate a given part type from the sporangium. For example, in the hypothetical structure from Fig. 1, sterile bracts (Part Type 3) are attached to the cone axis (Type 4), which is in turn attached to a seed integument (Type 1) that attaches to the megasporangium (Type 0); the sterile bracts are then two nodes and three attachment edges from the megasporangium (Fig. 1a,b). Across all reproductive structures, we used backbone networks to assign each part type a distance representing the shortest pathlength in number of attachment edges from that part type to the 0th node, which by convention here is the sporangium or megasporangium (Fig. 1a,b). We then used the corresponding complete network to calculate degree for all nodes and recorded the maximum degree node at each distance from the sporangium. After combining distance information from the backbone matrix and node degree information from the corresponding complete matrix, each reproductive structure can be represented as a line showing maximum degree for part types that are increasingly distant from the sporangium in terms of their physical attachment (Fig. 1c). To visualize patterns of node degree with distance from the sporangium or megasporangium in our entire dataset, we also added a slight amount of noise to the maximum degree values at each distance so that lines representing individual reproductive structures did not entirely overlap. This noise is a small fraction of maximum degree values and does not alter any patterns, but it does allow lines from different taxa with the same sequence of values to be
visible. These and other network analyses, including computation of summary statistics, used scripts written in Python with Networkx (45). Statistical procedures for the calculation of temporal patterns were performed using R version 4.1.3 (46). CSV files for raw data, all adjacency matrices, Python, and R scripts are available on Dryad (https://doi.org/10.5061/dryad.n02v6wx36).

Results

General patterns in reproductive networks

In aggregate, networks from more than 1450 fossil and extant reproductive structures (Figs. 2a; S1) show a proportional increase in edges with increasing nodes, suggesting that complexity as measured by part types is correlated with complexity as measured by interactions among parts. The steep slope of this relationship is due to engagement edges (fusion, conforming, envelopment, and enclosure interactions); in networks where these have been removed (the “backbone” networks) the slope is positive but much shallower and more similar to the minimum expected increase for a network of connected parts (n nodes - 1; Fig. 2a).

Maximum degree (the number of edges incident to a node) in reproductive networks also increases sharply with node number (Fig. 2b), meaning that reproductive structures with many parts often have edges concentrated on specific part types rather than spread diffusely through the network. Again, this effect is driven by engagement edges; backbone networks show a much shallower increase in maximum degree with node number (Fig. 2b).

Temporal patterns in part interactions
Both average degree and number of engagement edges increased from the Late Silurian through the Late Carboniferous (Fig. 3). This increase reflects two superimposed processes: 1) a general trend towards the evolution of compact strobili with specialized, tightly packed sporophylls in multiple lineages including lycopsids, horsetails, and progymnosperms, and 2) the evolution of seed plants at the end of the Devonian with their characteristic megasporangium-enveloping organ integument (seed coat). Among free-sporing plants, average measures of complexity after the Carboniferous remains relatively unchanged (Figs. 3, S2), although the appearance of highly specialized aquatic Salviniales ferns (47) in the late Mesozoic substantially increased maximum complexity (Fig. 3; note Cretaceous maximum increase in leftmost panels). The apparent decline in average degree over the Mesozoic (Fig. 3b), which is significant based on resampled median values for some time period comparisons, is primarily due to differences in sampling. Mesozoic intervals are dominated by leptosporangiate ferns, which generally produce simple reproductive structures, although lycopsids and horsetails must also have been present given that they survive to the present day.

In contrast to free-sporing plants, seed plants show consistent increases in the number of engagement edges through time (Fig. 3). Engagement edges in their pollen-producing structures rise over the Mesozoic but generally plateau at values similar to those of free-sporing Carboniferous lycopsids and sphenophytes (see Fig. S2, S3), reflecting the convergent evolution of a suite of conforming edges between pollen sacs and protective sporophylls (see Fig. S4). By contrast, seed-producing structures generally produce more parts and engagement edges and reach much higher values over the Mesozoic due to a more diverse set of interactions between seeds and various covering structures. In gymnosperms (here used in a broad sense to refer to all non-angiosperm seed plants) this rise is primarily caused by two different patterns: 1) an increase
in the abundance of compact seed cones and therefore conforming edges in some crown conifer
clades, and 2) a greater variety of seed-covering structures creating both conforming and
envelopment edges in various “seed fern” groups (e.g., Caytoniales, Doyleales, and Petriellales
and some crown conifer clades (Podocarpaceae, Taxaceae) with reduced, fleshy cones (Figs. 4,
S4). The median number of nodes in gymnosperm reproductive structures varies little through
time, however (Fig. S5, see also [3]), meaning that the steep slope in the relationship between
node number and engagement edges across plants generally (Fig. 2a) was not initially present in
seed plants (Fig. S6). This stronger relationship emerged only over the later Mesozoic with the
appearance of more compact and physically integrated gymnosperm reproductive organs (Figs.
3,4).

The relationship between nodes and engagement edges among modern seed plant
lineages was further strengthened by the appearance of angiosperms in the Early Cretaceous,
whose seed-producing structures (here meaning both pistillate and bisexual flowers) show
extremely high numbers of both part types and engagement edges (Figs. 3,S5; see also [3]).
Angiosperms are characterized by many fusion, envelopment, and enclosure interactions (note
the increase in these interactions in Fig. S4 coinciding with the diversification of angiosperms)
due to developmental and physical integration of seeds, individual carpels, and (in many
lineages) syncarpous gynoecia. Although some gymnosperms produce reproductive structures
that rival average angiosperms in complexity (e.g. certain Gnetales and extinct Bennettitales;
Figs. 4, S1), angiosperms reach much higher maximum values than any other group of seed
plants (Fig. 3). These unprecedented maximum values were apparently achieved rapidly by
angiosperms, and median values for the group remain similar over their history in our data (Fig.
3).
Spatial patterns in part interactions

The concentration of engagement edges within reproductive networks also shifts through time, generally moving away from the sporangium (Fig. 5). This change reflects the incorporation of various novel organs into functional relationships with sporangia and later with seeds, which themselves represent a close interaction between a megasporangium (called the nucellus in seed plants) and the integument. Among free-sporing structures, the Paleozoic rise in edges occurs primarily on the sporangium itself and then to a lesser extent on nodes several attachment edges from it (Fig. 5a), consistent with sporangia, microsporangia, or megasporangia being the locus of physical interactions with closely packed sporophylls. High-degree peaks on more distant part types do occur (Figs. 5a) but these tend to be found in highly specialized taxa such as the previously mentioned aquatic Salviniales and Late Carboniferous lycopsids that had evolved seed-like structures (e.g. Miadesmia, [48]). Pollen-producing structures show a similar overall pattern as free-sporing structures (Fig. 5a), although as noted earlier (Fig. 3a), increases in maximum degree occur later in their history when compact pollen cones become more abundant.

By contrast, the largest number of edges in seed-producing structures occur on part types other than their megasporangium/nucellus, and the concentration of these edges gradually increases and shifts towards more distant structures over time (Fig. 5a). The Mesozoic rise in engagement edges among gymnosperms is clearly visible on the seed coat (one edge from the nucellus; Fig. 5b), as these structures are the locus for physical interactions with surrounding cone scales and cupules that cover them. Angiosperm integuments are comparable to those of later gymnosperms in degree, but the evolution of the carpel creates a new locus of engagement.
edges at a greater distance from the nucellus (Fig. 5b; the exact distance varies among taxa due to differences in the number of integuments, funiculus development, and differentiation of placental tissues). Additional organs such as enveloping calyces or receptacles may also contribute to peaks at even greater distances. Our sampling within angiosperms is not detailed enough to resolve smaller-scale trends within the group, but derived clades do appear to show a shift towards increasing interactions on their carpels (Fig. 5b, right panel). In particular, early-diverging extant lineages (with the exception of some derived Nymphaeales) show lower degree carpel peaks (positions 4 and 5 in Fig. 5b, right panel) than magnoliids, monocots, and eudicots. These groups are notable for their lack of fusion among carpels, resulting in fewer engagement edges, lower degree, and generally less complex ovary organization.

Discussion

Our plant reproductive networks are independent of homology by design, but a basic understanding of phylogenetic relationships (e.g., 49-51) is nevertheless important when interpreting our results. Namely, increased engagement edges must have arisen independently among many major lineages of free-sporing and seed plants given the morphology of their early members combined with phylogenetic tree topology (Figs. 4, S2, S3). Although the phylogenetic placement of extinct seed plant lineages is not well resolved due to conflicting topologies and low support (e.g., 38, 52-56), increased engagement edges in various Mesozoic gymnosperm groups likely evolved independently given that they frequently link different specific organs (e.g., various types of bracts, fertile scales, or cupules). At minimum, engagement edges in the flowers of Bennettitales, the cupules of “seed ferns”, and the seed cones of crown conifers are almost certainly not homologous given fundamental differences in their construction (see 25).
Even within crown conifers, increased conformable interactions among cone scales (meaning their bract-scale complexes; 57, 58) must have evolved independently because interlocking scales became abundant in the fossil record only in the Jurassic (22), postdating the divergence of major extant lineages (59).

These results are consistent with some long-standing hypotheses in plant evolution; in particular, the co-option and functional integration of novel organs into reproductive structures is thought to be an important part of their morphological diversification (60, 61) and has often been invoked, either explicitly or implicitly, to interpret major patterns among seed plants. For example, the earliest seeds used an elaborate projection of the exposed megasporangium to capture pollen grains (54, 62) while later gymnosperms shifted this function to a fully enclosing integument (e.g. 63). Floral synorganization, or the fusion of carpels, stamens, and perianth elements, is a widely recognized trend among derived angiosperms (2, 4, 13, 64) and the group also commonly transfers functional roles in floral protection and display to a wide variety of auxiliary organs (61, 65). Our analyses place these observations of seed plants in a wider phylogenetic and temporal context, showing that integration, at least as characterized by engagement edges, is common across vascular plant lineages but is expressed most strongly among seed plants and especially among angiosperms.

The correlation between part types and engagement edges in our networks (Fig. 2a) suggests that morphological complexity is related in some way to the evolution of interactions among parts. For example, differentiated sporophylls in Paleozoic lycopsids and horsetail strobili, as well as Mesozoic seed plant pollen cones, occur with evidence of greater physical contact (e.g., faceted sporangia or pollen sacs and sporophyll phlanges that closely adhere to them). Seeds and carpels, whose origins represent the largest increases in part types (3), are
characterized by their high degree of integration between sporangia and various enclosing organs. These patterns at face value suggest that the evolution of reproductive complexity emerges in the context of specialized interactions among parts. But our results also show that this relationship may not be straightforward; one of the most notable temporal patterns in our data is an increase in engagement edges among Mesozoic seed plants without a change in part type numbers (Figs. S4-S6). For much of seed plant history, then, the evolution of reproductive structural complexity, at least as measured by part type diversity, was decoupled from changes in physical and functional interactions among these parts.

The gradual emergence of greater integration among seed-producing structures over the Mesozoic could be explained as a response to animal groups that diversified long after seed plant origins. For example, the Jurassic to Early Cretaceous increase in conformable edges among many conifer clades coincides with a shift towards thicker and more heavily armored seed cones (22, 66), a pattern that mirrors microevolutionary changes in extant conifer populations experiencing high predation pressure from birds and mammals (67, 68). At the same time, fleshy seed-covering organs and compact fruit-like cones that likely functioned in vertebrate seed dispersal become notable features of other conifer clades (69, 70), members of the Gnetales (71), and extinct gymnosperms like Caytonia, whose cupules have been recorded in coprolites (72). Stronger ecological interactions with vertebrates, particularly as small-bodied birds and mammals radiated from the Jurassic onwards (70, 73, 74), may have favored the evolution of more compact reproductive structures better adapted to either protect seeds or function as small edible diaspores (75, 76). These morphologies appear to be largely derived from existing morphological elements, however, rather than new suites of organs, resulting in a diverse range
of morphological strategies for covering seeds (e.g., 33, 34, 56) but little fundamental change in
structural complexity.

Specialized animal pollination syndromes are also thought to have been common over the
Mesozoic (77-79) and are associated with high reproductive part type numbers generally (3).
Insect pollination was likely important in the evolution of the most complex gymnosperm
reproductive structures; the high part type counts in the bisexual flowers of some extinct
bennettitaleans and the staminate strobili of some extant Gnetales (see Figs. 4, S3) reflect the
presence of both pollen and seed organs as well as enveloping perianth elements. But obvious
morphological adaptations to insect pollination were rare among Mesozoic gymnosperms and
many extant insect-pollinated gymnosperms like cycads likewise do not show obvious
specializations, instead relying on temperature and volatile cues to attract pollinators 80, 81).

Although pollination biology may have played an important role in the evolution of extremes
among Mesozoic gymnosperms, the general increase in engagement edges appears to have been
driven more by the evolution of compact reproductive structures that presumably performed
fruit-like functional roles in seed protection and/or dispersal.

Why gymnosperm lineages have rarely evolved highly integrated seed
protection/dispersal modules in combination with specialized perianth and staminate organs is
unclear, but may relate to their pollination syndromes. Pollinators like moths and bees, which
often interact with highly specialized perianth parts and intricate flower geometries, are thought
to have diversified with derived angiosperm clades (79, although see 82). By contrast, early-
diverging angiosperm lineages generally produce less complex flowers and are primarily visited
by ovipositing flies and beetles (83), more akin to proposed Mesozoic pollinators (77,78). These
pollination syndromes are thought to rely less on specialized floral geometries and more on cues
like odor or food rewards (84). But further exploring these potential associations and
mechanisms behind them requires a much better sampling of angiosperms and pollination
syndromes than provided in this study.

In a broader biological sense, the trend towards increasing integration within some
lineages of vascular plants is consistent with patterns in other groups (e.g., 85-89), particularly
vertebrate crania that are perhaps the most well-studied system in this regard (e.g., 90-93). It is
worth noting that this similarity occurs despite fundamental differences in how integration is
measured (see 88); for example, our engagement edges commonly represent more diffuse
generic and spatial relationships among parts (as do functional interactions in plants generally;
see 94) than size or shape covariation in cranial elements due to a shared joint. The exact
relationship between integration and complexity in plants and animals, however, appears to be
different. Fusion among vertebrate bones is often regarded as increasing integration and reducing
the number of skeletal elements (95, 96), whereas plant engagement edges most analogous to
bone fusion, such as strong interactions among the parts of seeds and fruits, are associated with
an increase in part types when they evolve. Interestingly, network-based analyses that define
complexity in terms of interactions among skeletal elements rather than their number reach a
conclusion more similar to ours, linking higher integration with increased complexity (97, 98).

The approach that we develop here represents a different way to analyze broad
macroevolutionary patterns in vascular plants, as well as a means of quantifying the evolution of
reproductive complexity with more granularity than analyses of part types alone (3). Our
approach provides a fresh window into functional integration in plants and highlights the key
role of major reproductive innovations as the origin of seeds and angiospermy are clearly
associated with increased interactions among parts. These results also demonstrate the interplay
between close physical integration and the evolution of complexity, as various organs become
incorporated into compact, modular structures that in turn become loci for new interactions,
ratcheting up structural complexity and functional integration through time. Focused studies of
individual plant clades, where the homologies of these part types and phylogenetic histories of
interactions among them are better constrained, may provide further tests of whether this
complexity ratchet hold true on smaller scales. But the flexibility of networks like those used
here means that studies of virtually any group of organisms where anatomy or morphology can
be atomized into discrete interacting parts can build on this approach.

Acknowledgements

We thank Ana Andruchow-Colombo and two anonymous reviewers for their comments and
suggestions to improve this work. We thank Rebecca Morrison and Carl Simpson for discussions
and scripts relating to the 'distance to zero' analysis, and we thank Dan McShea for helpful
comments on the manuscript.

Literature Cited

1. Endress, P. K. Evolutionary diversification of the flowers in angiosperms. Am J Bot 98, 370-
396 (2011).
flowers and diversity: a comparison of Apocynaceae and Orchidaceae. Ann Bot 117, 749-767
(2016).
3. Leslie, A. B., Simpson, C. & Mander, L. Reproductive innovations and pulsed rise in plant


**Figure Captions**

**Fig 1.** Network approach to characterizing interactions among reproductive parts. (a) Cross-section and external view of hypothetical reproductive structure with part types numbered from zero (the sporangium). (b) Networks (left) and adjacency matrices (right) based on (A), with part types depicted as numbered nodes and the five types of interactions among them as undirected, unweighted edges. Nodes interacting with themselves indicated by self-loops. Adjacency matrices represent the number of edges connecting pairs of nodes in the networks. For each reproductive structure, we scored a 'complete' network depicting all interactions among nodes.
and a 'backbone' network with only attachment interactions; red numbers indicate distance of the
nodes in attachment edges from the sporangium. (c) Plot of degree (total edges incident to a
node) in complete network from (a, b) as a function of shortest backbone distance from the
sporangium.

**Fig 2.** Basic characteristics of vascular plant reproductive networks. (a) Relationship between
node number and minimum possible number of edges (n nodes -1), backbone edges, and
complete edges across all networks. The number of 'engagement' edges (fusion, conforming,
enveloping, and enclosure interactions) in complete networks shown as a color gradient. (b)
Relationship between node number and maximum degree across all complete and backbone
networks.

**Fig 3.** Temporal trends among vascular plant reproductive networks. (a) Engagement edges
through time among free-sporing, pollen-producing, and seed-producing structures. (b) Average
degree of reproductive networks through time. In (a, b), colored bars depict maximum and
minimum values, interquartile range, and 95% confidence intervals (CI) on median values for
each geologic time bin. Confidence intervals were calculated from a bootstrap resampling
procedure (see Materials and Methods).

**Fig 4.** Engagement edges in the seed-producing structures of extinct and extant groups.
Reproductive structures from each group are plotted by their average age and colored by the
number of engagement edges. Extant groups are shown in bold and a provisional phylogeny is
provided to the left. Double branches leading to taxa indicate either paraphyletic or polyphyletic
groups. ANA grade angiosperms refers to the *Amborella*, Nymphaeales, and Austrobaileyales
lineages. Phylogenetic relationships among extinct groups are largely unresolved; overall
topology primarily reflects (50, 51, 54, 55).

**Fig 5.** Location of maximum degree nodes within reproductive structures through time. (a)
Maximum degree as a function of the number of attachment edges away from the sporangium for
all free-sporing, pollen-producing, and seed producing reproductive structures. (b) maximum
degree location within only the seed-producing structures of gymnosperms (left), all angiosperms
(center), and early-diverging angiosperm lineages (left; ANA grade plus Chloranthaceae).
Nymphaeaceae are generally resolved as sister to a grade of Cabombaceae and Hydatellaceae. In
(a, b) each line represents an individual reproductive structure color coded by age. Values at zero
distance represent the sporangium itself.